Pronouns, prepositions and probabilities:
A multivariate study of Old English word order

Rhona Jayne Alcorn

Doctor of Philosophy
The University of Edinburgh
2010/11

I, Rhona Jayne Alcorn, hereby declare that:

- the thesis has been composed by myself;

- the thesis is my own work;

- the work has not been submitted for any other degree or professional qualification.

Abstract

It is widely accepted that Old English personal pronouns often turn up in 'special' positions, i.e. positions in which functionally equivalent nominals rarely, if ever, appear. Leading theories of Old English syntax (e.g. van Kemenade 1987, Pintzuk 1991, 1996, Hulk & van Kemenade 1997, Kroch & Taylor 1997) account for the syntax of specially placed pronouns in different ways, but all treat special placement as a freely available option. Focusing on pronominal objects of prepositions in particular, this thesis shows, firstly, that current theories fail to account for the variety of special positions in which these pronouns appear and argues that at least three special positions must be recognised. The central concern of this thesis, however, is whether special placement is the freely available option that leading theories assume. Drawing on evidence from a number of descriptive studies of the syntax of pronominal objects of prepositions (e.g. Wende 1915, Taylor 2008, Alcorn 2009), statistical evidence is presented to show that, in a number of contexts, the probability of special placement is either too high or else too low to be plausibly ascribed to free variation. The thesis explores the linguistic basis of each of the statistically significant parameters identified, finding answers in some cases and intriguing puzzles in others.

Acknowledgements

Chapter 1 Foundations

1.1   Introduction

This chapter begins by explaining why the placement of personal pronouns relative to a governing preposition in Old English prose is a worthy topic for detailed research before outlining the particular goals of this thesis (section 1.2). Section 1.3 introduces some terminological conventions that will allow different types of data to be referred to in a straightforward way. Limitations of previous descriptive studies of the phenomenon of interest are described in section 1.4, where I identify a method for transforming previous findings into a powerful heuristic. The penultimate section describes four types of data that share at least one feature in common with the data of interest and explains why these data are not included within the scope of this study (section 1.5). The chapter concludes with a summary of the organisation of the remainder of the thesis (section 1.6).

1.2   Aims

In terms of placement relative to a governing preposition, three types of object can be recognised in Old English. One type has an almost completely predictable distribution to the immediate right of the preposition, and includes NPs headed by a noun, a nominalised element or a demonstrative pronoun. According to Taylor (2008: 343, fn. 1) such objects are situated immediately after the preposition in 99.9% of cases. Some examples are given in (1).[1]

---

[1] All examples are taken from the York-Toronto-Helsinki parsed corpus of Old English Prose ('YCOE') (Taylor, Warner, Pintzuk & Beths 2003) and follow the referencing conventions used by the editors of that corpus. A description of the YCOE is provided in Chapter 3. Each reference indicates: the title of the YCOE's text file, the short title used by the Dictionary of Old English where this differs from the YCOE title and information to allow the example to be located in the relevant base edition, e.g. by page and/or line number. A list of base editions used to compile the YCOE is included in the corpus documentation.

(1) a. he his eagan lythwon **fram ðære eorðan** upahof

       he his eyes   little   from the   earth   raised

       'He raised his eyes a little from the earth'

                              (comary,LS_23_[MaryofEgypt]:270.181)

     b. Hig   cwæðon eft   **to þam blindan**

       they said   again to the   blind

       'They said again to the blind one'

                              (cowsgosp,Jn_[WSCp]:9.17.6539)

     c. ac se apostol Paulus andwyrde **be**   **þysum**

       but the apostle Paul   answered concerning this

       'but Paul the apostle answered concerning this'

                          (cocathom1,ÆCHom_I,_32:456.162.6483)

Placement to the preposition's immediate right is also the rule with relative pronouns and with interrogative pronouns or phrases (Mitchell 1985: §1062), although these PPs are invariably placed clause-initially due to *wh*-movement of the preposition's object with obligatory pied piping of the preposition (Allen 1980: 268–73, 284–5, van Kemenade 1987: 149–53), as in (2).

(2) a. hæfde heo dohtor þa   Ercongotan, **bi ðære** we nu   syndon sprecende

       had   she daughter then Ercongote   of whom we now are   speaking

       'she had a daughter then, Ercongote, of whom we are now speaking'

                            (cobede,Bede_3:6.172.23.1695)

     b. **Mid hwylcum fotum** gæst þu   on Godes halgan flore

       with which   feet   walk you on God's holy   floor

       'With which feet do you walk on God's holy floor?'

                            (coaelhom,ÆHom_27:63.3966)

The second type of object has a completely predictable distribution to the preposition's left. This type consists of just two elements: *þær* 'there' and *her* 'here', which are variously referred to as R-pronouns (e.g. Vat 1978, van Kemenade 1987: 109), locative pronouns (e.g. Allen 1980: 291, van Kemenade 1987: 108) and (locative) adverbs (e.g. Clark Hall 1960, Mitchell 1985: §1155).

These elements are sometimes found immediately to the left of a governing preposition, e.g. (3a) and sometimes further to its left, e.g. (3b).

(3) a.   and tigdon hine **þærto** mid heardum bendum
         and tied   him thereto with hard   bonds
         'and tied him thereto with hard bonds'

                                        (coaelive,ÆLS_[Edmund]:106.7027)

    b.   Þa   eodon **þær** ma   manna **to**
         then went   there more men   to
         'Then more men went thereto'

                                        (cogregdC,GD_2_[C]:9.123.2.1480)

The third type of object consists of simple i.e. unmodified and uncoordinated, personal pronouns. These pronouns often occur to the immediate right of the preposition, as in (4), but they also occur to the preposition's left: sometimes immediately to its left, as in (5a), and sometimes with one or more intervening constituents, as in (5b).

(4) a.   Þa   his gebroþru **to him** comon
         when his brethren to him came
         'When his brethren came to him'

                                        (cocathom1,ÆCHom_I,_21:346.24.412)
    b.   God cwæð **to him**
         God said   to him
         'God said to him'

                                        (cocathom1,ÆCHom_I,_1:181.75.72)

(5) a.   ... oððæt se halga gast   **him to** com
              until   the holy spirit him to came
         '… until the holy spirit came to him'

                                        (cocathom1,ÆCHom_I,_21:346.24.4121)

b. Þæt wif   **him** cwæð þa   **to**

   the woman him said   then to

   'the woman then said to him'

<div align="right">(coaelhom,ÆHom_5:21.690)</div>

By comparing (4a) with (5a) and (4b) with (5b), one can see there is no apparent correlation between pronoun position and PP semantics (Wende 1915: 70–1, Mitchell 1978: §7). Such evidence has lead to the generally accepted assumption that pronoun placement in Old English is a syntactic variable.

The option of pronoun placement to the preposition's left rather than its right is mentioned in many Old English grammars (e.g. Quirk & Wrenn 1957: §141, Visser 1970: §402, Mitchell 1985: §1062), and even in some introductory textbooks (e.g. Mitchell & Robinson 1992: §213, Hogg 2002: 93–4), but even the grammars provide no more than a few lines of discussion. Leading theories of Old English syntax (e.g. van Kemenade 1987, Pintzuk 1991, 1996, Kroch & Taylor 1997) aim to account for the *variety* of positions in which personal pronoun objects occur — more or less successfully for those governed by prepositions as we will see in the next chapter — but treat placement of pronouns to the preposition's left as a freely available option.

Alongside the theoretical accounts are a number of studies that examine the placement of pronominal objects of prepositions in Old English very closely. These studies, summarised in Chapter 3, identify a number of factors that appear to correlate statistically (in some cases) or absolutely (in others) with pronoun position. These findings suggest that pronoun placement is not the freely available option that theory predicts, and suggest instead there may be some degree of structure to the variation. To date, however, there has been little attempt to make sense of these findings or to investigate if and how correlating factors identified relate to one another.

The present study investigates sixteen independent variables that have been associated with the positioning of simple personal pronouns relative to a governing preposition in Old English prose. The study reveals a number of contexts in which pronoun placement can be accurately predicted at least 95% of the time, and identifies which of the other variables correlate statistically with

pronoun placement in what appears to be a linguistically meaningful way. For each of the categorical and probabilistic variables identified, the thesis attempts to make sense of the patterns and trends uncovered.

1.3 Terminology

From this point onward, I adopt a few terminological conventions for ease of reference. Firstly, I will refer to prepositions as 'prepositions' regardless of where their object is situated. This allows me to avoid terms variously used to indicate prepositions with a specially placed object, as in (5), including: 'postposition' (Fakundiny 1970: 139, fn. 1); 'preposition in post-position' (Mitchell 1978: §48, 1985: §1062); 'adposition in post-position' (Colman 1991: 56); 'post-adposition' (Miranda-García & Calle-Martín 2010: 93); 'postpositioned preposition' (Lapidge 2006: 154); and, in a different vein, 'inverted PP' (Allen 1980: 288). Instead, I use the terms 'left-of-P' and 'right-of-P' to indicate the position of the object. Left-of-P objects include those situated immediately to the preposition's left, as in (5a), as well as those which are somewhere further to its left, e.g. (5b). Secondly, it will be helpful to have shorthand terms to refer to the three types of object identified in the previous section. For the simple personal pronouns, I use the abbreviation 'PPOP' (**p**ersonal **p**ronoun functioning as the **o**bject of a **p**reposition'): this distinguishes them from other types of (pro)nominal objects of prepositions and from personal pronouns in other functions. The term 'nominal objects' will be used to denote the type of object with an (almost) completely predictable distribution to the immediate right of the preposition, including nominals, demonstrative pronouns, modified personal pronouns and coordinated personal pronouns, and for *þær* and *her* I employ the commonly used term 'R-pronoun'.

1.4 Evaluating existing evidence

Although many factors have been found to correlate with a greater, or in some cases lesser, frequency of left-of-P placement of PPOPs in Old English prose, the relevance of these observations for a linguistic analysis of this variability is not at all clear. There are several reasons for this. Firstly, different studies have used different sets of texts. The extant prose provides evidence of written English

produced in different parts of England over a period of more than 300 years, and it is presently unknown whether what was observed in one sample is representative of what would be found in another. Secondly, there are differences in the amount of descriptive detail provided. Some observed correlations are backed up by quantitative data, but others are described only in very general terms, e.g. '[w]ith pronouns, the prepositions (especially those of more than one syllable) quite frequently follow' (Quirk & Wrenn 1957: §141). This observation could lead us to expect PPOPs governed by *æfter* 'after' to appear more frequently in a left-of-P position that those governed by *to* 'to'. In Chapter 3 we will see that this is not so. Thirdly, no study has yet sought to estimate the independent effect of these factors, i.e. the effect of one when the effects of others are held constant, on PPOP placement. Taylor (2008) estimates the independent effects of some independent variables but not all, so her results cannot be generalised beyond her model, nor beyond her sample of data. For all of these reasons, we simply do not know which observations ought to be taken seriously.

With respect to the first problem, i.e. that of inconsistent materials, the York-Toronto-Helsinki parsed corpus of Old English Prose ('YCOE') (Taylor, Warner, Pintzuk & Beths 2003) provides the 21st-century scholar with direct access to syntactically-annotated versions of scholarly editions of some 100 Old English texts of varying lengths, dates, dialects and genres. As a 'superset' of most if not all of the texts from which previous observations have been made, the YCOE allows the results of smaller scale studies to be checked on an unprecedented scale. The YCOE is a wonderful resource for word order studies in particular and makes an ideal choice of corpus for the present study.

The second and third problems, i.e. those concerned with methods of measurement, are far from unique and many studies have shown how sophisticated statistical software can be exploited to quantify independent and combined effects of multiple variables in order to produce a comprehensive description of linguistic phenomena. Hinrichs & Szmrecsanyi (2007) (henceforth 'H&S'), for example, itemise a host of variables variously hypothesised to influence the choice between the *s*-genitive, as in *Gordon Brown's political career*, and the *of*-genitive, as in *the political career of Gordon Brown*, in Present Day English. Although H&S had no reason to doubt the fact of a correlation

between any of these variables and genitive choice, little was known about how their effects correlated with each other. This lack of clarity had, in turn, allowed two competing views to emerge about the significance of the 'animacy effect', i.e. the observation that the *s*-variant is preferred to the *of*-variant when the possessor is more animate than the possessum. According to one view, this effect is epiphenomenal to the 'end-weight effect', i.e. the observation that the *s*-variant is preferred when the possessum phrase is more complex, or 'heavier', than the possessor phrase, and that the *of*-variant is favoured when the possessor phrase is the heavier of the two (e.g. Hawkins 1994: 424). According to another view, genitive choice exhibits animacy as well as end-weight effects (e.g. Rosenbach 2005). The difference between these two views is important: in Rosenbach's view, semantics plays a central role in determining choice of genitive, whereas in Hawkins' view the animacy principle can be reduced to a processing constraint. In order to bring clarity to this debate, H&S modelled the choice of variant in newspaper prose and included all previous identified variables as independent variables in their model. As predicted by Rosenbach (*ibid*), H&S found a significant effect of end-weight <u>and</u> of animacy, showing that choice of genitive is indeed partly conditioned by semantics. Moreover, the results of this multivariate analysis additionally allowed variables to be ranked according to their relative importance in genitive choice, a considerable enhancement to previous descriptions of the phenomenon.

Hoffmann (2005a) takes a similar approach to this same problem of disconnected observations in order to evaluate the role of a number of variables purported to influence the choice between preposition stranding and pied piping in Present Day English *wh*-relative clauses. Drawing data from the British English component of the International Corpus of English, Hoffman estimates the independent effect of each variable previously identified and his results, like those of H&S, provide the field with three types of new evidence: confirmatory, e.g. Hoffman found the probability of pied piping is indeed much greater in formal contexts than in informal contexts, as was widely supposed; disconfirmatory, e.g. contra Johansson & Geisler (1998: 76), Hoffmann found no independent effect of clause complexity; and clarifactory, *viz.* the ranking of variables according to relative importance in variant choice.

H&S (2007), Hoffmann (2005a) and other studies in this vein, e.g. de Sutter (2009), demonstrate most effectively how a collection of unconnected observations and assumptions about a linguistic variable can be transformed, firstly, into a more comprehensive description of the data and, secondly, into a powerful heuristic. Such an approach seems ideal for exploring previously reported correlations involving PPOP placement in Old English.

1.5 Excluded data

1.5.1 Modified pronouns and coordinated pronouns

The regular placement of modified personal pronouns and coordinated personal pronouns to the right of a governing preposition in Old English has been noted on several occasions (e.g. Wende 1915: 64–6, Koopman 1992: 61, 1997a: 87, 2005: 50–1). Taylor (2008: 360) reports that pronouns modified by *self* 'self' always follow the preposition in her sample of nineteen Old English texts, but is silent about coordinated personal pronouns and those modified by elements other than *self*. Wende's examples indicate that right-of-P placement is the regular rule when the pronoun is modified by an adjacent adjective, as in (6a), a (possibly non-adjacent) relative clause, e.g. (6b), and when the pronoun is coordinated, e.g. (7).

(6)   a. **to him anum** we scolun us gebiddan

  to him alone we should us pray

  'we should pray to him alone'

  (cocathom1,ÆCHom_I,_11:270.116.2084)

  b. and **on him** ic gelefa, **þe   is fæder and sunu and halig gast**

  and in him I believe that is father and son   and holy ghost

  'and I believe in him, who is the father and son and holy ghost'

  (comargaC,LS_14_[MargaretCCCC_303]:6.8.78)

(7)   a. ... þæt ge **mid him and his halgum** þæt ece   lif habban moton

   that you with him and his saints   the eternal life have   may

  '… that you, with him and his saints, may have eternal life'

  (cocathom2,ÆCHom_II,_6:59.199.1184)

b. **betwux us &   eow** is gefæstnod micel þrosm

between us and you is fixed   great vapour

'between us and you is fixed a great vapour'

(cocathom1,ÆCHom_I,_23:368.84.4596)

Drawing data from the YCOE (Taylor *et al*. 2003), which I describe in greater detail in Chapter 3, I identified 34 coordinated and 772 modified personal pronouns functioning as the object of a preposition. All of the coordinated pronouns and all but four (0.5%) of the modified pronouns are situated to the preposition's right. The fact that, in terms of their syntax, modified and/or coordinated personal pronouns pattern with nouns and demonstrative pronouns rather than with unmodified and uncoordinated personal pronouns in Old English would come as no surprise to pronoun typologists. Cardinaletti & Starke (1996, 1999), for example, show that the syntax of modified and/or coordinated personal pronouns is exactly like that of functionally equivalent full noun phrases in many different languages. As modified and/or coordinated PPOPs do not exhibit the type of variation with which this thesis is concerned, all 806 examples found in the YCOE are excluded from the present study.

As said, the YCOE provides four examples involving a modified left-of-P pronoun. According to the YCOE editors' parse of the first example, (8), the emboldened initial NP is the object of the emboldened instance of *to*.

(8)   **Him þa   gyta ferendum** sume dæge on Grecalande wæs **to** broht   to

him then yet travelling some day   in Greece   was to brought to

lacnianne an man

treating   a   man

'one day in Greece, while still travelling, a man was brought to him for treatment'

(cogregdC,GDPref_and_3_[C]:3.183.24.2268)

In several places in this thesis I present a number of PPs extracted from the YCOE that exhibit some or other unexpected feature but that may not actually involve a PP at all. For some of these PPs, I argue that the word parsed as a preposition by

the corpus editors could instead be interpreted as a verbal prefix or particle, and the element parsed as the preposition's object could instead be interpreted as the object of the complex verb. Example (8) is the first of these examples. Many Old English verbs consist of two elements: the verb itself and an element referred to as a prefix (where inseparable from the verb) or particle (where separable) (e.g. Elenbaas 2006: 105–74). Many of the Old English prefixes and particles are identical in form to a preposition, which can make it difficult to decide whether a particular clause contains a complex verb and its object or a simplex verb plus a PP (e.g. Mitchell 1978). The YCOE's editors' approach to this problem is addressed in Chapter 3. Returning to the example at (8), the fact that three authoritative Old English dictionaries — Bosworth & Toller (1898), Clark Hall (1960) and the *Dictionary of Old English* (Cameron, Amos, Healey *et al*. 2007) (the '*DOE*') — list *tobringan* 'to bring to', either in its own right or as a derivative of *bringan* 'to bring', provides support for interpreting the emboldened initial NP as the object of a complex verb. In further support of this interpretation it may be noted that in the early eleventh-century *Cambridge Psalter*, Latin *adtulit eis* 'gave them' is glossed interlinearly as *tobrohte him* (Ogura 1992: 377, fn. 3). As right-of-P objects always occur immediately to the preposition's right, *him* cannot be the object of *to* in this gloss.

The second example could be explained by similar means. This example, (9a), is noted by Wende (1915: 64, fn. 1), who suggests we may be dealing with *tosprecan*. *Tosprecan* is not listed in Bosworth & Toller, Clark Hall or the *DOE*, but the example at (9b), from the same text as (9a), supports Wende's suggestion. In (9b), the nominal object — *Sigeberhte þam cyninge* — is more likely to be the object of *tosprecan* than of prepositional *to* given the position of the nominal relative to *to*, and this is also how this nominal is parsed in the YCOE.[2]

---

[2] The relevance — or rather the irrelevance — of word division for distinguishing between free and bound morphemes in Old English is discussed in Chapter 3.

(9) a. þa cleopode he ðriga & aane from þæm gehalgedum fæmnum Criste
       then called he thrice and one of the consecrated virgins Christ
       hire agne noman cegde, swa swa he **hire ondweardre to** spræce
       her own name called as if he her present   to speak
       'he then called out thrice and summoned one of the consecrated virgins of
       Christ by her own name, as if he was speaking to her in person'

                                        (cobede,Bede_4:9.286.1.2879)
     b. Ðas word & eac monig þysses gemetes mid þy Osw1o se cyning
        these words and also many of-this manner with this Oswio the king
        Sigeberhte þam cyninge mid freondlicre & mid broðorlicre geþeahte
        Sigeberht the king   with friendly   and with fraternal advice
        oft   & gelome **to**spræc, þa æt nehstan mid fultume & mid
        often and repeatedly conveyed then at last   with support and with
        geþafunge his   freonda þæt he gelyfde
        consent of-his friends it he believed
        'Once king Oswio had conveyed these words and also many of this
        manner often and repeatedly to king Sigeberht with friendly and fraternal
        advice, then, at last, with the support and the consent of his friends, he
        believed it'

                                        (cobede,Bede_3:16.224.26.2305)

Wende (*ibid*) also records the example at (10) and accepts it as a clear exception
to the otherwise regular right-of-P placement of modified PPOPs. The last
example, from a text not included in Wende's study, is given at (11). This too
appears to be an exception.

(10)   ... gif hie gemunan willað hiora ieldrena   unclænnessa, & heora
          if they consider wish their ancestors' impurities   and their
        wolgewinna,   & hiora monigfealdan unsibbe,   & hiora
        calamitous-wars and their manifold   dissensions and their
        unmiltsunge þe hie to Gode hæfdon, ge eac **him selfum betweonum**
        impiety   that they to God bear   and also them selves between

                                                                16

'… if they wish to consider the impurities of their ancestors and their calamitous wars and their manifold dissensions and their impiety which they bear to God, and also between themselves'

<div align="right">(coorosiu,Or_2:1.38.25.741)</div>

(11)　　&　he swa sona þone sceoccan adræfde of　þam earman menn, **heom**
　　　　and he thus soon the　demon　drove　from the　poor　man,　them
　　　　eallum tomiddes
　　　　all　　towards
　　　　'and he thus soon drove the demon from the poor man towards them all'

<div align="right">(coaelhom,ÆHom_18:293.2653)</div>

1.5.2　'Understood' objects

Like PPOPs, the second type of object not considered in this thesis does not regularly appear in the right-of-P position. Unlike PPOPs, however, this type of object does not appear in <u>any</u> position: rather it is absent, but interpreted by reference to an element in a higher clause. Three sub-types can be identified.

The first type of 'understood' object appears in what is usually referred to as the *tough* or *easy-to-please* construction. The use of this construction in Old English is discussed in Mitchell (1985: §§928–31), and is characterised in Fischer, van Kemenade, Koopman & van der Wurff (2000: 256–7) as a construction that 'consists of a subject followed by a predicate formed by an adjective plus an infinitival clause with a non-subject gap. The interpretation of this gap is provided by the subject.' An example in which the non-subject gap functions as the object of a preposition is given in (12). In this example, *on* occurs without an overt object in an infinitival clause, and the interpretation of its object is provided by *heo*, the subject of the subordinating clause.

(12)　　... þeah　　heo　　ær　　　　gladu　　wære **on** to locienne
　　　　although it-NOM previously pleasant was　on　to look
　　　　'… although it was previously pleasant to look upon'

<div align="right">(coboeth,Bo:6.14.12.214)</div>

There already exists an extensive literature on the theory of the *tough* construction, including its manifestation in Old English (e.g. Allen 1980: 283–4, van Kemenade 1987: 152, 163–70, van der Wurff 1987, 1990) to which the interested reader is referred.

The second type of 'understood' object occurs in relative clauses. Earlier in this chapter we saw that when there is overt *wh*-movement of a preposition's object, we also find pied piping of the preposition, e.g. (2). When the preposition's *wh*-object is covert, however, the preposition is invariably stranded. Relativised prepositional objects are always covert in infinitival clauses, so in these clauses we invariably find preposition stranding rather than pied piping (Allen 1980: 272–6, van Kemenade 1987: 151), e.g. (13).

(13)   þeah   he nu nanwuht ellese næbbe   **ymbe** to sorgienne
       though he now nothing else   not-have about to worry
       'though he now should have nothing else to worry about'

                                                     (coboeth,Bo:11.24.15.410)

In finite relatives, relativised objects of prepositions are sometimes covert and sometimes overt. An example without a relative pronoun is given in (14), in which the preposition is predictably stranded. Note that the initial element of the finite relative in (14), i.e. *þe*, is not a relative pronoun: it is an indeclinable relative complementiser (Allen 1980: 266–8, van Kemenade 1987: 147–8).[3]

(14)   Þa   eode ut in dagunge of   þam huse, þe ða untruman menn **in**
       then went out in daybreak from the house that the sick   men   in
       reston
       rested
       'Then at daybreak (he) went out of the house, which the sick men slept in'

                                                     (cobede,Bede_3:19.242.23.2483)

---

[3] *Þe* is the usual form of the relative complementiser but it is sometimes realised as indeclinable *þæt*.

The history of preposition stranding in relative clauses has also been studied in detail. For Old English studies in particular, see e.g. Allen (1980) and van Kemenade (1987: 144–72).

The third type of 'understood' object is illustrated by the examples in (15).

(15) a. &   niðer **mid** geweotan in   midde   ða niolnesse ðæs   byrnendan leges
         and down with went   into middle the abyss   of-the burning   fire
         'and went down with (them) into the abyss of burning fire'

         (cobede,Bede_5:13.428.3.4299)

    b. Eft   wið   þon ilcan celeþonian seaw &   sæwæter, smire **mid** þa eagan
         also against the same celandine   juice and seawater anoint with the eyes
         'Likewise for the same (ailment): celandine juice and seawater. Anoint the
         eyes with (it)'

         (colaece,Lch_II_[1]:2.3.1.182)

This type of example is strongly reminiscent of null object constructions, which are commonly found in the language of modern recipes, e.g. (16) (the null objects are indicated by Ø), although objectless prepositions are now generally ungrammatical, e.g. (17).[4]

(16)   Take the cake mix, 1 cup of water, and 3 eggs. Mix Ø well and beat Ø for
        5 minutes. Pour Ø into a well-greased cake pan and bake Ø for 20
        minutes. Remove Ø from oven and cool Ø.

        (Massam & Roberge 1989: 135, ex. 2)

(17)   Mix the lemon juice and chopped parsley. *Then sprinkle scallops with Ø

        (Massam & Roberge 1989: 136, ex. 12b)

I have found just one mention of the type of objectless preposition illustrated by the examples at (15) in the Old English literature although examples are far from uncommon. There are almost 400 in the YCOE. The single mention

---

[4] Some present day non-standard varieties do allow elliptical *with* (Adams 1997), as in *I'm going to the cinema. Do you want to come with?*

occurs in Visser (1970: §626), who claims it occurs frequently in Old English charms and rarely elsewhere. The text with by far the largest share of examples in the YCOE is *Bald's Leechbook* (colaece, N=100), a collection of medicinal recipes, but the others are spread across a range of genres, e.g. there are nineteen examples in the *Peterborough Chronicle* (cochronE), eleven in *Lives of Saints* (coaelive) and six in the *Heptateuch* (cootest). *Mid* 'with' is by far the most commonly involved preposition, accounting for about 50% of the examples in the YCOE, e.g. (15). *To* 'to' is the second most commonly involved preposition, accounting for another 25%, e.g. (18). *On* 'in, on' is the only other preposition to provide more than a handful of examples, e.g. (19).

(18)   and hi   ferdon   ða **to**
         and they travelled then to
         'and they travelled then to (that place)'

                                              (coaelive,ÆLS_[Maccabees]:378.5083)

(19)   Þa   se halga wer nam þæt hors   &   healfne þone weg **on** geferde.
         then the holy man took the horse and half   the   way on travelled
         'Then the holy man took the horse and travelled half the way upon (it)'

                                              (cogregdC,GD_1_[C]:10.78.24.877–8)

The correct description and analysis of the type of the objectless prepositions illustrated in (15) will not be settled here, although both are interesting questions for future research. One possibility is that these objectless prepositions — or at least some of them — are not prepositions at all. For example, adverbial *mid* 'together' (Clark Hall 1960) would work for (15a) ('and (they) went down together ...'), although not for (15b).

1.5.3   R-pronouns

Old English *þær* 'there' and *her* 'here' are difficult words to classify, with the literature divided about whether they function as adverbs, as locative pronouns or, indeed, as both. Their status as adverbs is assumed, for example, by Wende (1915: 23–35), Clark Hall (1960) and Pintzuk (1991: 187–286). This is also the position

of Bosworth & Toller (1898) and Mitchell (1985: §1155, fn. 267), although both recognise the pronominal force of *þær* when governed by a preposition. Those espousing the pronominal view include van Kemenade (1987: 108–41) and Allen (1980: 291–7), although Allen considers *þær* only when governed by a preposition. Vat (1978) takes a different stance altogether and argues that *þær* has a 'double role' (*ibid*: 702): as pronoun when governed by a preposition, otherwise as locative pro-PP.

Despite differing views on the categorial status of *þær* and *her*, it is generally accepted that *þær* shows evidence of the same sort of special syntax as exhibited by specially placed personal pronouns in some, if not all, of its uses although, as van Bergen (2003: 144) points out, this has yet to be systematically tested. In addition, the following chapter shows that the syntax of *daar*, one of the Modern Dutch cognates of Old English *þær*, demonstrates that 'there' can exhibit a special syntax in one respect but a non-special syntax in another. As the syntax of Old English 'here' and 'there' is presently rather poorly understood, I make no assumptions about the significance of what appear to be similarities between the syntax of R-pronouns when governed by a preposition and the syntax of left-of-P PPOPs.

1.5.4   PPOPs in the poetry

Many syntactic differences between the language of Old English prose and the language of Old English poetry have been identified (e.g. Blockley 2001, Carlton 1963: 778, Gneuss 1991: 49, van Kemenade 2002), but there is no consensus about how these differences should be interpreted. Some argue that at least some differences reflect change over time, with the language of poetry assumed either to be particularly conservative (e.g. Lehmann 1992: 240) or else of an earlier date than the extant prose. Pintzuk (1991), for example, compares word order in the language of *Beowulf* to that of late Old English prose and interprets differences as evidence as diachronic variation. She acknowledges (*ibid*: 192), however, that not everyone would agree that the language of *Beowulf* is representative of a stage of Old English for which little prose is extant. Others argue that the prose and poetry should be treated as two separate languages, not least because the poetry conforms to a particular metrical structure that dictates the basic rhythm of the text. Such is

the position of van Kemenade (1987: 4), although in later work on the history of negation van Kemenade (1997a), like Pintzuk (1991), takes *Beowulf* as her earliest source of evidence.

The metrical structure to which much of the Old English poetry conforms is an important consideration for word order studies. Words and phrases have to be integrated into this structure, and this imposes particular constraints that do not apply to the prose, potentially leading to word order choices that might have been regarded as marked in other contexts. Leaving Old English to one side for the moment, consider the opening verses of the original non-metrical version of Psalm 23:

(20)   The Lord is my shepherd; I shall not want. He maketh me to lie down in
        green pastures: he leadeth me **beside the still waters**.

                                        (Psalm 23:1–2, King James Bible)

When these verses were adapted to ballad metre — four lines with an a-b-x-b rhyming pattern, the first and third having four beats, the second and fourth having three — in the mid-17th century (Patrick 1949: 103), the emboldened PP underwent both lexical and syntactic transformation:

(21)   The Lord's my shepherd, I'll not want.
        He maketh me down to lie
        In pastures green: he leadeth me
        The quiet waters by.

                                        (Psalm 23:1–2, Scottish Metrical Psalter of 1650)

There is good reason to suppose that metre plays an important role in determining word order, including the position of objects relative to a governing preposition, in the Old English poetry. In his detailed discussion of the regulation of stress in early Germanic poetry, Kuhn (1933) draws a distinction between *Satzpartikeln*, i.e. words which are normally stressed but which can be de-stressed when placed before the first stressed position of the clause, and *Satzteilpartikeln*, i.e. words which form a syntactic phrase with a following word. *Satzteilpartikeln*,

including prepositions, are normally unstressed when their phrase-mate is *in situ*, but they can acquire stress when their phrase-mate is elsewhere. Given the need to conform to a particular metrical structure, one can appreciate how useful it would be to have freedom to position an object before rather than after a governing preposition when composing verse. Thanks to Lapidge (2006), we have ready access to an exhaustive list of all lines containing a preposition with a left-of-P object in the extant Old English verse. Lapidge's metrical analysis of these lines reveals that in each case the preposition does indeed carry metrical stress (*ibid*: 174), which brings a different perspective to Pintzuk's (1991: 193) analysis of left-of-P placement of PPOPs in *Beowulf*. While Pintzuk shows that these pronouns are invariably unstressed, it is not apparent to those unfamiliar with Old English meter that their left-of-P placement invariably goes hand-in-hand with the assignment of stress to the preposition. Lapidge's focus on the prosodic properties of the prepositions rather than their left-of-P objects reveals the possibility that objects may sometimes be positioned to the left of a governing preposition to allow their governor to carry a stressed syllable. Lapidge additionally notes that in 49% of examples occurring in a first half-line and 34% of examples occurring in a second half-line, the stressed preposition participates in the line's alliteration. Since a word cannot participate in alliteration unless it carries one of the two stressed syllables of its half-line, we may infer this to be an additional motivation for placing objects before a governing preposition in the poetry. An example from *Beowulf* is given in (22). Note that the preposition's object is positioned to its left, allowing the preposition, *mid* — the first accented syllable of the second half-line — to alliterate with the initial, accented syllables of *madma* and *mænigo* in the first half-line.

(22)   madma   mænigo,   þa **him mid** scoldon …
         of-treasures many   that him with must
         'many treasures, which were to go with him'

(*Beo* 41)

Pintzuk (2002a) reports the results of a similar study to that of Lapidge (2006) and shows that in more than 40% of cases involving left-of-P placement of a nominal object, right-of-P placement would still yield a metrical half-line. This

indicates that placement of objects relative to a governing preposition is not wholly due to metrical constraints. However, as Pintzuk (*ibid*) provides no comparative data for PPOPs and as Lapidge (2006) does not consider whether right- rather than left-of-P placement would violate the meter, further work would be needed to determine how often left-of-P placement of PPOPs in particular is for reasons of meter overall.

Another important feature of Lapidge's data, although one he does not draw attention to, is that 33% (N=85) of objects situated to the left of a governing preposition are nominal.[5] This is a substantially larger proportion than may be gauged for the prose. According to Taylor (2008: 343, fn. 1), the number of nominal left-of-P objects in the whole of the YCOE is about 100, which represents less than 0.1% of all that occur in that corpus. It is especially striking that the number of nominal objects in a left-of-P position in Lapidge's poetry corpus (Bessinger 1978) is not far short of the number in the YCOE given that the word count of the latter (*c*. 1.5 million words) is approximately 8.5 times that of the former (*c*. 175,000 words).

Lapidge (*ibid*: 179) tentatively suggests that left-of-P placement of objects in the Old English prose could be explained as 'a poeticism adopted from the practice of Old English poets', but this seems unlikely for several reasons. Firstly, if it were a poeticism, there would be no obvious reason for it to be restricted almost entirely to simple personal pronouns. Secondly, left-of-P placement is far too widespread to be regarded as a poeticism: it occurs in all but nine of the ninety-one YCOE text files that have at least one PPOP, and of the nine that supply right-of-P PPOPs only, just two supply more than ten examples in total: colawafint (*Alfred's Introduction to Laws*) x17; and cobyrhtf (*Byrhtferth's Manual*) x11. Thirdly, at least some of the eighty-two texts with at least one left-of-P PPOP are not the sort one might generally associate with poetic expression, e.g. the four versions of *Anglo-Saxon Chronicles* — in which the frequency of left-of-P placement of PPOPs ranges from 33% (*Peterborough Chronicle*) to 55% (*Parker Chronicle*) — and medical texts such as *Bald's Leechbook* (colaece, 71% left-of-P). Fourthly, with the help of Professor Bremmer of Leiden University

---

[5] I discount the 16 stranded prepositions that Lapidge classifies as governors of clause-initial *þe*. As shown by Allen (1980), *þe* functions as a relative complementiser, not as a relative pronoun.

(personal communication), I have found evidence that suggests placement of simple personal pronouns to the left of a governing preposition occurs in Old Frisian prose also. The three examples, provided by Professor Bremmer, are given below together with his glosses and translations. The first example comes from the First Hunsingo Ms., which dates to *c*.1325-50, and the second is from the Fivelgo Ms., dated between 1427 and *c*.1450 (Bremmer 2009: 13). Prof. Bremmer cautions, however, that in these examples, *to* could possibly be a separable prefix of *tospreka* 'to address'.

(23) Ik sprek **iu** **to** fon tha liudum end fon tha frana.
I speak you-DAT-SG to of the people and of the *frana*
'I accuse you on behalf of the people and of the *frana* (a kind of legal official)'

(H1 XX.1 [Hoekstra 1950: 166])

(24) and sprecht ma **him** **to** thet hi alle ewela deda den hebbe.
and speaks one him-DAT-SG to that he all evil deeds done have
'and if they accuse him of having done all kinds of evil deeds'

(F IV.20 [Sjölin 1970: 258])

The third example, (25), is better since it does not have this parsing ambiguity. This example can be directly compared with that at (26), since both are taken from near-contemporary (*ca*. 1300) redactions of *Brocmonna Bref*, a text consisting of detailed legal regulations (Bremmer 2009: 11). Note that these two examples vary only by pronoun position and spelling.

(25) and stonde hia **him** naut **mith**, sa resze hi fiuwer merc tha
and stand they him-DAT-SG not with, so reach he four marks the
liuden and thene tichtega vpriuchte hi.
people and the accusation pay-damage-for he
'and if they do not support him, then he should pay four marks to the people and pay damages for the accusation'

(B 1:76 [Buma 1949: 49-50])

25

(26) and stonde hia   nawet **mith him**,   sa resze hi fiuwer hagesta

and stand   they not   with him-DAT-SG, so reach he four   highest

merc   tha liudum and thene tichtega   vpriuchte   hi.

marks the people and the   accusation pay-damage-for he

'and if they do not support him, then he should pay four highest marks to

the people and pay damages for the accusation'

<div align="right">(B 2:76 [Buma 1949: 49-50])</div>

This minimal pair suggests that left-of-P placement of simple personal pronouns was a grammatical alternative to right-of-P placement in Old Frisian prose and, further, that the pronoun's position does not alter the PP's semantics.

In summary, the widespread attestation of left-of-P PPOPs in the YCOE and the identification of a probable example from an Old Frisian legal text suggests that left-of-P placement in Old English prose is not a feature borrowed from the poetry, but rather reflects the tail-end of a transition from a West Germanic post-positional structure to the predominantly pre-positional structure found in all modern Germanic varieties, including Present Day English (e.g. Vat 1978: 704, fn. 8, Lehmann 1992: 240–2, Algeo 2010: 69). Evidence of left-of-P placement in Old Saxon verse is presented in Wende (1915) and in Kuhn (1933) and the latter source provides evidence of left-of-P placement in Old Icelandic verse also. To the best of my knowledge, however, this is the first time that evidence of left-of-P placement of PPOPs has been presented for a variety of Germanic prose other than early English.

With respect to the objectives of the present study, my analysis of data in Lapidge (2006) indicates that placement of all types of objects relative to a governing preposition in Old English poetry is influenced by the prescribed metrical system and system of alliteration, i.e. factors specific to that genre. Accordingly, I focus solely on evidence from the prose.

1.6   Structure of thesis

The thesis is organised into six further chapters. Chapter 2 considers two aspects of the theory of PPOP placement in Old English: firstly, how left-of-P placement is accounted for and, secondly, whether all left-of-P PPOPs have the same status

in the syntax. As well as highlighting some shortcomings of existing theories, Chapter 2 provides the rationale for the way in which the dependent variable, i.e. PPOP position, is represented in the statistical analyses presented throughout this thesis. Chapter 3 introduces the source of my data on PPOP placement in Old English prose and considers the reliability of its component syntactic annotations on which I depend. Chapter 3 also introduces and contextualises each of the independent variables that define the dimensions of the statistical model used to investigate structure in variable PPOP placement and quantifies their effects in terms of raw frequencies. This reveals a number of 'knockout' (or near knockout) factors (Young & Bayley 1996: 273), i.e. factors that, when present, correlate with right-of-P placement more than 95% of the time or less than 5% of the time. These knockout factors are explored in greater detail in Chapter 4, where I consider why the syntax of PPOPs exhibits little variation, if any, in their presence. Chapters 5 and 6 deal with non-knockout factors and quantifies the relationship between each factor and PPOP placement when all of the other relationships are controlled for. Chapter 5 begins with some background to the generation and interpretation of these results before addressing those that do not appear to require a linguistic explanation or else do not seem to readily admit one. Chapter 6 then deals with those results that do appear to be linguistically meaningful. The main findings of the thesis are summarised in Chapter 7.

Chapter 2 Theoretical foundations

2.1   Introduction

This chapter considers two aspects of the theory of PPOP placement in Old English. Section 2.2 surveys treatments of the special placement of personal pronouns and compares predictions for the range of special positions in which PPOPs can appear against empirical evidence. As well as highlighting some theoretical shortcomings, this section provides the rationale for the way in which the dependent variable, i.e. PPOP position, is represented in the statistical analyses presented throughout the thesis. Section 2.3 considers whether there is more than one syntactic type of left-of-P PPOP. I show that a number of current theories predict two different types despite compelling empirical evidence that there is only one. On the basis of the empirical evidence, I conclude that all left-of-P PPOPs have equal status in the syntax, which poses a considerable challenge to derivations of the verb second constraint. This chapter additionally provides the opportunity to compare and contrast different theories of Old English word order, which will serve as a useful point of reference in later chapters.

2.2   Pronouns as clitics

PPOPs often appear somewhere to the left of a governing preposition in Old English whereas nominal objects do so very rarely. When they appear in a 'special', i.e. left-of-P position, PPOPs are generally treated as special clitics, i.e. grammatically independent elements that attach syntactically and phonologically to an adjacent host (Zwicky 1977), in theories of Old English syntax.[6] Special clitics contrast with simple clitics, which are phonologically dependent but syntactically independent words, while non-clitic words have phonological and syntactic independence.

Old English morphology does not distinguish clitic pronouns from non-clitic pronouns. This is unlike the situation in French, for example, in which the

---

[6] Following Klavans (1985), Pintzuk (1991: 234–6) argues that the phonological host of a special clitic is not necessarily the same as its syntactic host in Old English. The phonological host of Old English clitic PPOPs is not addressed in this thesis.

third person sg. masc. clitic object pronoun, *le*, contrasts with its non-clitic counterpart, *lui* (Cardinaletti & Starke 1999: 174). Orthographic systems may indicate the phonological dependency, and thus the clitic status, of a word, e.g. *show'em how to do it*, but Old English orthography provides no evidence of the phonological reduction of pronouns apart from the occasional spelling of the third person sg. nom./acc. neuter pronoun *hit* as <it> (Koopman 1992: 83).[7] Consequently, the only diagnostic for the status of an Old English pronoun is its syntax, but while special clitic PPOPs can be recognised on the basis of left-of-P placement, there is no independent way to determine whether any given right-of-P PPOP is a simple clitic or a non-clitic pronoun.

The syntactic host of Old English special clitics is generally accounted for in structural terms, e.g. Spec-IP, and clitic attachment is accounted for in directional terms. Proclitics attach to the left of their host, i.e. clitic=host, and enclitics to their right, i.e. host=clitic. The existence of a host in the PP domain to which PPOPs alone can optionally procliticise is widely assumed for Old English. Van Kemenade (1987: 132–3) accounts for this as procliticisation to P^0 and Pintzuk (1991: 276–7, 1996: 395) as procliticisation to the first constituent of PP, although Pintzuk provides no examples in which the first constituent is not the preposition. Procliticisation to P^0 is wholly consistent with the empirical evidence. According to Wende's (1915: 82–107) analysis of the 482 left-of-P PPOPs in his corpus of Old English prose, 322 (66.8%) occur immediately to the preposition's left. Almost exactly the same proportion is found in the much larger YCOE, in which 1,844/2,775 (66.4%) left-of-P PPOPs are left-adjacent to P. The assumption of P^0 as a host for proclitic PPOPs thus neatly captures the placement of two-thirds of left-of-P PPOPs.

There is also broad consensus that clitic PPOPs can move away from P^0 to attach to a host situated at the CP/IP boundary. Structural analyses of this host vary, as I will show in a moment, but it is generally agreed that clitics in this position include those that appear in the following three contexts: firstly, between topic and finite verb in main clauses, as in (1);

---

[7] *Hit* occurs as a PPOP just seven times in the YCOE (Alcorn 2009: 439) and with initial <h-> in each case.

(1)   min God **me** asende **to** sona   his engel

        my God me sent   to immediately his angel

        'my God at once sent to me his angel'

                                              (coaelhom,ÆHom_22:326.3470)

secondly, between finite verb and subject in verb-initial clauses and in operator-initial clauses, i.e. clauses beginning with a *wh*-word or one of a small set of adverbs, mainly *þa* 'then' or *þonne* 'then' (van Kemenade 1987: 111, Pintzuk 1991: 143), as in (2);

(2)   Þa   asende **him** God **to** swyðlice steore

        then sent   him God to severe   punishment

        'Then God sent a severe punishment to him'

                                              (coaelive,ÆLS_[Book_of_Kings]:437.3975)

and thirdly, immediately following the complementiser in subordinate clauses, as in (3).

(3)   ... siþþan **him** cristendom **to** com

          after   them Christianity to came

        '... after Christianity had come to them'

                                              (coorosiu,Or_6:4.136.21.2873)

Van Kemenade (1987) is one of the earliest generative accounts of Old English syntax and I begin my survey of analyses of clitic placement at the CP/IP boundary there.[8] Van Kemenade likens Old English clause structure to that of Modern German and Dutch, hence variation in finite verb (V_{FIN}) placement is accounted for as competition between V_{FIN} and complementiser for placement in C^0. Van Kemenade also identifies C^0 as a clitic host. Topicalisation — the clause-initial placement of a constituent to indicate its pragmatic import — is analysed as movement to Spec-CP, which van Kemenade assumes is possible only in main

---

[8] Although formulated in an earlier version of the CP/IP model, Van Kemenade (1987) is easily translated into the latter as I do here.

clauses. A partial analysis of (1) within this framework is given at (4). In this, and other, syntactic trees I ignore traces within VP except for that associated with the PPOP. My trees indicate that the surface position of topics and special clitics is derived rather than base-generated, but the difference is unimportant here.[9] In the analysis given at (4), the topic — in this case the subject argument — has moved from its base position to Spec-CP and the PPOP is treated as proclitic on C^0.

(4)

```
                        CP
              ┌─────────┴─────────┐
            spec                  C'
             △            ┌───────┴────────┐
          min God_i      C°                IP
                      ┌───┴───┐      ┌──────┴──────┐
                   clitic=   C°     spec           I'
                      │       │      │         ┌────┴────┐
                    me_j   asende_k  t_i       VP        I°
                                          ┌────┴────┐     │
                                     to t sona his engel  t_k
```
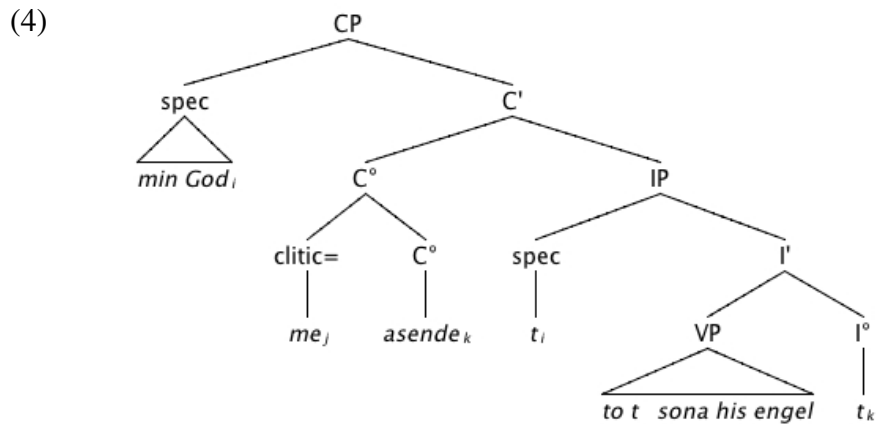
C^0 can also host clitic pronouns in operator-/verb-initial clauses and in subordinate clauses according to van Kemenade (*ibid*), but the direction of attachment is enclitic rather than proclitic in such cases. This is illustrated in the analysis of example (2) given at (5).
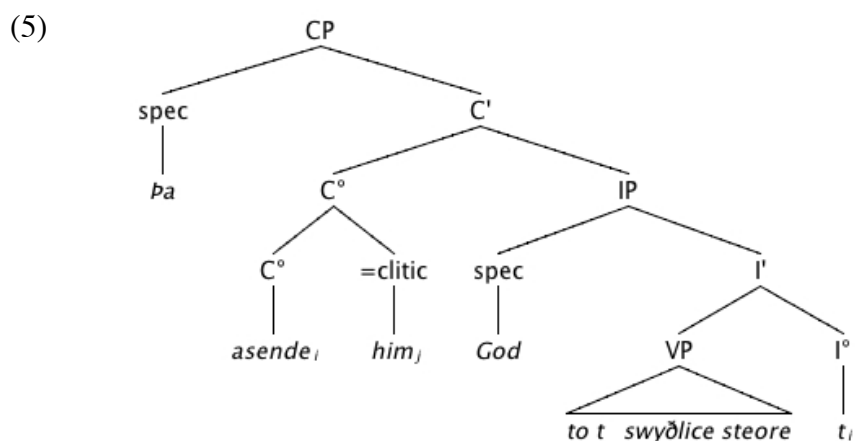
(5)

```
                        CP
              ┌─────────┴─────────┐
            spec                  C'
             │             ┌──────┴────────┐
            Þa            C°                IP
                      ┌───┴───┐       ┌──────┴──────┐
                     C°     =clitic  spec           I'
                      │       │       │        ┌────┴────┐
                  asende_i   him_j   God       VP        I°
                                          ┌────┴────┐     │
                                     to t swyðlice steore  t_i
```

Pintzuk (1991, 1993, 1995, 1996) rejects the idea of a main/subordinate clause asymmetry in V_{FIN} placement for Old English on the basis of the not-infrequent occurrence of verb final main clauses (underivable in van Kemenade 1987) and on the basis of the frequency of verb second subordinate clauses (which are too frequent to be explained satisfactorily by van Kemenade *ibid*).[10] Pintzuk instead likens Old English to Modern Icelandic and Yiddish, in which V_{FIN} moves to I^0 in all types of clause, and moves further to C^0 only where the clause is verb- or operator-initial. Variation between verb second and verb final structures in V-in-I^0 clauses is explained as variation in the internal structure of IP, with head-initial IPs producing verb second word order, and head-final IPs producing verb final word order. In this model, topicalisation is obligatory in main clauses and in subordinate clauses and involves movement to Spec-IP. Spec-IP is identified as a host of clitics and optionality in the direction of attachment is assumed (Pintzuk 1991: 285, 1996: 395). According to this framework, the PPOP in (1) is enclitic on Spec-IP, as shown in (6).

(6)
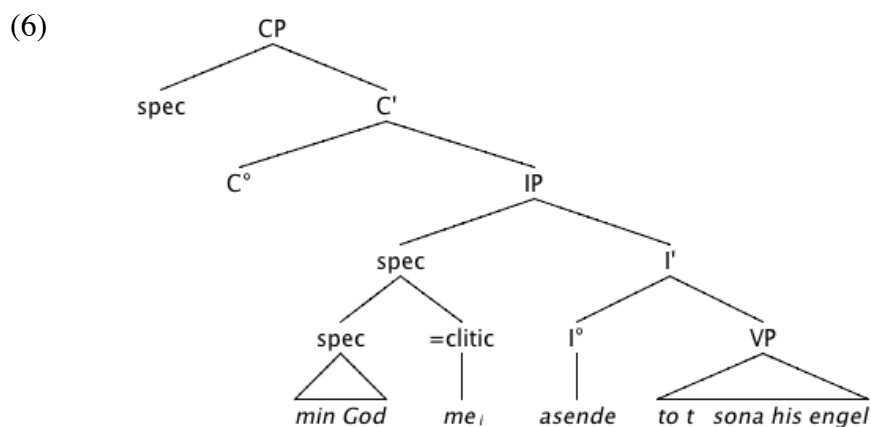
In (2) and (3), the PPOP is proclitic. This is shown for (2) in (7).[11]

---

[10] For data on the frequency of verb final main clauses and verb second subordinate clauses, see Pintzuk (1991: 312, Table 5.7).

[11] The internal structure of IP in (2) is ambiguous. I represent it in (7) as head-initial as this is the more common variant in main clauses (see previous fn.). In (3), IP is head-final.

(7)

```
                    CP
                 /      \
              spec       C'
               |       /    \
              Þa      C°      IP
                      |      /   \
                   asendeᵢ  spec    I'
                          /    \   /  \
                     clitic=  spec I°    VP
                        |      |   |   /    \
                      himⱼ    God  tᵢ  to t swyðlice steore
```
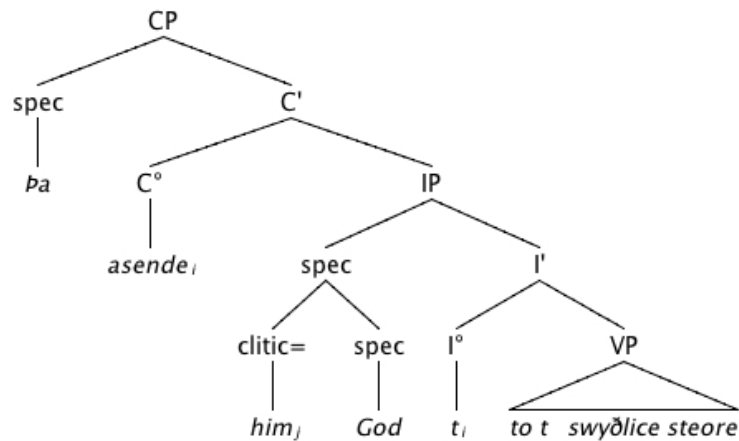
Kroch & Taylor (1997) follow Pintzuk (1991, 1996) in terms of V_{FIN} placement but make an important modification to her analysis of main clause topics. Kroch & Taylor agree that topics always move to Spec-IP, but propose that in topic-initial main clauses the topic moves further, to Spec-CP. In terms of the position of main clause topics, Kroch & Taylor thus agree with van Kemenade (1987). Kroch & Taylor do not actually specify the host at the CP/IP boundary, but as they claim their analysis overcomes Pintzuk's (1991, 1996) requirement that clitic pronouns are sometimes enclitic and sometimes proclitic on this host (*ibid*: 305), it would appear that they see these clitics as positioned regularly at the left edge of IP or, alternatively, regularly to the right of C^0.

So far I have considered analyses of PPOPs placed immediately to the left of the preposition and those placed as in (1)–(3). Sometimes, however, clitic PPOPs follow rather than precede the finite verb in topic-initial main clauses, as in (8), follow rather than precede the subject in operator-initial clauses, as in (9), and appear non-adjacent to a complementiser, as in (10).

(8)   Hermes cwæð **him** eft   **to**
        Hermes said   him again to
        'Hermes again said to him'

        (coaelhom,ÆHom_24:68.3797)

(9)   þa   clypode se cyning **him** drymen   **to**
        then called   the king   him sorcerers to

'then the king called the sorcerers to him'

<div align="right">(cocathom2,ÆCHom_II,_37:278.196.6265)</div>

(10)   ... forþon þe þæt halige wif   **him** ær   **to** cwæð þæt …

         because   the holy   woman him previously to said   that

     '... because the holy woman had previously said to him that …'

<div align="right">(comary,LS_23_[MaryofEgypt]:773.524)</div>

These are not isolated examples. The YCOE provides: 60 examples like (8), i.e. main clauses with [Subject V_{FIN} (...) PPOP ... P] word order; 23 examples like (9), i.e. main clauses with [(operator) V_{FIN} Subject (...) PPOP ... P] word order; and 77 examples like (10), i.e. subordinate clauses with [COMP Subject (...) PPOP ... P] word order.

For Kroch & Taylor (1997), PPOP placement in each of the 160 examples represented by (8)–(10) is a problem as no host is identified between that at the CP/IP boundary and that at the left edge of P^0.[12] In allowing for optional variation in the direction of attachment of clitic pronouns to Spec-IP, Pintzuk (1996) can account for PPOP placement in (9) and (10) as enclitic on (the topic in) Spec-IP, but PPOP placement in (8) cannot be derived. On the other hand, variation in the direction of attachment to Spec-IP means that Pintzuk (1996) predicts a possibility for clitic placement that is not represented by any of the examples presented so far, i.e. [clitic=Topic V_{FIN} ...] in topic-initial main clauses. She gives two examples of this option: (11a), in which the clitic is a PPOP; and (11b), in which it is the object of the verb.[13]

---

[12] Van Bergen (2003: 30–58) shows that subject pronouns sometimes follow rather than precede V_{FIN} in topic-initial main clauses where the verb is a subjunctive or negated indicative form. She points out (*ibid*: 197) this is straightforwardly accounted for by Kroch & Taylor (1997) as V-to-C^0 movement, with clitic placement of the subject pronoun at the CP/IP boundary. This analysis could be extended to object pronouns placed immediately after a subjunctive or negated indicative V_{FIN} in topic-initial main clauses and to the occasional PPOP placed after a positive indicative V_{FIN}, as in (8), but not to any significant number of PPOPs in the latter category.

[13] Pintzuk (1991: 286, 1996: 396–400) treats clitic pronouns as phonological enclitics and allows for their phonological attachment to sentential conjunctions and subordinators. Where there is no initial conjunction, Pintzuk proposes a rule of prosodic inversion to rule out pre-topic placement of clitics in topic-initial main clauses. In other words, Pintzuk (1991, 1996) allows for clitic=topic in main clauses but only if there is an initial conjunction, as in (11).

(11) a. & **him** man nam syððan   frið **wið**

      and him one made afterwards peace with

      'and afterwards peace was made with him'

      (cochronA-5,ChronA_[Plummer]:993.7.1419 [Pintzuk 1996: 392, ex. 33a])

  b. and **hi**   man mæg wenian wundorlice   to gefeohte

      and them one may   tame   wonderfully to battle

      'and one may tame them wonderfully for battle'

        (coaelive,ÆLS_[Maccabees]:569.5203 [Pintzuk 1996: 390, ex. 26a])

As topicalisation is obligatory in main (and subordinate) clauses according to Pintzuk (1991, 1996), one of the two initial pronouns in each of these two examples must be in topic position. Pintzuk (1991: 280–4) recognises the difficulty of determining which of two clause-initial *personal* pronouns is topic, but takes it for granted that the indefinite subject pronoun *man* is the topic in examples like those in (11) (see also Pintzuk 1991: 275, ex. 126). Subsequent work on *man* by van Bergen (2003) shows *man* very often behaves as a clitic, consequently its status as topic in the examples at (11) cannot be taken for granted. It *may* be the topic, in which case the initial pronouns could be treated as proclitic on *man* as Pintzuk (1996: 390, 392) suggests, but it is equally possible that the initial pronouns are in topic position, with *man* an enclitic pronoun.

      Examples like (11a), then, force us to consider the possibility that some specially placed PPOPs are topics rather then clitics, at least under the assumption that topicalisation is obligatory. However, the ability of PPOPs to topicalise in Old English has seldom been considered — and never in detail. The possibility is clearly of relevance for the identification of relevant data for the purposes of this thesis and there are potentially quite far-reaching theoretical implications too. Given the importance of this issue, I deal with the possibility of PPOP topicalisation separately in section 2.3.

      Pintzuk (1991) differs from Pintzuk (1996) in one respect: in the former, Spec-VP is also regarded as a host, but only for clitic PPOPs (1991: 285). As with Spec-IP clitics, clitic PPOPs can optionally attach proclitically or enclitically to Spec-VP (*ibid*: 276). Cliticisation to Spec-VP could account for PPOP placement in (8), which is impossible in Pintzuk (1996), but it could also account for PPOP

placement in (9) and (10). In other words, by allowing for encliticisation to Spec-IP and procliticisation to Spec-VP, Pintzuk (1991) introduces some redundancy to her analysis of PPOP placement. In later work, Pintzuk (1996: 392, fn. 14) rejects Spec-VP as a host on the basis of examples like (12), highlighted by Allen (1990: 148), where the PPOP appears to be VP-internal — between a nominal object and adverb in (12) — but non-adjacent to Spec-VP. Such examples are instead said to involve 'some sort of reanalysis and scrambling within the VP' (Pintzuk *ibid*), an argument which Pintzuk would presumably extend to the example at (8).

(12)   and ofclypode his diacon **him** hrædlice **to**
        and called   his deacon him quickly   to
        'and quickly called his deacon to him'
        (cocathom2,ÆCHom_II,_11:107.528.2324 [Pintzuk 1996: 392, fn. 14, ex. ii])

Aside from the question of what Pintzuk means by 'reanalysis', I do not see how PPOP placement in (8) or (12) can be accounted for in her model without assuming at least one other clitic position situated below Spec-IP and distinct from that on P^0. Van Bergen (2003: 126, fn. 6) suggests it may still be possible to treat the PPOP in (12) as adjacent to Spec-VP on the assumption that the nominal object has scrambled out of VP, a possibility supported by Koopman (1991: 114–7), Haeberli (1999) and Trips (2002: 188–97). So perhaps Spec-VP need not — and should not — have been abandoned as a host for clitic PPOPs after all.

Like Pintzuk (1991), van Kemenade (1987: 126–33) identifies three hosts for clitic PPOPs. In van Kemenade's model, these position are: P^0, available only to PPOPs; Spec-VP; and C^0. Unlike Pintzuk (1991), van Kemenade claims Spec-VP is available to clitic objects of verbs as well as clitic PPOPs and that the direction of attachment to Spec-VP is invariably proclitic. Van Kemenade (*ibid*: 112) additionally allows personal pronoun objects of verbs to procliticise to V^0, but it is unclear whether she extends this option to PPOPs. Although she initially claims that PPOPs can move 'to precisely those positions where other object pronouns can appear' (*ibid*: 115), when the positions for clitic PPOPs are itemised (*ibid*: 115–6, 132–3), V^0 is not mentioned. As noted above, the possibility of procliticisation to Spec-VP means that the position of the PPOPs in examples at

(8)–(10) is unproblematic for Pintzuk (1991). The same can be said for van Kemenade (1987). Assuming VP-external scrambling of the nominal object in (12), this example is unproblematic too. There is also no redundancy in van Kemenade's analysis of clitic placement, i.e. there is only one possible position for the clitic PPOPs considered so far: proclitic on C^0 in (1); enclitic on C^0 in (2) and (3); and proclitic on Spec-VP in (8)–(10) and (12). On the other hand, this parsimony comes at the expense of requiring clitic pronouns at the CP/IP boundary to vary in their direction of attachment in main clauses according to the type of initial constituent: proclitic in topic-initial clauses, as shown in (4), and enclitic in operator-/verb-initial clauses, as shown in (5).

Unlike the other accounts surveyed so far, Hulk & van Kemenade (1997) treat specially placed personal pronoun objects as weak pronouns rather than as special clitics, although they do not discuss PPOPs in particular. They maintain van Kemenade's (1987) analysis of topicalisation but assume a more elaborate structure between CP and VP, including a functional projection (FP) immediately below CP. Spec-FP is identified as the position for weak pronouns, and F^0 is identified as the position for V_{FIN} in main clauses unless operator-initial, for which V_{FIN}-in-C^0 is assumed. As Spec-FP is the only position identified for weak object pronouns, the placement of the majority of left-of-P PPOPs is unaccounted for by Hulk & van Kemenade, including: the two-thirds of specially placed PPOPs that are left-adjacent to their governor; PPOPs that follow V_{FIN} in clauses where V_{FIN} is in F^0, i.e. where the PPOP must be below Spec-FP, as in (8) and (12); and PPOPs that follow a non-topicalised subject, as in (9) and (10), since non-topicalised subjects are assumed to be in their case position which is also below Spec-FP.

A non-clitic analysis of left-of-P PPOPs is considered in some detail by Harris (2006: 37–9). Harris observes that left-of-P placement occurs rarely when the PP is unambiguously outside the VP domain, specifically when the PP is in extraposition or is a constituent of an NP, and argues that this may be interpreted as evidence that left-of-P placement involves weak pronoun movement to the verbal domain. In Chapter 4, we will see that left-of-P placement is indeed rare when the PP is a constituent of an NP, but two groups of examples show that left-of-P placement is possible when the PP is VP-external. One group, noted by Harris himself (*ibid*: 38), involves PPs in extraposition with a reflexive PPOP.

Although Harris found only eight such examples, four of these PPOPs are specially placed. The other group involves examples like (13), where the PP appears to be in topic position. Such examples are also not numerous (the YCOE provides just twenty), but they do suggest that PPOP movement need not necessarily involve movement to the verbal domain.

(13)   … and **him æfter** ferde Iudas mid fultume

        and him after   went Judas with support

   'and after him went Judas with support'

                                    (coaelive,ÆLS_[Maccabees]:498.5168)

Overall, I see no advantage in treating some or all left-of-P PPOPs as weak pronouns rather than as special clitics. The same variety of special positions is needed under both approaches, but a weak pronoun analysis can achieve this only (i) by assuming a considerably more elaborate structure than is needed under a clitic analysis and (ii) by introducing an alternative explanation for the fact that two-thirds of left-of-P PPOPs are immediately adjacent to their governor regardless of where the PP happens to be placed.[14] For the purposes of this thesis, I therefore assume that left-of-P PPOPs are special clitics rather than weak object pronouns. The possibility that some left-of-P PPOPs are topics rather than special clitics is considered in the following section.

The data considered in this section suggest that at least three hosts need to be recognised for clitic PPOPs: P^0; one at the CP/IP boundary; and at least one other somewhere below the CP/IP boundary. Two of the clitic analyses surveyed, i.e. Pintzuk (1991) and van Kemenade (1987), provide for all three options although there is some redundancy in the former and, it turns out, in the latter too. Various scholars have shown that it is extremely difficult to show that simple personal pronoun objects of verbs can occur in positions inaccessible to their nominal counterparts unless situated in one of the positions illustrated by examples (1)–(3) (e.g. Pintzuk 1991: 222–3, 1996: 389–91, Koopman 1992: 53–4,

---

[14] See van Bergen (2003: 171–8) for a further argument against a weak pronoun analysis of Old English pronouns.

1997a: 79–80). Van Kemenade's (1987) claim that Spec-VP and V^0 is available to clitic objects of verbs in particular thus has little support.

Arguably, the most parsimonious account of PPOP placement would follow from the identification of Spec-VP as an additional host, as in Pintzuk (1991) and van Kemenade (1987), within the overall framework of Kroch & Taylor (1997).[15] Such a modification to Kroch & Taylor's model could certainly account for examples (1)–(3) and (8)–(10) without assuming variation in direction of clitic attachment. Examples like (12) would require an additional assumption about scrambling of nominal objects, but such an assumption for Old English is not without independent support. Whether this refinement to Kroch & Taylor (1997) would be sufficient to account for the placement of <u>all</u> special clitic PPOPs is a question requiring further research, as is the question of whether it is possible to restrict the option of cliticisation to Spec-VP to clitic PPOPs only. Pintzuk (1991) incorporates such a restriction but it appears to be stipulative. One of the theories underpinning her analysis of clitic attachment requires clitics to subcategorise for their host (Klavans 1985) but Pintzuk does not address how PPOPs can subcategorise for Spec-VP when clitic objects of verbs cannot.

Having illustrated some of the problems in accounting for the variable placement of special clitic PPOPs, I now turn to the treatment of this variability in the present study. Following van Kemenade (1987), Pintzuk (1991, 1996) and Kroch & Taylor (1997), I assume that regardless of the particular position they occupy, special clitic PPOPs have equal status in the syntax. This assumption is already supported by data in Taylor (2008: 344). Her multivariate analysis of variation in PPOP placement in a sizeable subset of the YCOE found factors that favour *him to* over *to him* also favour *him ... to* over *to him*, and factors that favour *to him* over *him to* also favour *to him* over *him ... to*. In other words, Taylor found no evidence of structure in the variation between placement immediately to the left of the preposition and placement somewhere further to its left. The same is true for the results of the present study, which uses a more complex statistical model and a larger set of data. Although differences might emerge if a three- rather than two-way distinction among left-of-P data were to be made, I have not

---

[15] Kroch & Taylor (1997) would perhaps prefer to identify this additional host as being located at the IP/VP boundary.

investigated the possibility: the more distinctions that are made, the fewer the examples available for each variant and consequently the harder it becomes to distinguish random fluctuations from linguistically significant trends. As the statistical evidence supports the assumption that placement of special clitic PPOPs in one position rather than another is a freely available option, this thesis will focus on detecting structure in variation between the realisation of PPOPs as special clitic pronouns on the one hand and as simple or non-clitic pronouns on the other.

2.3 Pronouns as topics

In the previous section, I noted that Pintzuk (1996: 392) avoids a topic analysis of the initial PPOP in example (11a) by treating it as proclitic on topicalised *man*. The same analysis is given for the example at (14) in Pintzuk (1991: 210).

(14)   & **heom** man syððan   þær frið **wið** nam

and them one afterwards there peace with took

'and afterwards peace was made with them there'

(cochronA-5,ChronA_[Plummer]:1001.16.1432 [Pintzuk 1991: 210, ex. 39b])

I pointed out, however, that once it is accepted that *man* itself exhibits clitic-like behaviour, there is no aspect of Pintzuk (1991, 1996) that would preclude treating *man* in (14) (and 11a) as the clitic and the initial PPOP as topic. Moreover, a topic analysis of the initial PPOP in (11a) and (14) is not only possible within Kroch & Taylor's (1997) model, it is unavoidable given their assumption of obligatory topicalisation and their particular analysis of clitic placement at the CP/IP boundary. As Hulk and van Kemenade (1997) are primarily concerned with what negation patterns reveal about Old English clause structure, they do not discuss whether main clause topicalisation is obligatory in their view. Accordingly, it is not clear whether a topic analysis of the initial PPOP in (11a) and (14) is avoidable in their framework.

So, on the one hand we have Pintzuk (1991, 1996) and Kroch & Taylor (1997) who assume that specially placed personal pronouns are special clitics and that topicalisation is obligatory in all types of clause. For them, a topic analysis of

the initial PPOPs in (11a) and (14) is at least a possibility (for Pintzuk) if not a requirement (for Kroch & Taylor), although in none of these accounts is the possibility/requirement recognised for PPOPs in particular. On the other hand, however, we have van Kemenade (1987: 117, 132), who makes the explicit claim that PPOPs can topicalise, although it transpires — ironically — that this claim is entirely unnecessary within her particular framework, as I will show.

The question central to this section, then, is whether simple personal pronoun objects of prepositions can function as topic. Despite different predictions for the possibility of embedded topicalisation, as outlined in the section 2.2, it is generally agreed that subordinate clause-initial personal pronoun objects are special clitics rather than topics. This follows from the fact that they often appear immediately after the complementiser (van Kemenade 1987: 111, 113, 116, Pintzuk 1991: 203, 208, 211, Koopman 1992: 46–8) whereas nominal objects rarely do. For this reason, this section will concentrate on whether PPOPs can function as topic in main clauses in particular. In addition to the theoretical implications of this question which I identify, this question is clearly of direct relevance to the type of data that is appropriate to this thesis: if PPOPs can topicalise, then left-of-P placement cannot be regarded as a sufficient condition for recognising special clitic PPOPs.

As discussed in the previous section, theoretical treatments of topicalisation in Old English differ in a number of respects that reflect differences in the formulation of V_{FIN} placement and different predictions for topicalisation. According to van Kemenade (1987: 43–8, 55, 1997b: 333), topicalisation is (a) optional and (b) possible only in main clauses and in embedded main clauses, e.g. complements of so-called 'bridge verbs', mainly verbs of saying. According to Pintzuk (1991: 73, 1996: 379) and Kroch & Taylor (1997: 305–10), topicalisation is obligatory in subordinate as well as main clauses, although Kroch & Taylor suggest that non-syntactic, specifically discourse-based information structure, considerations produce a low frequency of non-subject topics in subordinate clauses in comparison to their frequency in main clauses (*ibid*: 309).

Different predictions aside, there are two generally agreed characteristics of main clause topicalisation in Old English. Firstly, more often than not the topic position is filled by the subject argument (e.g. Kohonen 1978: 154, Kroch &

Taylor 1997: 301–2). Examples of topicalised nominal subjects can be found in the previous section at (1) and (8), but personal pronoun subjects can topicalise too (van Kemenade 1987: 109, Pintzuk 1991: 201, fn. 75), e.g. (15).

(15)   He eode æfter mæssan ut   of   þam temple
        He went after mass   out from the   temple
        'He went out from the temple after mass'

                                            (coaelive,ÆLS_[Basil]:169.562)

The examples at (16) illustrate some non-subject topics, which are far from exceptional in main clauses, as we will see a little later.

(16) a. Ælc   riht   sculon gehadode men lufian
        Each virtue must   ordained men love
        'Each virtue ordained men must love'

                                            (cowulf,WHom_10a:10.771)

    b. On þam æfene   sæt se hælynd mid hys twelf   leorningcnihtum æt
       on that evening sat the saviour with his twelve disciples   at
       gereorde
       meat
       'On that evening the saviour sat with his twelve disciples at dinner'

                                            (cowsgosp,Mt_[WSCp]:26.20.1854)

Secondly, main clauses with a non-topicalised nominal subject, as in (16), typically show a feature that reflects the verb second ('V2') tendency of Old English main clauses, namely subject-verb inversion. Whereas non-topicalised nominal subjects typically invert with the verb in topic-initial main clauses, non-topicalised subject personal pronouns and the indefinite subject pronoun *man* do so only rarely, e.g. (17).[16]

---

[16] For data on the frequency with which nominal subjects fail to invert in topic-initial main clauses, see Koopman (1997b: 311–15) and Haeberli (2002: 249–52). For data on the frequency with which personal pronoun subjects do invert in topic-initial main clauses, see Koopman (*ibid*) and van Bergen (2003: 30–58).

(17)　Ðas　godnysse　　we sceolan simble lufian

these goodnesses we must　　always love

'These virtues we must always love'

(coaelive,ÆLS_[Christmas]:93.74)

As said, van Kemenade (1987: 117, 132) claims PPOP topicalisation to be possible, a claim that is accepted by Koopman (1997a: 77). In support of her claim, van Kemenade provides just one example, which I give at (18).

(18)　and **me** com　þærrihte **to** Godes encgel mid　rode

and me came directly　to God's angel　with rood

'and God's angel came directly to me with a cross'

(coaelive,ÆLS[Agnes]:355.1960 [van Kemenade 1987: 118, ex. 12a])

That van Kemenade (1987) should analyse this particular pronoun as topic is surprising on two counts. The first reason requires some background. Van Kemenade (*ibid*: 43–8) treats V2 and topicalisation as separate phenomena. This enables her to provide for the possibility of topic-less main clauses, which she exemplifies with a number of examples including (19).

(19)　Wæs Hæsten þa　　þær　cumen mid his herge, þe　ær　　　æt Middeltune

was　Hasten　then there come　with his host　that before at Milton

sæt

sat

'Hasten had come there then with his host, which had previously encamped at Milton'

(cochronA-2a,ChronA_[Plummer]:894.43.1045 [van Kemenade 1987: 44, ex. 65a])

In addition, she accounts for the placement of pronouns between topic and finite verb, as in (20), to procliticisation of the pronoun (*ibid*: 113, 116), as shown in the previous section at (4).

(20)   min God **me** asende **to** sona   his engel

       my God me sent   to immediately his angel

       'my God at once sent to me his angel'

<div align="right">(coaelhom,ÆHom_22:326.3470)</div>

Given the twin possibilities of procliticisation to a fronted finite verb and V2 without topicalisation, the initial pronoun in (18) could be analysed as a clitic rather than topic. Moreover, as exactly these aspects of van Kemenade's framework are used to treat the clause-initial PPOPs in (21) as clitics (*ibid*: 116), it is not clear why she treats the one in (18) differently.

(21) a. and **him** com þæt leoht **to**, þurh   Paules lare   syððan

      and him came the light to   through Paul's teachings after

      'and afterwards he was enlightened through Paul's teachings'

<div align="center">(coaelive,ÆLS_[Denis]:17.5790 [van Kemenade 1987: 116, ex. 11b])</div>

   b. ac **him** com fyr **to** færlice   ehsynes

      but him came fire to suddenly visibly

      'but suddenly a light came to him visibly'

<div align="center">(coaelhom,ÆHom_10:170.1495 [van Kemenade 1987: 116, ex. 11c])</div>

The second reason why van Kemenade's topic analysis of the pronoun in (18) is surprising can be put in her own words: '[t]opicalisation of a NP prepositional object always involves pied piping in OE' (*ibid*: 152). The reference to topicalisation with pied piping is an unnecessary complication for present purposes. The more usual term for this is PP topicalisation. Example (16b) has a topicalised PP with a nominal object. In (22) the object is a PPOP.

(22)   **for ðe** arærde se ælmihtiga God us of   eorðan ær   ðam micclum

       for you raised the almighty God us from earth   before the great

       dæge

       day

       'for you the almighty God raised us from earth before the great day'

<div align="right">(cocathom2,ÆCHom_II,_31–32:248.219.5529)</div>

As topics are syntactically (and phonologically) independent constituents, van Kemenade's claim that simple personal pronouns can topicalise independently of a governing preposition when their nominal counterparts cannot is distinctly odd, yet no explanation is given.

Van Kemenade's clitic analysis of the examples at (21) rests entirely on her analysis of examples like (19) as topic-less clauses. This type of clause, i.e. positive verb-initial declaratives in which all arguments are present, are commonly said to illustrate 'narrative inversion' (e.g. Los 2000: 263). Analyses of narrative inversion in Old English are more often aligned with analyses of other verb first word orders, e.g. imperatives and direct questions, which in turn are aligned with V2 word orders by assuming the presence of a covert clause-initial operator (e.g. Pintzuk 1991: 139, Kroch & Taylor 1997: 303, Kroch, Taylor & Ringe 2000: 364–5). As co-author of a later textbook in which clauses with narrative inversion are grouped with operator-initial clauses (Fischer *et al*. 2000: 106–7), it would appear that van Kemenade herself would now treat (19) as operator-initial. Crucially, if the topic-less analysis of (19) is now abandoned, then her (1987) argument for optional topicalisation disappears and, consequently, so too does her argument for treating the pronouns in (21) as special clitics, since clitic pronouns always follow rather than precede the finite verb in operator-initial clauses (e.g. van Kemenade 1987: 111, 113, 116, Pintzuk 1991: 203, 208, 211, Koopman 1992: 46–8). So van Kemenade herself would seem to need a topic analysis of the pronoun in (18) — and presumably those in (21) — after all, which leaves unexplained the lack of examples involving topicalised nominal objects of prepositions.

Pintzuk (1991, 1996) and Kroch & Taylor (1997) make no explicit claim about the (im)possibility of topicalisation of objects of prepositions, but they too are unable to give a clitic analysis for the initial PPOPs in (18) and (21): assuming a topic-initial structure for these examples, the pronouns cannot be clitics since obligatory topicalisation is presumed; assuming, instead, the presence of some covert initial operator, the pronouns again cannot be clitics since clitic pronouns are (correctly) predicted to follow rather than precede V_{FIN} in operator-initial clauses. So a topic analysis of certain clause-initial PPOPs seems unavoidable in these accounts too.

In short, examples like (18) and (21) present something of a theoretical conundrum. On the one hand a topic analysis of their initial PPOPs is desirable on the basis that derivations of the V2 constraint require the finite verb to be preceded by *something*, but that 'something' cannot be a (null) operator in these particular cases, and I have found no account of V2 that would allow it to be a clitic. On the other hand, a clitic analysis of all left-of-P PPOPs in Old English, i.e. regardless of their particular position, is desirable on the basis that it explains their freer word order in comparison to their nominal counterparts. The problem is that theories of Old English syntax do not permit a pronoun to be a syntactically independent topic and a syntactically dependent special clitic at the same time.

The problem at hand is not unlike the one posed by the Dutch R-pronoun *daar*. Unlike non-R-pronouns and nominals, *daar* and its unstressed counterpart *er* invariably precede a governing preposition (e.g. van Riemsdijk 1982, van Kemenade 1987: 119–26, den Dikken 2010, Koopman 2010), e.g. (23).

(23) We hadden {**daar/er**} liever niet **op** gewacht
      We had   {there/there} rather not for waited
      'We had rather not have waited for {that/it}'

                                                (van Eynde 1999: 143, ex. 20)

However, as object of P, *daar* can topicalise but *er* cannot, e.g. (24).

(24) {**Daar/*er**}   had ze niet **aan** gedacht
      {there/there} had she not of   thought
      'Of {that/it} she had not thought'

                                                (van Eynde 1999: 143, ex. 22)

Further, non-R-pronouns and nominals cannot topicalise out of a PP either (van Riemsdijk 1982: 138), e.g. (25).

(25) ***Mijn moeder** heb je   deze plaat **voor** gekocht
       my   mother   have you this record for   bought
      'My mother you bought this record for'

                                                (van Riemsdijk 1982: 138, ex. 10a)

The syntax of *daar* thus appears to be 'special' in terms of its invariable left-of-P placement, but its ability to topicalise indicates that it is also a syntactically independent word. The difference between *daar* and Old English left-of-P PPOPs, however, is that the special syntax of *daar* (and of *er*) is generally associated with some feature peculiar to R-pronouns (e.g. van Riemsdijk 1982, den Dikken 2010, Koopman 2010), whereas the special syntax of pronominal clitics is generally associated with some structural deficiency of the pronoun (e.g. Cardinaletti 1994, 1999, Cardinaletti & Starke 1996, 1999). The difference between *daar* and *er*, then, is that *er*, like personal pronoun clitics, can be viewed as the structurally deficient counterpart of non-deficient *daar* (van Eynde 1999, van de Visser 2002), as is evident from their contrasting behaviour with respect to modification and coordination, e.g. (26).[17]

(26) a. We zijn precies {**daar/*er**} waar de Greenwich lijn de evenaar kruist
   'We are exactly there where the Greenwich line crosses the equator'

   (van Eynde 1999: 143, ex. 21)

   b. Wil je liever hier of {**daar/*er**} zitten?
   'Would you rather sit here or there?'

   (van Eynde 1999: 144, ex. 23)

Another problem similar to the one at hand is described in Haegeman (1999: 261, fn. 1), where the syntactic status of the West Flemish deficient third person subject pronoun *ze* is called into question. This pronoun can satisfy the V2 constraint, as shown in (27), which suggests it is a syntactically independent XP.

(27)   Ze goa dienen boek kuopen
   'She is going to buy that book'

   (Haegeman 1999: 261, ex. 1a)

---

[17] As Old English *þær* 'there' is morphologically invariant, it is possible that it represents a strong form, equivalent to Dutch *daar*, as well as a deficient form, equivalent to Dutch *er*. This is one good reason for making no assumptions about the syntactic status of *þær* 'there' (and *her* 'here') when governed by a preposition without firstly undertaking the type of study recommended by van Bergen (2003: 144).

However, *ze* also shows clear clitic (as opposed to weak and strong pronoun) behaviour by requiring repetition under sentence coordination, as in (28), and in allowing its referent to be doubled, e.g. (29).

(28)   Ze goa dienen boek kuopen and *(ze) goat em vanoavend nog lezen
        'She is going to buy that book and (she) is going to read it tonight'

                                                    (Haegeman 1999: 261, ex. 1b)

(29)   Ze goa (zie) dienen boek kuopen
        'She is going to buy that book'

                                                    (Haegeman 1999: 261, ex. 1c)

This suggests *ze* may also be (or is) a clitic, which, as Haegeman (*ibid*: 261) points out, would be problematic for the derivation of the V2 word order of these examples. Haegeman (*ibid*) suggests one solution would be to treat the subject clitic as licensing *pro*, which would in turn satisfy the V2 constraint, but this solution could not be extended to the clause-initial PPOPs in (18) and (21). As the licensing element, *pro* would be PP-internal in the case of (18) and (21), which is too low down in the clause for *pro* to satisfy the V2 constraint. A weak pronoun analysis of these pronouns would not provide a solution either. Cross-linguistic evidence shows that weak object pronouns, unlike weak pronoun subjects, cannot topicalise (e.g. Weerman 1998: 62, Cardinaletti 1999: 50, Fanselow 2009: 111).

So the PPOPs in (18) and (21) cannot be clitics (because of the need to satisfy the V2 constraint), nor can they be topics (because their nominal counterparts don't topicalise), nor are they *daar*-like (because there is no identifiable feature that left-of-P PPOPs possess that right-of-P PPOPs do not — other than their special syntax of course), nor can they be weak pronouns (on the basis of cross-linguistic evidence). But these pronouns must have the status of *one* of these types, unless there is some other option that has yet to be identified by the linguistics community at large.

In what follows, I offer four pieces of empirical evidence that, when taken together, suggest that a special clitic analysis of the pronouns in question is highly

desirable. Before doing so, I firstly explain how relevant examples were identified within the YCOE.

I count as clause-initial those PPOPs that occur as the first word of a main clause or the first word of a main clause following a clausal conjunction. I also looked for examples in which the pronoun is preceded only by a vocative element or by a vocative and clausal conjunction but found none. I did not seek examples from clauses lacking an overt subject or finite verb: in such cases, there is little or no hope of distinguishing between topic and clitic on empirical grounds. In addition, for those that lack an overt subject in particular, it may be possible to assume that an empty category occupies the initial position and that this empty category satisfies the V2 constraint, similar to the analysis suggested in Haegeman (1999: 261, fn. 1) for clause-initial placement of *ze* in (27)–(29). Lastly, I do not count as potential topics the twelve main clause-initial PPOPs that are adjacent to P, e.g. (30), as each could be derived via PP topicalisation with procliticisation of the pronoun to P^0.

(30) a. and **him to** com se   halga gast

and him to came the holy   spirit

'and to him came the holy spirit'

(cocathom2,ÆCHom_II,_3:22.108.520)

b. and **heom betweonan** cwædon, þa ealdras and þa mæssepreostas: la ...'

and them between   said   the elders   and the mass-priests   lo

'and among themselves the elders and the mass-priests said, "Lo ..."'

(conicodA,Nic_[A]:15.2.4.313)

The total number of clause-initial PPOPs identified is 127. In 90 of these examples (71%), the PPOP is the only element (excluding a clausal conjunction) to precede the finite main verb.

One type of evidence that would suggest that these main clause-initial PPOPs are topics rather than special clitics would be their ability to be separated from the finite verb by a personal pronoun subject, just as the personal pronoun subject separates topic and V_{FIN} in (17). According to Pintzuk (1991: 284), an independent surface string constraint on adjacent personal pronouns rules out the

possibility of personal pronoun object > personal pronoun subject word order regardless of the pronouns' individual status as topic or clitic. Personal pronoun objects do indeed characteristically follow rather than precede an adjacent personal pronoun subject: in the YCOE there are more than 7,500 such examples. But there are also about 40 that violate Pintzuk's ordering constraint, as in (31).

(31) a. hine ic lufige ofer eallum oðrum þingum

      it   I love over all   other things

      'it (i.e. wisdom) I love above all other things'

<div align="right">(cosolilo,Solil_1:43.21.553)</div>

    b. þe hi clypiað to him

      you they call   to them

      'you they call to themselves'

<div align="right">(coaelive,ÆLS_[Christmas]:72.58)</div>

This type of example is noted by van Bergen (2003: 186–7), who concludes that the initial object pronoun is 'almost certainly' in topic position in such cases (*ibid*: 186). Consequently, she argues that the ordering constraint proposed by Pintzuk is too strong and should be restricted to apply only to adjacent personal pronoun *clitics*. So, the type of example we are looking for to support a topic analysis of main clause-initial PPOPs should, in principle, be possible.

      None of the 127 main clause-initial PPOPs precedes a personal pronoun subject (nor do any subordinate clause-initial PPOPs). This is exactly what we would expect if the clause-initial pronouns are clitics, at least on the assumption of a subject > object surface order constraint on clusters of clitic personal pronouns. However, the absence of examples which would support a topic analysis of clause-initial PPOPs does not mean that a topic analysis can be ruled out: their absence could simply be accidental to the sample rather than because of ungrammaticality.

      Another type of example that would support a topic rather than clitic analysis of main clause-initial PPOPs would be those with clear evidence of inversion of nominal subject and V_{FIN}, just as the non-subject topics trigger inversion in the examples at (16). It is well known that V_{FIN} is more likely to be in

second position, and so available for inversion, in uncoordinated than in coordinated main clauses (e.g. Mitchell 1985: §§904–5, Traugott 1992: 277, Koopman 1995, Pintzuk & Haeberli 2008), so I divide the data accordingly. Each of the 41 uncoordinated main clauses with an initial PPOP has a nominal subject, but only four (10%) show clear evidence of inversion, i.e. [PPOP V_{FIN} Subject ...] word order. One of these four examples, (32), almost certainly involves the prefixed verb *onbelædan* 'to inflict upon' rather than *belædan* 'to lead astray' plus *on* 'on, in', so should probably be excluded.

(32)   **Us** is unlytel broga **an** beled

       Us is great   terror inflicted-upon

       'Great terror is inflicted upon us'

                                           (cochdrul,ChrodR_1:79.77.987)

Another of the examples with apparent inversion is given at (33). There are six similar examples among the 41 uncoordinated main clauses with an initial PPOP, all from the same text (*Orosius*), except that in those cases the subject is clause-final.

(33)   Corsica **him** is Romeburh **be eastan**[18]

       Corsica him is Rome-city by east

       'Corsica, the city of Rome is to the east of it'

                                     (coorosiu,Or_1:1.21.17.425)

All seven appear to involve what is commonly referred to as Hanging Topic Left Dislocation (HTLD) (e.g. Grohmann 2003, Boeckx & Grohmann 2004). HTLD is characterised by a number of features cross-linguistically: the left-dislocated element ('the dislocate') appears in a default case and co-refers with a clause-internal resumptive pronoun (RP); the RP is usually, but not always, a

---

[18] Although written as two words in (33), *beeastan* 'to the east of' is recognised as a preposition in Toller (1921), Clark Hall (1960) and Campbell (1969: §669, fn. 1), as are *benorþan*, *besuþan* and *bewestan*, e.g.:

(i)   Burgendean habbað þone sæs earm **be westan him**
        Burgundians have   the sea's arm by west   them
        'The Burgundians have the arm of the sea to the west of them'
                                        (coorosiu,Or_1:1.13.22.217)

demonstrative pronoun; and the RP appears in the case that its governor would normally assign to a functionally equivalent (pro)nominal argument. The RP often appears at the left edge of the clause, but again not always. According to Boeckx & Grohmann's (2004) analysis of HTLD, the RP spells out a trace of the dislocate's movement from its base-generated position. While agnostic about the structural positions of the dislocate and RP, Boeckx & Grohmann (*ibid*: 144) note that RPs are often equated with topic position. It is certainly interesting that topicalisation appears to be possible under the special circumstance of the left dislocation of the preposition's object, but as there are so few examples, all from a single text and each involving a clause-initial RP in particular, I do not think they provide much by way of insight into the syntactic status of clause-initial PPOPs.

The other two examples with apparent inversion are at (34). These are without question the best examples in the YCOE to support the claim that PPOPs can topicalise.

(34) a. **Him** com stemn **to**, þus clypiende þriwa

him came voice to thus calling   thrice

'A voice came to him, thus crying thrice'

(coaelive,ÆLS[Peter's_Chair]:87.2330

b. **Him** cwæð Nichodemus **to**, swiðe þæs   ofwundrod

him said   Nichodemus to, very   of-this astonished

'Nichodemus, very astonished by this, said to him'

(coaelhom,ÆHom_13:10.1886)

To calculate the frequency of subject-verb inversion in the 86 coordinated main clauses with an initial PPOP, we must firstly exclude the five examples in which the subject is *man* as, like personal pronoun subjects and unlike nominal subjects, *man* normally inverts only, although not always, in operator-initial clauses (van Bergen 2003: 95). We have already seen two examples in which an initial PPOP immediately precedes *man* in a coordinated main clause, i.e. (11a) and (14). Another example is given at (35).

(35)   and **him** man lædde þone witegan **to** Danihel

and him one led   the   prophet to Daniel

'and the prophet Daniel was led to him'

(cocathom2,ÆCHom_II,_33.253.142.5660)

In the other 81 examples, the subject is nominal (recall that there are no examples in which a clause-initial PPOP precedes a personal pronoun subject). Of these 81 examples, 22 (27%) have [PPOP V_{FIN} Subject ...] word order. Two examples have already been given at (21). The three examples at (36) illustrate another ten.[19]

(36) a. and **him** beah   god   dæl þæs   folces **to** þe ær   under

and him submitted good deal of-the people to that previously under

Deniscra manna anwealde wæron

Danish   men   power   were

'and a good deal of the people who were under the power of the Danish men submitted to him'

(cochronC,ChronC_[Rositzke]:913.1.3.1070)

b. and **him** comon englas **to**

and him came   angels to

'and angels came to him'

(cocathom1,ÆCHom_I,_11:267.25.2002)

c. Ac **him** cwæð se Hæland **to**

but him said   the Saviour to

'But the Saviour said to him'

(coaelhom,ÆHom_13.13.1888)

---

[19] Example (36a), from the entry for AD 913 in the C-text of the *Anglo-Saxon Chronicle*, also occurs in the entry for AD 913 in the A-text (cochronA-2c,ChronA_[Plummer]:913.7.1221) and the D-text (cochronD,ChronD_[Classen-Harm]:913.4.1002). Three other examples involving *bugan to* 'to submit to' with subject-verb inversion occur in entries for AD 1016: in the C-text (cochronC,ChronC_[Rositzke]:1016.44.1675); the D-text (cochronD,ChronD_[Classen-Harm]: 1016.47.1644); and the E-text (cochronE,ChronE_[Plummer]:1016.46.1975). *Tobugan* is not listed as a prefixed verb in Clark Hall (1960), Bosworth & Toller (1898) or Toller (1921) so the examples represented by (36a), probably cannot be interpreted this way. An identical example to the one at (36b) occurs in the same text, *Ælfric's Supplementary Homilies* (coaelhom,ÆHom_13. 95.1928), and an identical example to the one at (36c) occurs in that same text, *Ælfric's Catholic Homilies I* (cocathom1,ÆCHom_I,_11:270.127.2093).

So, there are a number of examples with subject-verb inversion to lend support to a topic analysis of the initial PPOPs — those at (34) in particular — but the proportion is rather small: 2 out of 39 relevant examples in uncoordinated main clauses and 22 out of 81 relevant examples in coordinated main clauses, an overall rate of 20%. But, once again, the fact that there is no clear evidence of inversion in the majority of cases does not mean that those initial pronouns are not topics: it simply means that in 80% of cases the evidence is inconclusive.

I now turn to the third piece of evidence that suggests that clause-initial PPOPs are clitics rather than topics. This evidence comes from their distribution by clause type. As said, there are 127 main clause-initial PPOPs in total: 41 (32%) in uncoordinated main clauses and 86 (68%) in coordinated main clauses. After adjusting for the example involving a prefixed verb, (32), and the seven examples involving HTLD, exemplified by (33), the figures are 33 (28%) in uncoordinated main clauses and 86 (72%) in coordinated main clauses. Neither van Kemenade (1987), Pintzuk (1991, 1996) nor Kroch & Taylor (1997) predict that non-subject topicalisation should be any less (or, indeed, more) frequent in uncoordinated than in coordinated main clauses, at least not by means of their syntactic apparatus. Nevertheless, the results of two separate corpus studies of word order in Old English prose indicate that non-subject topicalisation does indeed occur at different frequencies in the two types of main clause, but both studies show that non-subject topicalisation is more frequent when the clause is <u>un</u>coordinated.

Kohonen (1978: 154) calculates non-subject topicalisation at a frequency of 12% (103/895) in coordinated main clauses and 40% (524/1,325) in uncoordinated main clauses.[20] Table 2.1 gives comparative figures derived from Bech (2001). The first two rows of data are from Bech (*ibid*: 89, Table 4.10). Bech's X-initial data, which I give in row two, exclude verb-initial clauses but include *þa-* and *þonne*-initial, i.e. clear operator-initial, clauses. *Þa-* and *þonne*-initial clauses are quantified in row three (from *ibid*: 100) and deducted from data in row two to calculate the number of main clauses with non-subject topicalisation, given at row four.

---

[20] Kohonen's definition of topicalisation (1978: 151) is compatible with that assumed here.

Table 2.1 Frequency of topicalisation in main clauses (Bech 2001)

	Coordinated	Uncoordinated	Total
Subject topic	**(68%) 468**	**(57%) 660**	**(61%) 1,128**
X-initial	227	847	1,074
less operator-initial	10	340	350
Other topic	**(32%) 217**	**(43%) 507**	**(39%) 724**
Total	685	1,167	1,852

Both studies indicate a similar frequency for non-subject topicalisation in uncoordinated main clauses (40% per Kohonen, 43% per Bech), but they vary considerably with respect to the frequency in coordinated main clauses (12% per Kohonen, 32% per Bech). As the phenomenon is significantly more frequent in uncoordinated main clauses according to both sets of data, I have not attempted to reconcile the difference in their estimations for coordinated main clauses.

In the following chapter, it is shown that the number of PPOPs in uncoordinated main clauses (2,558) is approximately equal to the number in coordinated main clauses (2,670).[21] Consequently, estimations of non-subject topicalisation in Kohonen (1978) and Bech (2001) would lead us to expect the number of clause-initial PPOPs to be significantly lower in coordinated main clauses than in uncoordinated main clauses if these pronouns were topics. The evidence does not fit this pattern at all: more than twice as many clause-initial PPOPs occur in coordinated main clauses (N=86) than in uncoordinated main clauses (N=33). This distribution is, on the other hand, generally consistent with the number of main clause PPOPs that appear in some other, i.e. non-initial, left-of-P position. There are 1,497 such examples: 808 (54%) in coordinated main clauses and 689 (46%) in uncoordinated clauses main clauses. Although this difference is less extreme, it shows that main clause-initial PPOPs behave more like special clitics than topics in terms of their frequency by main clause type.

To explain the significance of the final piece of evidence in favour of a clitic rather than topic analysis I must anticipate one of the most interesting findings to emerge from this thesis. In Chapter 3, I show that PPOPs very rarely precede a governing preposition unless they are clearly or very probably dative, a

---

[21] See Table 3.15.

result that is discussed in further detail in Chapter 4. In light of this finding, it is extremely interesting that none of the 127 main clause-initial PPOPs is definitely, or very probably, not dative. As 'the bulk of Old English prepositions prefer the dative' (Mitchell 1978: §27), it might be assumed that the shortage of accusative and genitive forms among main clause-initial PPOPs requires no special explanation. This is probably true for genitive PPOPs as there are only 31 in the YCOE in total, but the data suggest otherwise for accusative PPOPs. Among third person PPOPs in the YCOE, the ratio of dative to accusative is 5,937:811 or 7.3:1.[22] Of the 119 main clause-initial PPOPs (i.e. discounting the example involving a prefixed verb and the seven involving HTLD), 108 are clearly dative, and 10 of the 11 case-ambiguous forms are governed by a preposition that governs dative at least 95% of the time, so these 10 are very probably dative too.[23] Given 118 clear/likely dative PPOPs, a 7.3:1 ratio would predict around 16 accusative examples. The fact that none is unambiguously accusative, and only one is potentially accusative, is the fourth piece of evidence that suggests main clause-initial PPOPs are special clitics rather than topics.

My analysis of the distribution and case properties of main clause-initial PPOPs shows that they pattern (a) very like their non-initial left-of-P counterparts, and (b) differently from what we would expect if they were topics. There are also no examples in which a personal pronoun subject separates an initial PPOP from the finite verb, which would have given strong support for a topic analysis. The best evidence for PPOP topicalisation comes from examples with apparent subject-verb inversion. There are 24 such examples in all, although only two occur in an uncoordinated main clause, where non-subject topicalisation occurs most frequently in general. Taken together, the empirical evidence adds up to a strong case for treating the main clause-initial PPOPs as special clitics rather than topics. To this may be added the fact that clause-initial placement of nominal counterparts is extremely rare. The paucity of examples is already expected on the basis of Taylor (2008: 343, fn. 1), who quantifies the number of nominal objects situated somewhere to the left of a governing preposition (i.e. not necessarily in

---

[22] First and second person forms do not distinguish between dative and accusative. There are only 33 genitive PPOPs in total.

[23] Dative-favouring prepositions are identified in Chapter 3.

clause-initial position) at about 100 in the YCOE, and on the basis of van Kemenade (1987: 152), who found no clause-initial examples in her sample, but quantification of clause-initial examples in a large corpus such as the YCOE has hitherto been lacking.

The YCOE provides just fifteen examples of a (main or subordinate) clause-initial nominal object of P. In each case, left-of-P placement of the nominal is certainly unusual but their status as topics is not the only possibility. Two appear to involve HTLD, e.g. (37).[24] Although the resumptive element is a demonstrative rather than personal pronoun, I assume that clause-initial placement of the pronoun is tied up with the left dislocation of its co-referent.

(37)　Ða　ðe　þurh　ungehyrsumnysse oððe geleafleaste deafe wæron.
　　　those that through disobedience　or　unbelief　deaf　were
　　　þam he **on** ageat　andgites　hlyst
　　　those he in　poured knowledge's hearing
　　　'Those that were deaf through disobedience or unbelief, he poured into them the hearing of the knowledge'
　　　　　　　　　　　　　　(cocathom2,ÆCHom_II,_43:326.234.7325)

In another four examples, the element parsed as the preposition in the YCOE could instead be interpreted as a verbal prefix with the nominal re-interpreted as the object of the prefixed verb. We have already seen one of these four examples, at (8), where I suggested that the initial object could be interpreted as the object of *tobringan*. Given the referent of the subject in the second example, (38), I think it more likely that the clause has *midsiðian* 'to accompany, associate with' rather than *siðian* 'to go, travel' plus *mid* 'with', as the YCOE parse would have it.

(38)　**Þæt ilce　wundor** in þære spræce þæs　æþelan weres eac oðre
　　　the　same wonder in the　tale　of-the fine　hero　also other
　　　wundru　　　wæron **mid** siðiende
　　　wonders-NOM were　associated

_____

[24] The other HTLD example occurs earlier in the same text, at cocathom2,ÆCHom_II,_12.2: 121.398.2662.

'Other wonders were also associated (with) that same wonder in the tale of the fine hero'

<div align="right">(cogregdC,GDPref_and_4_[C]:28.302.5.4489)</div>

The accusative case of the initial NP in (38) adds weight to this analysis. *Mid +* accusative is an Anglian feature (Mitchell 1985: §1195) and so would not be entirely unexpected in this Anglian-influenced version of *Gregory's Dialogues*. However, *mid* + accusative occurs with any regularity in the prose only in *Bede*, and even there *mid* + dative is more common (Mitchell *ibid*).

I give the third example in its surrounding context to show that the morpheme tagged as a preposition could be the prefix of *togeðeodan* 'to adhere, cling to' (Clark Hall 1960), as my gloss and translation assume.[25]

(39)   He stod   in his gebede ealne dæg & **þam dæge** þa æfterfylgendan
        he continued in his prayer all   day and the day   the after-following
        nihte he **to** geþeodde. Eac swylce þone æftran dæg mid his nihte
        night he adhered.   likewise   the   next day with its night
        unwerig   on benum he þurhstod
        unwearying in prayers he continued
        'He continued in his prayer all day, and the following night to the day he persisted. Likewise the next day with its night he continued, unwearying in prayers'

<div align="right">(cogregdC,GDPref_and_3_[C]:14.200.5.2593–5)</div>

According to the YCOE's parse of (39), *to* is the head of a temporal PP that co-occurs with *ðeodan* 'to join, associate (with), attach or subject oneself to: come to: engage in'. Temporal relations can be expressed in Old English by case forms alone (Mitchell 1985: *passim*), so *þam dæge* need not necessarily be governed.

The fourth example, (40a), could involve either *ofaslean* 'to smite off' (Clark Hall 1960) or alternatively *aslean* 'to strike, cut' plus adverbial *of* 'off'. The example at (40b), in which *of* precedes but is not adjacent to the nominal

---

[25] Toller (1921) also lists *togeðeodan* (without a definition) as a derivative of *geðeodan*.

object, defends a non-prepositional interpretation of *of* in (40a), since right-of-P objects always occur immediately to the preposition's right.

(40) a. Gif **men** sie lim **of** aslegen

      if   man be limb off cut

      'If a limb should be cut off a man'

                                  (colaece,Lch_II_[1]:38.8.1.1232)

    b. ... þaþa he of asloh þæs forscyldigan eare

       when he off cut   the   wicked's   ear

      '... when he cut off the wicked one's ear'

                            (colwstan1,ÆLet_2_[Wulfstan_1]:190.258)

Prepositions with null objects were discussed earlier in this chapter, and the example at (41), from *Lacnunga*, comes from the genre in which they predominate, i.e. medicinal recipes and prayers. The emboldened preposition is interpreted with a null object in Cockayne (1866: 25) and Grattan & Singer (1952: 125), both of which treat the emboldened clause-initial NP as a constituent of the preceding clause.

(41)   ... &   do ceac innan in ða buteran, genim þonne ænne sticcan [...] Styre

       and put jug   inside in the butter   take then   a   spoon [...] stir

      þonne mid ðy sticcan ða buteran, **eal þæt fæt**,   ðu   sing **ofer** ðas

      then   with the spoon the butter   all the vessel you recite over the

      sealmas, Beati immaculati, ælcne ðriwa ofer, &   gloria in excelsis deo, &

      psalms *Beati immaculati* each   thrice over and *gloria in excelsis deo* and

      Credo in deum patrem

      Credo in deum patrem

      '... and put the butter into a jug. Then take a spoon [...] Then stir the butter with the spoon, the entire vessel (of it). Recite over (it) the psalms, each thrice over, *Beati Immaculati* and *Gloria in Excelsis Deo* and *Credo in deum Patrem*'

                          (colacnu,Med_3_[Grattan-Singer]:63.31.367–72)

Six of the remaining eight clause-initial nominal objects of P occur in a subordinate clause. One, (42), is given in van Bergen (2003: 199, ex. 19) to show how the nominal can escape a topic analysis according to van Kemenade (1997b).

(42)　　... oþ **þam burgwarum** com　mara fultum **to** utan　　to helpe

　　　　　until the　citizens　　came more help　　to outside to help

　　　　'... until more help came to the citizens from oustide as help'

　　　　　　　　　　　　　　　　　　　(cochronA-2c,ChronA_[Plummer]:921.43.1305)

Van Kemenade (1997b) argues that what looks like embedded topicalisation in Old English occurs in two particular contexts only. One is where the clause is the complement of a bridge verb, in which case the initial XP can be accounted for in the same way as main clause topics. The other context is when the predicate is unaccusative, i.e. has no external argument, including passives and verbs such as *cuman* 'to come', e.g. (42), *gan* 'to go' and forms of 'to be', (*ibid*: 332–8). In such cases, the nominative argument can be licensed VP-internally and may remain there. This frees up Spec-IP — normally the Case position for subjects according to van Kemenade (1987, 1997b) — which may then be filled by a non-subject constituent (*ibid*: 338). Accordingly, as long as Spec-IP is not identified as the topic position, as in Van Kemenade (1987, 1997b), a topic analysis of examples like (42) can be avoided. Van Kemenade's (1997b) argument could also be extended to the other five examples with a subordinate clause-initial nominal object of P, as each co-occurs with an unaccusative predicate.[26] However, for Pintzuk (1991, 1996) and Kroch & Taylor (1997), in which topicalisation is obligatory in subordinate (as well as main) clauses, a topic analysis of the initial nominal in the six examples represented by (42) is unavoidable.

A topic analysis of the initial nominal objects of P in the last two examples, given at (43), is similarly unavoidable in Pintzuk (*ibid*) and Kroch & Taylor (*ibid*). In both examples the predicate is unaccusative, so according to van Kemenade (1997b) the subject may be VP-internal. However, as each example

---

[26] Three examples involve *beon* 'to be' (cogregdC,GD_2_[C]:21.145.34.1750, cogregdH,GD_ 2_ [H]:21.145.35.1430 and cosolilo,Solil_1:30.17.396). One involves *cuman* 'to come' (cocathom1, ÆCHom_I,_17_[App]:540.171.3303) and the other involves *gesittan* with the sense 'to appear' (coherbar,Lch_I_[Herb]:1.15.63).

involves a main clause, the finite verb must be VP-external, which means the initial NP must be in topic position.

(43) a. **Ðam folce**   eode **ætforan** symle   Godes wolcn swilce ormæte   swer
the   people went before   always God's cloud like   immense pillar
'God's cloud went ever before the people like an immense pillar'

<div align="right">(cocathom2, ÆCHom_II,_12.1:113.110.2453)</div>

b. Ac **þisne** þa   sona   færinga   **in** geeode seo wræcenda gast
but this   then immediately suddenly in went   the avenging spirit
'But then directly forthwith the avenging spirit then entered into this one'[27]

<div align="right">(cogregdC,GDPref_and_3_[C]:14.200.17.2600)</div>

Against a topic analysis of the eight clause-initial nominal objects of P for which an independent and theory-neutral non-topic analysis has not been identified, is their distribution by clause type. Data in Kohonen (1978: 154) confirm the widely held view that non-subject topicalisation occurs least frequently in subordinate clauses: he found evidence of it in just 61/1,689 (4%) of subordinate clauses.[28] It is surely significant that six of the eight examples represented by (42) and (43) occur in the type of clause where non-subject topicalisation is especially infrequent. Nominal objects situated to the left of a governing preposition are already a problem for theories of Old English syntax. Taylor (2008: 343, fn. 1) quantifies them at about 100, so clearly the majority of problem cases are not situated clause-initially. Since a separate account is already required for examples for which topicalisation is not a possibility, it is conceivable that the position of apparently topicalised examples may be explained the same way. I am therefore not convinced that a topic analysis of the examples at (42) and (43) is necessary.

To sum up the empirical evidence so far: PPOPs never precede a personal pronoun subject; only 24 main clause-initial PPOPs occur in examples where there appears to be subject-verb inversion; there are no clear accusative or

---

[27] *Ingan* 'to go in, enter' occurs frequently in the YCOE but always intransitively or else with a directional PP.

[28] Bech 2001 provides no comparative data.

genitive examples among main clause-initial PPOPs; the number of initial PPOPs in coordinated main clauses is more than twice the number in uncoordinated main clauses; there is only a handful of examples involving a potentially topicalised nominal object of P, and most occur in subordinate clauses where non-subject topicalisation is relatively rare in general. This collection of observations is not what would be expected if objects of prepositions can topicalise in Old English. Accepting that left-of-P nominals are a problem in any event, the data on clause-initial PPOPs fall much more neatly into place if they are viewed as special clitics.

Having concluded that the empirical evidence points quite firmly towards the need for a special clitic analysis of the pronouns in question, we are left with the problem of how the V2 constraint is satisfied in examples like those at (18), (21), (34) and (36). One possibility can be seen in Axel's (2009: 30–3) use of subject *pro* to account for certain V1 constrictions in Old High German (OHG). In her discussion of the V2 constraint in OHG, Axel (*ibid*) shows that existential and presentational constructions do not have an overt expletive subject as is the case in Present Day German, e.g. (44): instead they are realised with V1 word order, e.g. (45).

(44)   Es spielen die Wiener Philharmoniker

      it play   the Vienna Philharmonic

      'The Vienna Philharmonic Orchestra is playing'

<div align="right">(Axel 2009: 31, ex. 28b)</div>

(45)   uuaram thô   hirta   In thero lantskeffi

      were   PARTICLE shepherds in that   country

      'there were shepherds in the same country'

<div align="right">(Axel 2009: 31, ex. 29b)</div>

Axel claims that most OHG V1 examples involve an unaccusative predicate, and suggests that, in such cases, the post-verbal position of the subject could be related to the fact that it is actually the underlying object. This, she suggests, admits the possibility of a covert expletive subject, i.e. *pro*, in first position in (45), co-indexed with the nominative NP, which would then satisfy V2.

Axel's idea basically entails that, when the predicate is unaccusative, the clause can be superficially topic-less but may in fact have an empty subject. This neatly captures van Kemenade's (1987: 43–8) idea of topic-less main clauses — discussed earlier in relation to the example with narrative inversion at (19) — which allowed her to treat the initial pronouns in (21) as clitics. Although van Kemenade (*ibid*) does not restrict topic-less clauses to unaccusative contexts, her three examples with a clause-initial PPOP, i.e. the two at (21) (which she treats as clitics) and the one at (18) (which she treats as topic), and the example of narrative inversion at (19), each involve unaccusative *cuman* 'to come'. A quick look at the predicate in main clauses with an initial PPOP suggests that most are, or are potentially, unaccusative: 21 examples have *cuman*, and others involve, e.g. *beon* 'to be', *nealæcan* 'to approach', *faran* and *gan* 'to go', and there are a couple of passives.

If Axel's suggestion can be maintained, then an expletive *pro* subject in clauses with an unaccusative predicate would, firstly, permit a V2 analysis of narrative inversion that does not require a null operator, i.e. narrative inversion could be analysed as [pro_{TOPIC} V$_{FIN}$ Subject] instead of [Ø$_{OPERATOR}$ V$_{FIN}$ Subject]. Certainly, an operator-initial analysis of narrative inversion is not universally accepted: Roberts & Roussou (2002: 137–41), for example, argue that the appropriate characterisation of a null sentential topic operator is far less obvious than that of the null operators assumed for direct questions and conditionals. Secondly, it would allow the position of the PPOP in (34a), for example, repeated here as (46), to be analysed as a clitic, i.e. [pro_{TOPIC} clitic V$_{FIN}$], analogous to [XP$_{TOPIC}$ clitic V$_{FIN}$] analyses of examples like (20), repeated at (47). I have revised the translation in (46) to signal a pro_{TOPIC} analysis.

(46)   **Him** com stemn **to**, þus clypiende þriwa
        Him came voice to thus calling   thrice
        'There came to him a voice, thus crying thrice'

                                    (coaelive,ÆLS[Peter's_Chair]:87.2330

(47)   min God **me** asende **to** sona   his engel
        my God me sent   to immediately his angel

'my God at once sent to me his angel'

<div align="right">(coaelhom,ÆHom_22:326.3470)</div>

While Axel's idea would appear to provide a way to treat main clause-initial PPOPs in Old English as clitics, it runs into some real and some potential problems. An operator-initial analysis would still be required to account for the post- rather than pre-V_{FIN} placement of clitic pronouns in other V1 constructions, e.g. direct questions, as well as in the presence of an overt operator, such as a *wh*-word or *þa/þonne*. It would also be needed for examples like (34b), repeated here at (48), in which the predicate is not unaccusative.

(48) **Him** cwæð Nichodemus **to**, swiðe þæs ofwundrod

him said Nichodemus to, very of-this astonished

'Nichodemus, very astonished by this, said to him'

<div align="right">(coaelhom,ÆHom_13:10.1886)</div>

And if a covert topic cannot account for clause-initial placement of the PPOP in (48), it is difficult to maintain that it accounts for clause-initial placement of the PPOP in (46). In addition, Axel's idea would seriously undermine van Kemenade's (1997b) analysis of (what appears to be) embedded topicalisation: if Spec-IP hosts a subject *pro* when the verb is unaccusative, then Spec-IP would not be available for non-subject constituents. Consequently, placement of anything other than the subject or a clitic immediately after the complementizer should not be possible when the embedded predicate is unaccusative, contrary to evidence in van Kemenade (*ibid*).

Since the empirical evidence very strongly suggests that clause-initial PPOPs are special clitics in main as well as in subordinate clauses, I include them in the sample to be analysed for variation between right- and left-of-P placement, despite the absence of a robust theory to explain how the V2 condition is satisfied in examples like (48). This type of example has been somewhat neglected in the relevant theoretical literature, but I have shown that their correct syntactic analysis has implications far beyond that which has been previously recognised.

Chapter 3 Derivation of data

3.1 Introduction

This chapter surveys the source of my data and provides the necessary background to, and the results of, my univariate analyses of variation in PPOP placement. The corpus from which my data is taken is introduced in section 3.2, where I additionally provide some evidence that the identification of prepositional phrases by the corpus editors yields a sufficiently reliable set of data for my purposes. The next three sections justify and describe the dimensions of the statistical model I use to analyse variation in PPOP placement. All variables but one are included on the basis of what has been observed in previous close studies of PPOP placement in Old English prose. These studies are contextualised in section 3.3. Section 3.4 quantifies the number of PPOPs included in the study and gives proportions for the two variants of interest, i.e. left-of-P and right-of-P. The independent variables are then presented within five broad groupings in section 3.5. Section 3.5.1 deals with variables relating to the pronoun, section 3.5.2 deals with variables relating to the PP, section 3.5.3 deals with variables relating to the clause and section 3.5.4 deals with extra-linguistic variables. Section 3.5.5 discusses a number of variables for which the data are not encoded. For each variable discussed in 3.5.1–3.5.4, I explain how the data are classified and show how the data distribute according to these classifications. The chapter concludes with a summary of knockout (or near knockout) factors, i.e. factors that, when present, correlate in at least 95% of cases with right-of-P placement only or with left-of-P placement only.

3.2 Materials

3.2.1 Corpus

The York-Toronto-Helsinki Parsed Corpus of Old English Prose (the YCOE) (Taylor *et al*. 2003) contains approximately 1.5 million words of running prose within 100 text files. Each text file represents a syntactically annotated version of

a scholarly edition of a particular Old English version of a particular text.[29] It is not the largest corpus of Old English available: that title belongs to the Dictionary of Old English corpus (diPaolo Healey 2009), which contains over 3 million words, including 2.1 million words of prose plus 0.9 million words from interlinear glosses, poetry, glossaries and inscriptions. The YCOE is, however, the only corpus of its type to be syntactically annotated. To understand how the YCOE's text files relate to the primary linguistic evidence, i.e. the Old English manuscripts, we need to recognise four levels of representation, namely: the texts; the manuscripts which represent the texts and which constitute the primary linguistic evidence; the scholarly editions which represent the manuscripts; and the YCOE text files which represent the scholarly editions.

In the simplest cases, a non-branching line can be drawn between text and text file. For example, the relationship between *Adrian and Ritheus* (a text) and coadrian (the corresponding text file) is mediated by Cross & Hill's (1982: 35–40) edition of the version found in London, British Museum, Cotton Julius A.II, an 11th century manuscript. In other cases, the relationship between text and text file is more complex. Firstly, one text may be represented by more than one text file. This is the case, for example, with *Gregory's Dialogues*, which is represented by two text files, cogregdC and cogregdH. CogregdC represents (an edition of) an 11th century copy of Bishop Wærferth's late-9th century translation. Wærferth's translation was subsequently revised by an unknown reviser, and cogregdH represents (an edition of) an 11th century copy of that revision. So *Gregory's Dialogues* is represented twice in the YCOE, but each text file represents (an edition of) a linguistically distinct version. The same is true for other texts which are represented more than once in the YCOE: if more than one copy is included, it is because the copies differ significantly in terms of date and/or dialect, if not also in terms of content. Secondly, one text file may represent (an edition of) more than one manuscript versions of a particular text. This is the case where the scholarly edition represented by the text file is a composite edition, i.e. an editor's version of a text which has been (re)constructed from two or more incomplete versions. For example, the Old English *Bede* is represented in the YCOE by

---

[29] A full list of the YCOE's text files and associated scholarly editions can be found at http://www-users.york.ac.uk/~lang22/YCOE/info/YcoeTextFile.htm.

cobede. Cobede is a syntactically-annotated version of a composite edition by Miller (1890–98), who takes the most authoritative extant Old English manuscript version as the base, with missing material taken from three later Old English manuscript versions.

Complicating the linguistic analysis of the material represented by the YCOE text files is the fact that most of the associated manuscripts are of unknown provenance. Although most of the manuscripts are reliably dated to the Old English period by Ker (1957) or, for charters and wills, Sawyer (1968), with a few dated to the early Middle English period by Ker, most are copies, many are written in more than one hand and few hands can be attributed to a named scribe. Consequently, it is usually far from clear how many and which 'sort(s)' of Old English a particular manuscript represents. In addition, the texts represent many different genres, e.g. homilies, annals, biblical works, medical texts, laws and rules. This mix is undesigned: the amount of extant prose material is finite so the corpus editors have simply exploited what is available. Whether this mix of genres is problematic for the present study is unknown. Studies of word order differences between genres have, for Old English, largely focused on differences between the language of prose and of poetry (see the discussion in Chapter 1) or on differences between translated and non-translated texts (e.g. Rissanen 2006, Taylor 2008), and there are simply no generalisations to be made. It is not even clear what categories of sub-genres ought to be recognised, nor how a genre effect could be differentiated from what may potentially be a distinctive style or register of an individual scribe or scriptorium. Such issues are part and parcel of the YCOE, and it is for the corpus users to decide how to handle them. My own approach to extra-linguistic variables is outlined in section 3.5.4.

It is also the case that scholarly editions can and do differ in quality as well as in the conventions employed by their editors to indicate particular features of the language of their base materials. Lacking the necessary resources to check all of the examples cited in this thesis against the notes and apparatus of the editions from which they derive, I have consulted the base editions only for a subset of crucial examples dealt with in Chapter 4, as I will indicate.

3.2.2 Identifying Prepositional Phrases

The ready availability of a large parsed corpus of Old English prose is an enormous advantage to studies of Old English word order. Using CorpusSearch 2 (Randall 2005), users can quickly identify and extract all clauses containing (or not containing) a particular linear configuration of constituents from all or any of the YCOE's 100 text files, leaving more time for the analysis of the results. Of course, the reliability of the results depends on the reliable classification of individual constituents by the corpus editors, which to some extent depends on the quality of the edition of the particular version of the text from which each text file was compiled. Without direct access to the original manuscripts, we can never be absolutely certain that there is a one-to-one correspondence between the linguistic units identified in a given text file and those of the base manuscript, or that the corpus editors' analysis is the only possibility, but without resources such as the YCOE it would be much more difficult to conduct the present type of study on such a large scale.

For most parts of speech there is no reason to question the methods of constituent classification employed by the YCOE's editors, but the Old English prepositions pose a particular set of potential problems. These problems, outlined in detail by Mitchell (1978) and summarised in Colman (1991: 56–7) and Miranda-García & Calle-Martín (2010: 90), result from the fact that many of the Old English prepositions are identical in form to certain verbal prefixes and/or adverbs. This formal ambiguity creates the potential for two types of parsing error: verbal prefixes or adverbs could be miscategorised as prepositions, and prepositions could be miscategorised as verbal prefixes or adverbs. For the purposes of my study, both types of error could potentially have quite serious consequences. If some items are incorrectly labelled as prepositions, then my data set would be corrupt. If, on the other hand, some prepositions are incorrectly labelled as prefixes and/or adverbs, the results could be skewed if, for example, this tended to happen more often with certain prepositions than with others. Accordingly, this section evaluates the methods by which the YCOE editors distinguish between prepositions — tagged P — on the one hand, and adverbs and verbal prefixes — tagged ADV and RP respectively— to gauge whether their decisions are reliable.

It should be noted, firstly, that the term 'verbal prefix' is not used by the YCOE editors: instead they use the deliberately neutral term 'adverbial particle' to avoid the finer-grained distinction between separable and inseparable prefixes. This is a sensible approach: differences between the two types of prefix are easily described, but they can be extremely difficult to distinguish in individual cases (e.g. Mitchell 1978, Elenbaas 2006: 105–74). The YCOE is not designed to provide a definitive syntactic analysis of its materials: rather its purpose is to provide a simple and atheoretical analysis from which more detailed analyses may proceed. The 'adverbial particle' category serves this purpose nicely. For my purposes, it is not necessary to distinguish between separable and inseparable prefixes, and I shall henceforth refer to them jointly as adverbial particles or just particles.

The corpus documentation includes an exhaustive list of words eligible to be tagged as adverbial particles, given here in (1). Those formally identical to a preposition are indicated by bold face. The corpus documentation does not explain the criteria for inclusion on this list, but I assume the list identifies all and only those elements that are listed as the first element of a complex verb in one or more of the Old English dictionaries.[30]

(1)   adun(e), **æfter**, aweg, dune, **fore**, **forð**, **fram**, **geond**, **in**, **mid**, niðer, **of**, dune, **ofer**, **on**, **ongean**, onweg, **to**, **þurh**, **under**, up, ut, **wið**, **wiðer**, **ymb(e)**.

When orthographically attached to the front of a verb in the base edition, a word on this list is always tagged as a particle unless the particle+verb combination is one of fourteen specified exceptions that are labelled always as simplex verbs.[31] To identify that a word labelled as a particle is orthographically attached to a verb, the word is tagged RP+, rather than just RP. When <u>not</u> attached to a verb, a word

---

[30] I have not glossed the items at (1) as the meaning of adverbial particles can vary according to the verb they accompany, and the meaning of inseparable prefixes in particular is often unpredictable (Elenbaas 2006: 114–5, 134–6). In addition to the items at (1), *abutan*, *æt* and *oð* are tagged as adverbial particles on several occasions in the YCOE. Whether these are tagging errors or whether the list of particles at (1) is incomplete is unclear.

[31] The exceptions are: *onbidian* 'to remain, wait', *onbryrdan* 'to excite, inspire', *onbyrgan* 'to taste, eat', *oncnawan* 'to understand, know', *ondrædan* 'to dread, fear', *onettan* 'to hasten', *ongierwan* 'to unclothe', *ongietan* 'to grasp, understand', *onginnan* 'to begin, attempt', *onhagian* 'to be possible, fitting', *onmunan* 'to esteem', *onscunian* 'to shun, avoid', *ontendan* 'to kindle, set fire to', *onþracian* 'to fear, dread'.

on this list is classified as a particle unless it is deemed to be transitive in the context in which it occurs, in which case it is classified as a preposition. The corpus documentation does not provide an exhaustive list of words eligible to be tagged as an adverb, so I assume that the corpus editors followed standard dictionary listings. The RP+ label always takes precedence over the ADV label where the morpheme could be interpreted either way. The RP label also takes precedence over the ADV label unless the word clearly functions adverbially.

Elements tagged RP+ are more likely than elements tagged either RP or ADV to be misidentified prepositions since RP+ elements are the only morphemes whose part-of-speech is determined purely by reference to word division. Word division is not a reliable indicator of word-hood in Old English (e.g. Hough 1991): spaces may appear in unexpected places, e.g. between the elements of a compound or at syllable boundaries; they may fail to appear where they might be expected, i.e. two or more independent words may be written continuously, especially if they are short; and sometimes the space is so narrow that it is hard to tell whether a space was intended. In addition, not every scholarly edition faithfully represents word division in its original materials: editors may add or remove spaces in a scholarly edition — sometimes silently — to reflect their particular interpretation of the text's 'words'.

The use of word division rather than potential transitivity for distinguishing between 'RP+' words and prepositions gives rise to some contrasting pairs, as in (2). In the base edition for (2a), *on* and *locige* are written as one word; consequently, *on* is tagged 'RP+' and the object pronoun is parsed as the object of a prefixed verb. In the base edition for (2b), by contrast, *on* and *locige* are written as separate words; consequently, *on* is parsed as the prepositional governor of *him*.

(2)   a. … ðæt he him onlocige

         that he him on looks

    '… that he should look upon them'

                                                                    (cocura,CP:17.111.18.745)

    b. … þeah ðe þu   **him on** locige

         though you him on look

'… though you should look upon him'

<div align="right">(coaelhom,ÆHom_13:156.1958)</div>

There is no easy way to determine how the examples at (2) should be parsed. The dictionaries do not help: Clark Hall (1960) lists *onlocian* 'to look on, behold', Bosworth & Toller (1898) and Toller (1921) do not, but all three cite *locian* 'to look, gaze' and *on* prep. 'on, upon'. We could look for other collocations of *on* and *locian* to see what word orders are attested when the object is nominal: nominal objects almost always appear to the immediate right of a governing preposition, so unless nominals occur to the immediate right of *on* in the presence of *locian* sufficiently often to support a PP analysis, we could probably conclude that the YCOE's parse of (2a) is the correct one. This is wholly impractical, however. Firstly, the same sort of test would have to be done for a huge number of combinations of preposition/particle and verb. Secondly, there is no guarantee that such a test would be conclusive: there may be no diagnostic examples; there may be too few to draw firm conclusions; or all examples may be structurally ambiguous between the two possibilities.

In order to gauge the proportion of words tagged as adverbial particles, i.e. RP or RP+, or as adverbs, i.e. ADV, that could potentially be prepositions, I have used the following five criteria to identify those that are highly unlikely to be prepositions:

1. ***no lexical ambiguity***: certain words tagged RP, RP+ or ADV do not share their form with an Old English preposition. This includes the unemboldened particles given at (1) and most of the frequently occurring adverbs such as *ða* 'then', *ðonne* 'then', *ðus* 'thus', *eft* 'again', *forðam* 'therefore', *her* 'here', *nu* 'now', *oft* 'often', *sona* 'immediately', *swa* 'so' and *swilce* 'as'.

2. ***morphological complexity***: prepositions do not undergo affixation, but some adverbs do. Consequently, words tagged as the comparative or superlative form of an adverb, e.g. *swiðor*, *swiðost*, as well as derived adverbs ending -*lice*, can be discounted as potential prepositions.

<div align="right">71</div>

3. ***no collocating object***: if the clause lacks an object, an RP, RP+ or ADV element is unlikely to be a preposition (although cf. the third type of 'objectless' preposition noted in section 1.5).

4. ***word order***:

(a) when an object is situated to the right of a governing preposition, it is always situated immediately to its right. Thus an RP, RP+ or ADV element is unlikely to be a preposition if it is non-adjacent to a following object, as in (3).

(3)   Hi   sceoldon þa   **under**hnigan nacodum swurde

they should   then under-fall   naked   sword

'They were then to submit to the naked sword'

                                      (coaelive,ÆLS_[Sebastian]:28.1227)

(b) more than 99.9% of nominal objects immediately follow a governing preposition. Thus an RP, RP+ or ADV element is unlikely to be a preposition if the nominal object precedes it, as in (4).

(4)   On ic þa   ða wynstran dælas Indie   wolde **geond**feran

and I   then the lefter   parts of-India would through-travel

'And I then wanted to traverse the lefter parts of India'

                                             (coalex,Alex:26.7.312)

(c) the verbal negator *ne* always immediately precedes a finite verb and its inseparable prefix (Mitchell 1978: §19, 1985: §1073, Elenbaas 2006: 120). We can therefore be confident that an RP or RP+ element is not a preposition in configurations illustrated by (5).

(5)   ... gif him lichoman untrymnis ne **wið**stode

   if   him bodily   infirmity   not against-stand

'... if bodily infirmity did not obstruct him'

                                     (cobede,Bede_3:19.242.30.2490)

(d) infinitival *to* almost always immediately precedes an inflected infinitive and its inseparable prefix (Mitchell 1978: §19, 1985: §1073, Elenbaas 2006: 112–3). We can therefore be certain that an RP or RP+ element cannot be a preposition in configurations illustrated by (6).

(6)   Ymb þone timan wæs gegaderad III   hund   biscepa & eahtatiene,
about that   time   was gathered   three hundred bishops and eighteen
hiene to **ofer**flitanne
him   to over-strive
'About that time, three hundred and eighteen bishops and were gathered
to confute him'

                        (coorosiu,Or_6:30.149.11.3159)

5. ***pre-modifying function***: certain RP elements are parsed as belonging to a PP, as in (7a), and certain ADV elements are parsed as the modifier of another adverb, as in (7b). Such elements are also unlikely to be prepositions.

(7) a. &   brohton eall **in** to Lundenbyrig
and brought all in to London-town
'and brought all {in to/into} London town'
                  (cochronA-2a,ChronA_[Plummer]:894.48.1054)

b. ... forþan heo hit heold **to** feste wið   hine
because she it held too strictly against him
'… because she kept it too strictly from him'
                  (cochronE,ChronE_[Plummer]:1042.5.2142)

Using these criteria, approximately 88% of the 11,000+ words tagged RP+, 96% of the 3,800+ words tagged RP and 97.5% of the 72,000+ words tagged ADV can be discounted as potentially prepositional. Of the remaining 1,300 or so RP+ elements, all but 39 of the particle+verb combinations with which these words are associated are listed as prefixed verbs in one or more of the standard Old English dictionaries, i.e. Clark Hall (1960), Bosworth & Toller (1898), Toller (1921), and together these 39 combinations account for just 48

individual examples. Of course, just because a particle+verb combination is listed in one or more of these dictionaries is not conclusive proof of its status as a prefixed verb of Old English or in the example in question, but it does provide reassurance that the YCOE editors' analysis of the vast majority of these 1,300 words as adverbial particles is reasonable. I have not examined the remaining 150 RP elements or the remaining 1,850 ADV elements individually, but I am confident that a very significant proportion of each type is unlikely to be re-analysable as a preposition: either because the object's case is not what is normally governed by that preposition, as in (8) (prepositional *to* 'to' usually governs dative), and/or because the word is clearly adverbial, e.g. (9).

(8)   Þæt ilce biþ nyttol ices   slite oþþe hundes gif hit   man sona
      the same is   useful frog's bite or   hound's if   it-ACC one immediately
      to deð
      to does
      'The same (treatment) is useful for bite of frog or hound if one applies it immediately'

                                  (colaece,Lch_II_[1]:35.4.6.1065)

(9) a. Dryhtyn, Dryhtyn, læt us **in**
      Lord   Lord   let us in
      'Lord, Lord, let us in'

                              (cowsgosp,Mt_[WSCp]:25.11.1739)

   b. &   asende hi   **forð** mid his mannum
      and sent   them forth with his men
      'and sent them forth with his men'

                            (cogregdH,GD_1_[H]:10.81.14.816)

   c. ... eallum þam mannum þe   him **ær**   abulgon
       all   the   people   that him previously angered
      '… all the people who previously angered him'

                          (coaelive,ÆLS_[Ash_Wed]:254.2848)

In conclusion, while some RP, RP+ or ADV elements may be compatible with a prepositional analysis, I am satisfied, firstly, that the proportion of such

elements is very small and, secondly, that the analysis given by the corpus editors for this small proportion is at least plausible, if not also the most appropriate.

Unfortunately, it is impossible to undertake a similar evaluation of the extent to which elements labelled as prepositions in the YCOE are compatible with either a particle or adverb analysis. 98% of words labelled P share their form with a particle or adverb, and there is no position in which a preposition can appear relative to its object and to other elements in the clause that would preclude treating the preposition as a particle or adverb. In short, there is no way to settle the matter without examining almost all examples individually. That said, I am confident that the P label has been used judiciously: having examined, for various purposes, hundreds of words and phrases labelled as PP constituents by the corpus editors, I have encountered no cause for concern, and for the genuinely ambiguous cases I am confident that in the majority of cases the application of the P label is entirely plausible. Overall, I conclude that constituents parsed as prepositional phrases by the YCOE editors are a suitably reliable resource for the investigation of PPOP placement.

3.3 Previous studies

There already exist a number of quantitative studies of various aspects of PPOP placement in Old English prose. Kitson (1996) provides an analysis for each of the major forms types of Old English 'between', which reveals two intriguing asymmetries that I explore in detail in Chapter 4. Ogura (1991, 1992) provides a similar analysis for objects of *cweðan to* 'to say to' constructions. Her findings are examined in Chapter 6.

As part of an evaluation of the syntactic status of personal pronoun objects in Old English, Harris (2006: 35–6) compares the placement of non-reflexive PPOPs in versions of texts composed in the early Old English period to their placement in texts composed in the late Old English period. He finds, firstly, that the overall frequency of left-of-P placement varies little according to text composition date. However, by distinguishing two left-of-P variants, Harris also finds that left-of-P PPOPs are separated from their governor as often (N=339) as not (N=304) in the early texts, whereas in the late texts non-adjacent left-of-P PPOPs (N=445) are half as frequent as adjacent left-of-P PPOPs (N=870). In

other words, there appears to be an increasing tendency to place left-of-P PPOPs immediately to the preposition's left over the course of the Old English period. As explained in Chapter 2, all left-of-P variants are treated as equivalent in this thesis, although the relationship between date and left-of-P placement in general is explored in this study. It was earlier noted that Harris (*ibid*: 37–9) also found left-of-P placement to be extremely rare when the PP is unambiguously outside the VP domain, i.e. when in extraposition or embedded in a complex NP, although the number of examples involved is rather small. PP placement is another variable included in this study.

Miranda-García & Calle-Martín (2010) is an exploratory study of factors contributing to left-of-P placement, but their focus on left-of-P placement to the complete exclusion of right-of-P placement severely limits the value of their findings for the present study. For example, in calculating left-of-P frequency in individual texts, they normalise their arithmetic counts to a common base of N (left-of-P) per 10,000 words of text (*ibid*: 95). A selection of their normalised frequencies are given in column two of Table 3.1. Column three shows that when left-of-P frequency is expressed instead as the proportion of all PPOPs (i.e. N (left-of-P) / [N (left-of-P) + N (right-of-P)] a very different picture emerges.

Table 3.1 Left-of-P frequency (Miranda-García & Calle-Martín 2010)

Text	N per 10,000 words (Miranda-García & Calle-Martín 2010)	As % of all PPOPs
West Saxon Gospels, Mark	26	15%
Ælfric's Supplementary Homilies	25	37%
West Saxon Gospels, Luke	24	21%
Lives of Saints	16	43%
West Saxon Gospels, Matthew	14	11%
Cura Pastoralis	13	28%
Orosius	10	41%
Bede	9	16%
West Saxon Gospels, John	8	3%
Alexander's letter to Aristotle	4	41%

This comparison shows very clearly that a text may have a relatively large number of left-of-P PPOPs per 10,000 words but a relatively low frequency of left-of-P placement, e.g. *West Saxon Gospels, Mark*. Conversely, another may have a relatively small number of left-of-P PPOPs per 10,000 words yet have a relatively high frequency of left-of-P placement, e.g. *Alexander's letter to Aristotle*. Miranda-García & Calle-Martín's finding that the majority of left-of-P PPOPs are (a) third person and (b) dative also means little without knowing how this compares to right-of-P PPOPs. Nevertheless, as we will see, grammatical person and pronoun case do indeed appear to be an important part of the story of PPOP placement in Old English.

Only three published studies attempt a quantitative analysis of a large and broad sample of data from Old English prose. The results of these three studies provide the motivation for most of the independent variables included in my statistical model, although observations and claims made elsewhere are taken into account too. Before I identify each of the variables found to correlate with PPOP placement in these various studies, I firstly give a brief summary of these studies so that their findings may be contextualised.

The earliest study is that of Wende (1915), who describes a number of trends and patterns he observed in the placement of PPOPs in *Cura Pastoralis*, *Catholic Homilies I*, *Catholic Homilies II*, *Bede* and entries to AD 871 in the *Parker Chronicle*. The first four of these texts are among the largest included in the YCOE, collectively accounting for some 24% of its total word count. Wende's thorough analysis of such a large volume of data in a pre-computer era is quite remarkable and his findings have proved to be extremely reliable.

The second study is that of Taylor (2008). Taylor approached the data with a specific question in mind: whether placement of PPOPs in translations from Latin is influenced by the usual head-initial word order of Latin PPs. Drawing data from a subset of YCOE texts, Taylor compared frequencies of left-of-P placement in twelve Latin translations to frequencies in seven non-translated texts. The translated texts were compared with their Latin source to determine whether each Old English PP corresponds to a Latin PP. In order to isolate translation effects from the effects of other factors that might also influence PPOP placement, Taylor performed a multivariate analysis in which a number of

'nuisance' factors, identified simply as 'the most likely candidates' (*ibid*: 349), were controlled for. As well as finding a clear effect of Latin PP word order on Old English PP word order, Taylor's analysis revealed most of her nuisance factors to correlate significantly with PPOP placement.

The third study is that of Alcorn (2009), which seeks an explanation for one particular pattern reported by Wende (1915: 76), namely a difference in left-of-P frequency according to the grammatical person of the pronoun. Although I was unable to identify an independent explanation for this difference, in the course of falsifying various hypotheses I identified three factors, not mentioned by Wende or Taylor, which correlate strongly with PPOP placement. Data for this study consisted of all unmodified and uncoordinated PPOPs in the YCOE.

3.4   The dependent variable

Using CorpusSearch 2 (Randall 2005), I identified all PPs occurring in the YCOE whose object consists of a simple, i.e. unmodified and uncoordinated, personal pronoun, and coded each pronoun for the dependent variable, i.e. the position of the pronoun relative to the preposition. As explained in Chapter 2, I assume all left-of-P pronominal objects of prepositions have equal status as special clitics in the syntax. Accordingly, each PPOP is coded as either left-of-P or right-of-P. Left-of-P PPOPs include those that are adjacent to the preposition and those that are not. All right-of-P PPOPs are adjacent to the preposition. The overall frequency of the two variants is summarised below.

Table 3.2 Overall distribution of PPOPs

Left-of-P	Right-of-P	Total
2,775 (28.6%)	6,928 (71.4%)	9,703

3.5 Independent variables

3.5.1 Variables relating to the pronoun

3.5.1.1 Person

A difference in frequency of left-of-P placement of PPOPs between third person pronouns on the one hand, and first and second person pronouns on the other, is noted by Wende (1915: 76–81), Taylor (2008: 350–1, 363), Alcorn (2009) and Miranda-García & Calle-Martín (2010: 98). Each reports, firstly, that a much higher proportion of third person pronouns are situated to the preposition's left in comparison to first and second person pronouns and, secondly, that first and second person pronouns appear to the preposition's left with approximately the same frequency as each other.

Personal pronouns are not tagged for grammatical person in the YCOE, nor is the corpus lemmatised. I therefore used CorpusSearch's 'make lexicon' feature to identify all spellings of all PPOPs. This lexicon showed that these pronouns could be accurately classified for person according to their initial letter. As pronoun form is relevant to three of the coded variables, the relevant parts of the personal pronoun paradigms are given in Table 3.3.

Table 3.3 Paradigm of Old English personal pronoun object form types

Person	Case	Singular	Dual	Plural
First	Acc.	*me*, *mec*	*unc*	*us*
	Dat.	*me*	*unc*	*us*
	Gen.	*min*	*uncer*	*ure*
Second	Acc.	*þe*, *þec*	*inc*	*eow*
	Dat.	*þe*	*inc*	*eow*
	Gen.	*þin*	*incer*	*eower*
Third (masc., fem., neut.)	Acc.	*hine, hi, hit*		*hi*
	Dat.	*him, hire, him*		*him, heom*
	Gen.	*his, hire, his*		*hira*

There are no *y*-spellings of *i*- forms nor *v*- spellings of *u*- forms among the YCOE's PPOPs, but *þ*- forms are often realised with initial *ð*-, and *e*- forms are

sometimes realised with initial *i-*.[32] There are only two instances of *h*-dropping with a third person object pronoun in the entire corpus, both involving *hit*. As neither is the object of a preposition, all *i*-initial PPOPs are unambiguously second person. PPOPs beginning *m-* or *u-* are unambiguously first person forms; those beginning *þ-*, *ð-*, *i-* or *e-* are unambiguously second person forms; and those beginning *h-* are unambiguously third person. The data were coded accordingly.

The univariate results for the independent variable PERSON, given in Table 3.4, confirm previous findings: the frequency with which first and second person PPOPs occur to the preposition's left is (a) almost identical and (b) significantly lower in comparison to third person PPOPs.

Table 3.4 Distribution of PPOPs by grammatical person

	Left-of-P	Right-of-P	Total
First person	170 (10.6%)	1,432 (89.4%)	1,602
Second person	136 (10.2%)	1,194 (89.8%)	1,330
Third person	2,469 (36.5%)	4,302 (63.5%)	6,771
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

3.5.1.2 Case

Wende (1915: 77, 80) reports that most of the left-of-P PPOPs in his sample are dative and none are genitive. A high proportion of dative forms among left-of-P PPOPs is evident from the third person data provided by Miranda-García & Calle-Martín (2010: 98, table 6) and has also been noted in unquantified terms by Visser (1970: §402, fn. 1) and Colman (1991: 77). Mitchell (1978: §27) suggests this simply reflects the preponderance of dative-governing prepositions, but comparative quantitative data published subsequently show this is clearly not so: although dative is indeed the most commonly found case with the prepositions of Old English, data given in Taylor (2008: 350–1) and Alcorn (2009: 443, fn. 21) indicate that dative PPOPs also occur much more frequently than accusative PPOPs to the preposition's left.

---

[32] Most of the *i-* forms are dual, but there are some *i-* spellings of plural forms, e.g. <iow>.

All third person personal pronouns in the YCOE are labelled for case.[33] As accusative and dative are not distinct for first and second person pronouns (apart from the infrequently occurring accusative forms *mec* and *þec*), the vast majority of first/second person PPOPs are not labelled for case. It was, however, possible to disambiguate case for a considerable number of first and second person PPOPs. Although many Old English prepositions govern accusative as well as dative, some strongly favour one case in particular and I assume it is reasonable to disambiguate case for at least some first and second person PPOPs on the basis of what can be shown to be a very strong tendency. Table 3.5 identifies eleven prepositions for which dative is clearly the norm.

Table 3.5 Prepositions for which dative is the norm

	Third person PPOPs		Nominal objects	
	N (dat./acc.)	% Dative	N (dat./acc.)	% Dative
æfter 'after'	171	99.4%	2,202	98.5%
æt 'at'	92	100.0%	1,651	95.1%
ætforan 'before'	56	100.0%	167	98.8%
be 'by, concerning'	264	100.0%	3,215	99.3%
beforan 'before'	133	94.8%	436	96.3%
BETWEEN[34]	297	99.3%	494	82.2%
fram 'from'	266	100.0%	2,428	98.7%
mid 'with'	1,083	94.6%	11,917	98.0%
of 'of'	177	99.4%	5,211	98.3%
to 'to'	2,558	99.8%	12,536	97.9%
togeanes 'against'	113	100.0%	63	93.7%

At least 95% of third person pronouns governed by each of these eleven prepositions are dative (allowing proportions for *beforan* and *mid* to be rounded up), as are at least 95% of the nominal objects with the exception of those

---

[33] The YCOE editors resolve the dat./gen. ambiguity of the third person fem. sg. pronoun *hire* in favour of dative. This is sensible: as noted in Taylor (2003: *Case with prepositions*), few Old English prepositions take genitive, and only *wið* 'against' does so with any real frequency. The number of examples of *hire* as object of *wið* in my sample is negligible.

[34] There are several form types of Old English BETWEEN. I treat them as a single category in Table 3.5, but variants are identified later.

governed by *togeanes* and BETWEEN, although these prepositions still assign dative in a significant majority of instances.[35] On the basis of the dative proportions in Table 3.5, I assume that the vast majority, and probably not less than 95%, of first and second person pronouns governed by these eleven prepositions are dative also. Since it is impossible to identify *which* 5% are likely to be accusative, I have simply coded all first and second person pronouns governed by these prepositions as dative. A total of 1,914 case-ambiguous first/second person PPOPs are thus classified as dative under government by these dative-favouring prepositions, and I assume that no more than 96 (i.e. 5%) are likely to be accusative.

There is only one preposition, *þurh*, for which accusative is the norm, as shown in Table 3.6.

Table 3.6 Prepositions for which accusative is the norm

	Third person PPOPs		Nominal objects	
	N	% Accusative	N	% Accusative
þurh 'through'	207	95.7%	2,817	94.3%

On the basis of the proportions in Table 3.6, I assume that the vast majority — again probably not less than 95% — of first and second person pronouns governed by *þurh* are accusative also. I therefore coded all first and second person pronouns governed by *þurh* as accusative. 54 first/second person PPOPs are classified as accusative under government by *þurh* and I assume that no more than 3 (i.e. 5%) are likely to be dative.

By using this method to disambiguate case, the proportion of first and second person PPOPs uncoded for case is substantially reduced from 99.7% (the 0.3% being genitive forms) to 32.6%. Taylor (2008: 350, fn. 10) disambiguates case for many first and second person PPOPs in her sample by the same principle, although she identifies case-favouring prepositions by reference to proportions among third person PPOPs only, i.e. without reference to proportions for full NPs. For infrequently occurring prepositions, she relies on the case norms identified by

---

[35] I return to the difference in case proportions for the different types objects of BETWEEN in Chapter 4.

Mitchell (1985: §§1178–1219). Consequently, she too assumes first and second person PPOPs governed by *þurh* to be accusative, but her list of dative-favouring prepositions excludes *mid* and — surprisingly — forms of BETWEEN, but includes *abutan* 'about', *ær* 'previously', *butan* 'out(side) of' and *wiðinnan* '(from) within'. As the last four prepositions collectively govern fewer than 100 third person PPOPs in my sample, I did not undertake a detailed analysis of their objects by case so first and second person PPOPs governed by these prepositions remain case ambiguous.

The univariate results for the independent variable CASE, given in Table 3.7, confirm previous findings: the majority of left-of-P PPOPs are indeed dative and there are no left-of-P genitive PPOPs, although there are only 31 genitive examples in total.

Table 3.7 Distribution of PPOPs by case

	Left-of-P	Right-of-P	Total
Dative	2,662 (33.9%)	5,189 (66.1%)	7,851
Accusative	39 (4.5%)	829 (95.5%)	868
Genitive	—	31 (100%)	31
Ambiguous	74 (7.8%)	879 (92.2%)	953
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

Table 3.7 also confirms that the large proportion of dative pronouns among left-of-P PPOPs is not simply due to the much larger proportion of dative PPOPs overall. Mitchell (1978: §27) was evidently convinced that reported associations between left-of-P placement and dative case were of no consequence since he could not 'imagine anyone having the time or the inclination to test these conclusions, even with the aid of a computer.' I am sure he would have been intrigued to see these raw results. The effect of pronoun case on PPOP placement is clearly significant and is discussed in further detail in Chapter 4.

Taylor's results differ slightly from those in Table 3.7, although she found the same overall trend, with 41.1% of dative PPOPs and 2.9% of accusative PPOPs in a left-of-P position. The higher frequency of left-of-P placement calculated by Taylor for dative PPOPs in comparison to that shown in Table 3.7

can be attributed to differences in the way we applied the same principle for disambiguating case and to differences in relative frequencies of individual dative-favouring prepositions in our respective samples.

Where one variant is very strongly favoured in a context that can be described by reference to a single dimension of one independent variable, it is common practice to exclude data occurring in that context from the variation analysis. Such a decision is entirely methodological: where the choice of variant is near categorical, there is simply 'little room for quantitative investigation' (Tagliamonte 2006: 84). The rule-of-thumb recommended by Guy (1988) and Tagliamonte (2006: 86–7) is to exclude all data associated with any factor that favours one particular variant at least 95% of the time. Since more than 95% of accusative and genitive PPOPs are right-of-P, I exclude these 899 pronouns from the analysis of variation, although I return to the linguistic analysis of the 39 left-of-P accusative PPOPs in Chapter 4. The 953 case-ambiguous PPOPs then present something of a problem as they doubtlessly include some dative as well as some accusative pronouns. Since the evidence clearly shows that PPOPs rarely appear in a left-of-P position unless they are dative, I have chosen to exclude all 953 case-ambiguous pronouns from the analysis of variation in pronoun placement. While we may be reasonably confident that most, if not all, of the 77 left-of-P examples are dative, any one of the 876 right-of-P examples could be accusative, and including the (assumed dative) left-of-P examples without including their corresponding dative right-of-P examples would create a skewed sample. As all of these 953 pronouns are either first or second person forms, this means discarding about one-third of the sample of non-third person PPOPs. The two-thirds that are not discarded, however, still amount to a large enough sample to allow PERSON effects to be estimated.

3.5.1.3 Number

Taylor (2008: 350, fn. 9) undertook a univariate analysis of PPOP placement according to grammatical number. Given number ambiguities among third person accusative and dative forms (see the paradigm at Table 3.3), her analysis is confined to first and second person data only. She found left-of-P placement to be about 10% more frequent for plural forms than for singular forms.

Although pronouns are not tagged for number in the YCOE, first and second person forms are easily classified on the basis of their spellings. PPOPs beginning *m-*, *þ-* or *ð-* were coded as singular, *inc*, *incer*, *unc* and *uncer* forms were coded as dual, and all other first and second person forms were coded as plural. Third person PPOPs were treated as follows: *hine*, *hire*, *his* and *hit* forms were coded as singular; *hira* and *heom* forms were coded as plural; and forms of *hi* and *him* were coded as number-ambiguous unless governed by a form of BETWEEN, which requires a semantically plural complement. The univariate results for the independent variable NUMBER are given in Table 3.8. The results are further analysed by person to allow a comparison with Taylor's findings for first and second person PPOPs. Although 1,852 pronouns were earmarked for exclusion from the analysis of variation in PPOP placement in the previous section, each of the univariate analyses provided in this chapter quantify the relationship between PPOP placement and the independent variable in question for the full sample. Pronouns identified for exclusion through the results of the univariate analyses are therefore excluded only from the multivariate analysis.

Table 3.8 Distribution of PPOPs by grammatical number

	Left-of-P	Right-of-P	Total
Singular			
- first/second person	*164 (8.8%)*	*1690 (91.2%)*	*1,854*
- third person	*190 (18.9%)*	*816 (81.1%)*	*1,006*
Total	354 (12.4%)	2,506 (87.6%)	2,860
Dual	6 (25.0%)	18 (75.0%)	24
Plural			
- first/second person	*136 (12.9%)*	*918 (87.1%)*	*1,054*
- third person	*295 (54.7%)*	*244 (45.3%)*	*539*
Total	431 (27.1%)	1,162 (72.9%)	1,593
Ambiguous	1,984 (38.0%)	3,242 (62.0%)	5,226
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

The results for first and second person PPOPs are broadly similar results to those reported in Taylor (*ibid*), except that the difference in left-of-P frequency

between plural and singular in my larger sample is 4% rather than 10%. Third person PPOPs — or rather those for which number can be determined — also occur more frequently in a left-of-P position when plural than when singular in my sample, except the difference, at 36%, is much larger than for non-third person PPOPs. One of the reasons for this sizeable difference is an imbalance according to pronoun case. In the previous section, I noted that PPOPs rarely occur in a left-of-P position unless dative. All but eleven of the 539 third person plural pronouns are dative (98%) compared to 446 of the 1,006 third person singular pronouns (44%).[36] As there is a much higher concentration of dative forms among third personal plural data than among third personal singular data, it is not surprising that the plural data show a much higher frequency of left-of-P placement. As I have already concluded that the analysis of variation in PPOP placement should focus on dative PPOPs only, the imbalance among third person forms in numbers according to pronoun case is of no further consequence.

3.5.1.4 Reflexivity

Pronoun reflexivity is one of the PPOP features included as a potential 'nuisance' factor in Taylor's (2008) analysis of Latin interference effects. Reflexive pronouns do not have a distinctive form in Old English; instead personal pronouns are used reflexively, either on their own, as in (10a), or with a form of 'self' (Mitchell 1985: §265, Traugott 1992: 215), as in (10b). Those modified by 'self' are excluded from my sample (see section 1.5.1).

(10) a. ... þætte good &   yfel bioð symle ungeþwære **betweox him**

      that good and evil are always discordant between them

   '… that good and evil will always be discordant between them(selves)'

(coboeth,Bo:37.113.25.2248)

  b. ... forðæm þu hit hæfst afunden **be þe   selfum**

     because you it   have found   by you self

   '… because you have found it by yourself'

(coboeth,Bo:31.70.27.1315)

---

[36] The paucity of third person accusative plural PPOPs is a consequence of the fact that these forms are ambiguous for number.

Reflexive pronouns are identified by the YCOE editors (and by Mitchell 1985: §§266–77) as non-possessive personal pronouns that co-refer with the subject of the clause. They are distinctively labelled in the YCOE. PPOPs were coded as reflexive or non-reflexive according to the presence or absence, respectively, of this 'reflexive' label and the univariate results for the independent variable REFLEXIVITY are given in Table 3.9.

Table 3.9 Distribution of PPOPs by pronoun reflexivity

	Left-of-P	Right-of-P	Total
Reflexive	399 (43.5%)	519 (56.5%)	918
Non-reflexive	2,376 (27.0%)	6,409 (73.0%)	8,785
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

These proportions are within a few percentage points of those reported by Taylor (*ibid*: 351). Although reflexive PPOPs appear significantly more frequently than non-reflexive PPOPs in a left-of-P position in both our samples, Taylor found that there was no statistically significant *correlation* between pronoun reflexivity and PPOP placement in her sample. In other words, the fact that reflexive PPOPs were more often left-of-P in Taylor's sample was due to their concentration in contexts where left-of-P placement is favoured for other reasons. It will be interesting to see if the same result obtains for my larger sample.

3.5.2   Variables relating to the PP

3.5.2.1 Preposition

Variation in frequency of left-of-P placement according to the particular preposition involved is noted by Wende (1915: 71–3) and Taylor (2008: 351). Both found that *þurh* 'through' never occurs with a left-of-P PPOP in their samples — as did Allen (1980: 316, fn. 58) — although neither found any preposition to occur only with left-of-P PPOPs. Wende (1915: 14) additionally notes that of the doublet forms *be~bi*(*g*) 'by, concerning' and *for~fore* 'before, because of', only *bi*(*g*) and *fore* occur with a left-of-P PPOP. Mitchell (1985: §1185) concurs that *for* always precedes its object, but does not observe the same

for *be* (*ibid*: §§1183–4). Quirk & Wrenn (1957: §141) have suggested that left-of-P placement occurs more frequently when the preposition consists of more than one syllable, while Kitson's (1996: 28–32) analysis of the placement of PPOPs governed by different forms of BETWEEN in the concordance of base material for the Dictionary of Old English (diPaolo Healey & Venezky 1980) reveals left-of-P placement to be much more frequent with *betweonum* forms than with others. Wende (1915: 71, 73) found a similar contrast among forms of BETWEEN.

As the YCOE is not lemmatised, I used CorpusSearch's 'make lexicon' function to identify all spelling variants of all PPOP-governing prepositions in the corpus. This enabled me to identify each preposition that occurs at least 100 times with a PPOP. Prepositions that govern a PPOP less than 100 times are assigned to the 'miscellaneous' category. Following Kitson's observation of a difference in frequency of left-of-P placement of PPOPs according to the form of BETWEEN, I distinguish two categories: *betweonum* and 'between' (other).[37] Following Kitson's (1993: 12) description of the major form types of this preposition, I classify *betweonum* forms as those with two nasal consonants — chiefly <betwynan>, <betweonum> and <betweonan> — and classify all other forms as 'between' (other).[38] The various form types of Old English BETWEEN are discussed further in Chapter 4. Following Wende, I also distinguish *bi*(*g*) from *be* and *for* from *fore*.

The univariate results for the independent variable PREPOSITION are given in Table 3.10, in which form types are listed in decreasing order of frequency. As *bi*(*g*) and *fore* govern less than 100 PPOPs between them, they are included in the miscellaneous category. The relationship between *bi*(*g*) and *be* and between *for* and *fore* is discussed further in Chapter 4.

---

[37] The data in Kitson (1996: 29) actually suggest a three-way contrast in frequency of left-of-P placement according to the form type of BETWEEN, i.e. *betweonum vs. betweo*(*h*)*n vs.* other. As *betweo*(*h*)*n* forms govern a simple PPOP only twice in the YCOE, I have grouped this form type with 'between' (other). Kitson (*ibid*) found 30 PPOPs governed by *betweo*(*h*)*n* forms, but his corpus (diPaolo Healey & Venezky 1980) is much larger than the YCOE as it includes data from poetry and interlinear glosses as well as from prose.

[38] There is one form of 'between' in the YCOE with only one nasal consonant that nevertheless belongs to Kitson's betweonum type. This form, <betweona>, occurs three times but never with a simple PPOP. My classification system is therefore appropriate for my particular purposes.

Table 3.10 Distribution of PPOPs by governing preposition

	Left-of-P		Right-of-P		Total
to 'to'	1,338	(41.5%)	1,889	(58.5%)	3,227
mid 'with'	129	(8.6%)	1,369	(91.4%)	1,498
on 'on, in'	248	(24.6%)	760	(75.4%)	1,008
miscellaneous	199	(31.8%)	427	(68.2%)	626
fram 'from'	141	(30.2%)	326	(69.8%)	467
wið 'against'	70	(17.2%)	337	(82.8%)	407
for 'before, because of'		—	290	(100.0%)	290
be 'by, concerning'		—	277	(100.0%)	277
ongean 'towards, against'	112	(45.3%)	135	(54.7%)	247
betweonum 'between'	229	(94.6%)	13	(5.4%)	242
æfter 'after'	47	(19.5%)	194	(80.5%)	241
beforan 'before'	42	(19.0%)	179	(81.0%)	221
of 'of'	58	(31.5%)	126	(68.5%)	184
þurh 'through'		—	182	(100.0%)	182
æt 'at'	25	(15.2%)	139	(84.8%)	164
'between' (other)	20	(13.4%)	129	(86.6%)	149
ofer 'over'	9	(6.1%)	139	(93.9%)	148
togeanes 'against, towards'	108	(86.4%)	17	(13.6%)	125
Total	2,775	(28.6%)	6,928	(71.4%)	9,703

Table 3.10 identifies three prepositions that are not attested with a left-of-P PPOP in the YCOE: *for*, *be* and *þurh*. As PPOPs governed by these three prepositions show no variation in placement, all 749 are excluded from the variation analysis. It seems reasonable to suppose that PPOPs always follow governing *þurh* because *þurh* almost always governs accusative (see Table 3.6), but in the following chapter we will see indications that there may be more to it than that. As neither *for* nor *be* favour accusative, their failure to appear with a left-of-P PPOP most certainly requires an alternative explanation. This too is considered in Chapter 4. Although left-of-P placement is rare with *ofer*, the majority of pronouns governed by this preposition are independently excluded

from the variation analysis on the basis that only a very small proportion is unambiguously dative.

At the other end of the scale lies *betweonum*, which very strongly favours left-of-P placement of PPOPs, and I exclude these 242 pronouns from the analysis of variation under the 95% rule-of-thumb discussed in section 3.5.1.2. The contrasting results for *betweonum* and 'between' (other) are in line with the descriptions in Wende (1915: 71, 73) and Kitson (1996: 28–32), and are sufficiently striking to merit separate discussion in Chapter 4.

Lastly, Table 3.10 provides little evidence for Quirk & Wrenn's suggestion that left-of-P placement occurs more frequently when the preposition consists of more than one syllable (1957: §141): certainly the highest frequencies of left-of-P placement are exhibited by objects of *betweonum* (95%), *togeanes* (86%) and *ongean* (45%), but objects of *to*, *of*, *fram* and *on* have higher frequencies of left-of-P placement (41%, 32%, 30% and 25% respectively) than objects of *æfter*, *beforan*, 'between' (other) and *ofer* (20%, 19%, 13% and 6% respectively).

3.5.2.2 Coordination

As well as noting that coordinated PPOPs are always situated right-of-P, Wende (1915: 66–8) found that the same is true 'wenn zwei oder mehrere Präpositional-verbindungen, deren Rekta materiell verschieden sind, miteinander irgendwie korrespondieren' ('when two or more preposition phrases, whose objects are materially different, somehow correspond to one another'). As Wende supplies numerous examples, it is possible to determine what he means. His examples indicate he found right-of-P placement to be the rule: when the PP is coordinated with a PP headed by the same preposition in the same clause, as in (11);

(11) & ic sette min wed   **to him** & **to his ofspringe on ecere**
      and I set   my covenant to him and to his offspring in eternal
      gefæstnunge
      protection

'and I will set my covenant with him and with his offspring in eternal protection'

<div align="right">(cocathom1,ÆCHom_I,_6:225.26.1077)</div>

when the PPs are headed by the same preposition and belong to parallel VPs, whether the verb is repeated, as in (12), or not, as in (13);

(12)   Þæt þæt ic **to eow** gecweðe. þæt ic cweðe **to eallum mannum**

        That that I to you say   that I say   to all   men

        'That which I say to you, that I say to all men'

<div align="right">(cocathom2,ÆCHom_II,_40:301.57.6852)</div>

(13)   he wunað **on me** and ic **on him**

        he dwells in me and I in him

        'he dwells in me and I (dwell) in him'

<div align="right">(cocathom2,ÆCHom_II,_15:152.71.3365)</div>

and when the PPs are headed by different prepositions but belong to parallel VPs, whether the verb is repeated, as in (14), or not, as in (15).

(14)   ic nelle   **mid ðe**   faran, ac ic wille faran **to minre cyððe**

        I not-will with you go   but I will go   to my   kinsmen

        'I will not go with you, but I will go to my kinsmen'

<div align="right">(cocura,CP:41.304.12.2025–6)</div>

(15)   and he wunað **betwux us**. and we **mid him**

        and he dwells between us and we with him

        'and he will dwell among us, and we with him'

<div align="right">(cocathom2,ÆCHom_II,_45:339.121.7604)</div>

Wende (*ibid*) provides no examples of coordinated PPs in the same clause headed by different prepositions. The YCOE provides a small number of examples involving a PPOP, and in each case the pronoun is right-of-P, e.g. (16).

(16)   ... forðan þe se sunu is þæs   fæder wisdom, **of   him** and **mid him**

       because   the son   is of-the father wisdom from him and with him

   '... because the son is the father's wisdom, from him and with him'

<div align="right">(coaelive,ÆLS_[Christmas]:35.26)</div>

The YCOE editors' parsing identifies coordinated PPs as those which are coordinated with another PP in the same clause, whether headed by the same preposition, as in (11), or not, as in (16), and those which are coordinated with some other constituent, e.g. an adverbial phrase. The example at (17) is the only one in my sample in which the PP is coordinated with something other than another PP. In this case the pronoun is left-of-P.

(17)   … ealle ða ricu   þe **him under** bioð oððe awer   on neaweste

       all   the mighty that him under are   or   somewhere in proximity

   '… all the mighty who are under him or somewhere in proximity'

<div align="right">(coboeth,Bo:16.34.20.628)</div>

The syntactic annotation of the YCOE's material does not annotate parallel structures in different clauses, as in (12)–(15), and there is no simple way to identify such parallelism.

In sum, PPOPs are coded as belonging to a coordinated PP where the PP conjuncts occur in the same clause, as in (11), (16) and (17). As there is no ready way to identify the other types of coordinated PPs identified by Wende, all other PPOPs are coded as belonging to an uncoordinated PP. The univariate results for the independent variable PP COORDINATION are given in Table 3.11. The example at (17) shows one of the two coordinated PPs with a left-of-P PPOP.

Table 3.11 Distribution of PPOPs by PP coordination

	Left-of-P	Right-of-P	Total
Coordinated	2 (1.7%)	118 (98.3%)	120
Uncoordinated	2,773 (28.9%)	6,810 (71.1%)	9,583
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

These results confirm that right-of-P placement is the norm when the PP is coordinated with another PP in the same clause. All 120 PPOPs belonging to such a PP are therefore excluded from the statistical analysis of variation in PPOP placement and I return to the linguistic analysis of this finding in Chapter 4.

3.5.2.3 Embedding

Wende (1915: 68–9) observes that right-of-P placement is near categorical when the PP is a constituent of a noun phrase, as in (18) and (19). Such PPs were excluded from Taylor's analysis (2008: 351) although she does not say why.

(18)   Ða   æt nehstan se foresprecena cyning self, &   se halga biscop

       then at last   the aforesaid   king   self and the holy   bishop

       Trumwine **mid him** &   monige oþre æfeste weras &   rice liðon on

       Trumwine with him and many   other pious men   and rich sailed on

       ðæt ealond

       the island

       'Then at last the aforesaid king himself and the holy bishop Trumwine with him and many other pious and rich men sailed onto the island'

                                     (cobede,Bede_4:29.368.9.3681)

(19)   Heald þu min wed: &   þin ofspring **æfter þe**

       hold   you my covenant and your offspring after   you

       'Hold my covenant, you and your offspring after you'

                               (cocathom1, ÆCHom_I,_6:224.20.1066)

Harris (2006: 39) makes the same observation as Wende about PPs embedded in a NP. Of 91 examples with a simple personal pronoun object in Harris's sample, just one involves a left-of-P pronoun.

     PPOPs are coded as belonging to an embedded PP where the PP has been parsed by the corpus editors as a sub-constituent of any non-verbal constituent. The univariate results for the independent variable PP EMBEDDING, given in Table 3.12, confirm Wende's findings.

Table 3.12 Distribution of PPOPs by PP embedding

	Left-of-P	Right-of-P	Total
Embedded	3 (1.6%)	181 (98.4%)	184
Unembedded	2,772 (29.1%)	6,747 (70.9%)	9,519
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

All 184 PPOPs belonging to the embedded PP category are therefore excluded from the statistical analysis of variation in PPOP placement and I return to the linguistic analysis of this finding in Chapter 4.

3.5.2.4 PP position

Quirk & Wrenn (1957: §141) claim that left-of-P placement 'is most frequent [...] when it enables the preposition to stand before a verb form', although they provide no quantitative data. This trend can be detected in data provided by Ogura (1991: 276, table 2), who shows left-of-P placement of PPOPs to be almost three times more frequent when the preposition precedes rather than follows the main verb in the prose section of the concordance to the Dictionary of Old English (diPaolo Healey & Venezky 1980), although Ogura's data are limited to PPOPs occurring in *cweðan to* 'to say to' constructions. I reported a similar result for this construction in the YCOE in Alcorn (2009: 445, table 7), where I additionally showed that left-of-P placement is also significantly more frequent when the PP precedes the main verb with other verb + preposition combinations, albeit to a much lesser extent. I further noted that adjacency of preposition and main verb appears to be another relevant factor, at least when the preposition is pre-verbal (*ibid*: 446, fn. 26).

The data were therefore coded, firstly, according to whether the PP precedes or follows the main verb, where I define the position of the PP by reference to the position of the preposition. For example, the PP in (20) is classified as pre-verbal, while the PP in (21) is classified as post-verbal even though the PPOP precedes the verb.

(20)　　&　sume **mid　heom** on Gallia læddon

　　　　and some with them　in　Gaul　led

　　　　'and some (they) took with them into Gaul'

<div align="right">(cochronE,ChronE_[Plummer]:418.1.108)</div>

(21)　　Hwilum　　hi　**him** bæron　**to** gold ond seolfor

　　　　Sometimes they him brought to gold and silver

　　　　'At other times they brought gold and silver to him'

<div align="right">(comart3,Mart_5_[Kotzor]:Ja17,A.20.133)</div>

The data were separately coded according to whether or not the PP and main verb are adjacent. Adjacent PPs are those which immediately precede or follow the main verb, as in (22), as well as those which are separated from the verb by the verbal negator *ne* or infinitival *to*, as in (23), neither of which can be separated from the verb by any constituent other than an inseparable prefix (Mitchell 1985: §§907, 1073, 1599, 1627–9).

(22) a. Þa　færinga　**beforan him** stod　se　eadiga　martir　Sanctus Iuticius

　　　　then suddenly before　　him stood the blessed martyr Saint　Juticius

　　　　'Then suddenly stood before him the blessed martyr Saint Juticius'

<div align="right">(cogregdC,GDPref_and_3_[C]:38.258.1.3724)</div>

　　b. and cwædon **heom betwynan**

　　　　and said　　　them　between

　　　　'and said among themselves'

<div align="right">(conicodA,Nic_[A]:13.3.1.256)</div>

(23) a. ic **eow fram** ne　fare

　　　　I　you　from　not go

　　　　'I will not go from you'

<div align="right">(coaelhom,ÆHom_7:44.1064)</div>

　　b. &　ne　geþristlæce he **mid him** to sittene

　　　　and not presumes　he with him to sit

　　　　'and he should not presume to sit with him'

<div align="right">(cochdrul,ChrodR_1:2.16.119)</div>

<div align="right">95</div>

As PPs embedded under a non-verbal constituent are excluded from the multivariate analysis by virtue of the fact they rarely govern a left-of-P PPOP (see previous section), such data were not coded for the PP's position relative to the main verb. However, a separate category was included for both variables to accommodate PPs that occur in clauses without a main verb, as in (24).

(24) Ne dyde God þis for me, ac **for þe** swiþor
      not did   God this for me but for you rather
      'God did not do this for me, but rather for you'

<div align="right">(coaelive,ÆLS_[Thomas]:393.7789)</div>

The univariate results for LINEAR ORDER OF PP AND V and ADJACENCY OF PP AND V, given below, are in line with previous observations: PPOPs more frequently occur in a left-of-P position when the PP precedes the main verb, and they more frequently occur when the PP is adjacent to the main verb.[39] The combined effects of these two variables is considered in Chapter 6.

Table 3.13 Distribution of PPOPs by linear order of PP and main verb

	Left-of-P	Right-of-P	Total
[PP(…)V]	1,556 (37.9%)	2,545 (62.1%)	4,101
[V(…)PP]	1,209 (22.9%)	4,075 (77.1%)	5,284
Embedded PP	3 (1.6%)	181 (98.4%)	184
Elided main verb	7 (5.2%)	127 (94.8%)	134
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

---

[39] Of the 1,209 left-of-P PPOPs in a [V(...)PP] context, 809 follow the main verb, i.e. the word order is [V(...)PPOP(...)P], and 400 precede the main verb, i.e. [PPOP(...)V(...)].

Table 3.14 Distribution of PPOPs by adjacency of PP and main verb

	Left-of-P	Right-of-P	Total
Adjacent	2,147 (33.4%)	4,276 (66.6%)	6,423
Non-adjacent	618 (20.9%)	2,344 (79.1%)	2,962
Embedded PP	3 (1.6%)	181 (98.4%)	184
Elided main verb	7 (5.2%)	127 (94.8%)	134
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

The data in these tables also indicate that main verb elision is yet another factor which strongly favours right-of-P placement of PPOPs. This too was noticed by Wende (1915: 68–9), although the 134 PPOPs identified in Table 3.13 and Table 3.14 as belonging to a clause with an elided main verb also include a subset of those which Wende classifies as part of a parallel VP, as discussed in section 3.5.2.2. As these 134 PPOPs show minimal variation in placement, they are excluded from the statistical analysis of variation and I consider the linguistic analysis of this correlation in Chapter 4.

3.5.3   Variables relating to the clause

3.5.3.1 Clause type

Taylor (2008) found an independent effect of clause type on probability of left-of-P placement, but the result is difficult to interpret. The difficulty stems from the fact that she provides two sets of results: one for PPOPs governed by *to* and one for all other PPOPs, and the results are not consistent. Taylor found that PPOPs governed by *to* are slightly more likely to occur in a left-of-P position in uncoordinated main clauses than in subordinate clauses (*ibid*: 364), whereas other PPOPs are significantly *less* likely to be in a left-of-P position in uncoordinated main clauses than in subordinate clauses (*ibid*: 351–2). For both sets of data, the probability of left-of-P placement was found to be unaffected by occurrence in a coordinated main clause. Differences according to clause type were not the cause of Taylor's decision to split her data this way: *to*-PPOPs were analysed separately from other PPOPs because of a diachronic difference in left-of-P frequency, as we will see in section 3.5.4.1. It therefore remains to be seen whether there is a

statistically significant correlation between PPOP placement when the data are not divided in this way.

Clause type is an easily encoded factor as each clause in the YCOE is labelled as either matrix, subordinate, coordinated subordinate, infinitival or small. Uncoordinated and coordinated main clauses do not have distinctive labels and are not always easy to distinguish on empirical grounds. The presence of an initial conjunction is a reliable indicator that a clause is coordinated, but a main clause without an initial conjunction is not necessarily uncoordinated. Elision of the subject is also not a reliable indicator of a paratactic relationship as uncoordinated main clauses can occur with an unexpressed subject (Mitchell 1985: §§1506–16). It is therefore often unclear whether two adjacent main clauses should be interpreted as independent or asyndetically paratactic clauses (e.g. Traugott 1992: 220, Mitchell 1985: §§1690–708). There is good reason to attempt to distinguish between coordinated and uncoordinated main clauses, however. Should it transpire that there is a statistically significant correlation between PPOP placement and clause type, then we might reasonably question whether this is connected to other word order differences between different types of clause. One very obvious difference concerns the position of the finite verb, which shows a strong asymmetry between main clauses (where verb-second is more likely than verb-final) and subordinate clauses (where verb-final is more likely than verb-second). It is generally agreed, however, that finite verbs are much more likely to be in second position in uncoordinated main clauses than in coordinated main clauses (e.g. Mitchell 1985: §§904–5, Traugott 1992: 277, Koopman 1995, Pintzuk & Haeberli 2008). I therefore distinguish between 'main clause conjunct', which are those main clauses with an initial conjunction, and 'main clause', which lack an initial conjunction and may or may not be independent.

Following Taylor (2008: 351), I additionally distinguish PPOPs occurring in participle phrases. The internal syntax of participle phrases is essentially sentential (Denison 1993: 372–80, Taylor 2003: *Participle Phrases*) and they are identified in the YCOE as adjuncts that are headed by a participle (past or present) and that have an adjectival function, e.g. the underlined phrase in (25).

(25) Ic stande on his gesihðe **to him** me_____gebiddende

I   stand   in   his sight   to him   me-REFLX praying

'I stand in his sight, praying to him'

(cocathom1,ÆCHom_I,_38:518.326.7785)

The univariate results for the independent variable CLAUSE TYPE are given in Table 3.15. Given that PPs governed by a non-verbal element and those co-occurring with an elided main verb are to be excluded from the main statistical analysis, these PPs were not coded for clause structure.

Table 3.15 Distribution of PPOPs by clause type

	Left-of-P	Right-of-P	Total
Main clause	722 (28.2%)	1,836 (71.8%)	2,558
Main clause conjunct	894 (33.5%)	1,776 (66.5%)	2,670
Subordinate clause	920 (27.2%)	2,467 (72.8%)	3,387
Subordinate clause conjunct	119 (29.7%)	282 (70.3%)	401
Infinitival clause	80 (32.1%)	169 (67.9%)	249
Participle phrase	24 (26.7%)	66 (73.3%)	90
Small clause	6 (20.0%)	24 (80.0%)	30
Embedded PP	3 (1.6%)	181 (98.4%)	184
Elided main verb	7 (5.2%)	127 (94.8%)	134
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

These results indicate there is little difference in the frequency of left-of-P placement between each of the clause types for which there are at least 100 examples involving a PPOP.

3.5.3.2 Main verb

Variation in PPOP placement according to the particular co-occurring main verb is noted by Taylor (2008: 351, 364), although no verb was found to correlate always with left-of-P placement or always with right-of-P placement. Using CorpusSearch's 'make lexicon' function, I obtained a list of all forms of all main verbs that co-occur with a PPOP in the YCOE. This list was then used to identify

all morphological and spelling variants of the verbs, from which all verb lemmas co-occurring at least 100 times with a PPOP were identified. Forms of *beon*, *wesan* and *(ge)weorðan* are classified together as BE. Verb lemmas co-occurring fewer than 100 times with a PPOP are assigned to the 'miscellaneous' category. YCOE annotations do not distinguish between auxiliary and main verb uses of forms of 'to be' or 'to have', but their auxiliary use is easily determined by the presence of a non-finite verb form. PPs governed by a non-verbal element and those co-occurring with an elided main verb were not coded for verb form as these PPs are independently excluded from the variation analysis.

The univariate results for the independent variable VERB are given in Table 3.16, in which the verb lemmas are listed in decreasing order of frequency.

Table 3.16 Distribution of PPOPs by co-occurring main verb

	Left-of-P	Right-of-P	Total
miscellaneous	1,117 (26.6%)	3,084 (73.4%)	4,201
cweðan 'to say'	569 (42.9%)	758 (57.1%)	1,327
BE	161 (19.4%)	670 (80.6%)	831
cuman 'to come'	336 (43.6%)	435 (56.4%)	771
habban 'to have'	37 (9.9%)	336 (90.1%)	373
sprecan 'to speak'	74 (29.5%)	177 (70.5%)	251
sendan 'to send'	80 (34.8%)	150 (65.2%)	230
gan 'to go'	61 (31.6%)	132 (68.4%)	193
lædan 'to lead'	31 (18.2%)	139 (81.8%)	170
liefan 'to allow'	4 (2.4%)	162 (97.6%)	166
niman 'to take'	50 (31.1%)	111 (68.9%)	161
biddan 'to ask'	26 (20.2%)	103 (79.8%)	129
faran 'to go'	35 (27.8%)	91 (72.2%)	126
don 'to do'	21 (17.5%)	99 (82.5%)	120
bringan 'to bring'	47 (39.2%)	73 (60.8%)	120
clipian 'to speak, call'	53 (45.7%)	63 (54.3%)	116
feohtan 'to fight'	63 (63.0%)	37 (37.0%)	100
Embedded PP	3 (1.6%)	181 (98.4%)	184
Elided main verb	7 (5.2%)	127 (94.8%)	134
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

Although there is very little variation with *liefan*, the majority of PPOPs that co-occur with this verb are independently excluded from further statistical analysis on the basis that only a very small proportion is clearly or very probably dative.

3.5.3.3 Narrative mode

In Alcorn (2009: 442–3), I noted that 84% of first and second person PPOPs occur in contexts of direct speech compared to just 16% of third person PPOPs. As first and second person PPOPs also appear in a right-of-P position much more frequently than third person PPOPs (see section 3.5.1.1), I questioned the possibility of a link between these two observations. Although I subsequently found that third person PPOPs are more frequently specially placed than first and second person PPOPs whether they occur in direct speech contexts or not, I did find that left-of-P placement is significantly less frequent in direct speech contexts than elsewhere. This evidence is the basis for treating NARRATIVE MODE as a variable in the present study.

In deciding whether a PPOP occurs in direct speech, I follow the decisions of the YCOE editors. Complements of verbs of saying are always labelled as direct speech unless introduced by *þæt* 'that' as in *he said that he would go*. Personal comments of the narrator are also labelled as direct speech, although the corpus documentation indicates this occurs only in texts in which the narrator's personal comments can easily be distinguished from the narrative (Taylor 2003: *Direct speech*), *Bede* and *Orosius* being the only two such texts identified. The direct speech label does not distinguish between complements of verbs of saying and personal comments of the narrator, and in many cases there is no simple way to differentiate the two other than by examining each instance individually. I have examined all of the PPOPs that occur in a clause of direct speech in the YCOE's *Orosius* text file (coorosiu) and found that only a very small number belong to the narrator's personal comments. The corpus documentation also points out that texts which could be characterised as composed largely, if not wholly, of representations of direct speech, e.g. homilies, are not labelled as speech. Consequently, it may be concluded that, in the vast majority of cases, the direct speech label indicates the clause is a complement of a verb of saying. The only exception I have found to this generalisation is *Alexander's letter to Aristotle*, the

entire content of which is labelled as direct speech. This is surprising since the corpus documentation asserts that letters are not labelled as speech (Taylor *ibid*), which is true of other letters represented in the corpus (and also of texts such as personal wills and charters). Nevertheless, I have adopted the YCOE editors' classification for the 58 simple PPOPs in coalex.

For ease of reference, I refer to contexts identified as part of a direct speech sequence as 'mimetic' (from *mimesis* 'imitation of another persons' words), and for 'elsewhere' contexts I use the term 'diegetic' (from *diegesis* 'the narrative presented by a literary work'). PPOPs in mimetic contexts may therefore be described as those that belong primarily to a complement of a verb of saying, while PPOPs in diegetic contexts are those occur in some other context.

The univariate results for the independent variable NARRATIVE MODE are given in Table 3.17. The results confirm my earlier findings (Alcorn 2009: 442, Table 3) that left-of-P placement is significantly less frequent in mimetic contexts than in diegetic contexts.

Table 3.17 Distribution of PPOPs by narrative mode

	Left-of-P	Right-of-P	Total
Mimetic	429 (12.7%)	2,960 (87.3%)	3,389
Diegetic	2,346 (37.2%)	3,968 (62.8%)	6,314
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

3.5.3.4 Subject form

In the course of processing the data for this study I formed a clear impression that left-of-P placement occurs noticeably less often when there is another personal pronoun in the clause. I therefore introduce a new variable to test whether there is indeed such a correlation. Of those PPOPs that co-occur with another personal pronoun, 85% co-occur only with a subject personal pronoun, 7.5% co-occur only with an object personal pronoun, and 7.5% co-occur with both a subject and an object personal pronoun. Since the vast majority of co-occurring personal pronouns are subjects, the data were coded according to the form of the subject only. Of course the 15% that co-occur with an object personal pronoun (with or without a co-occurring subject personal pronoun) might behave radically different

from the 85% that co-occur only with a subject personal pronoun, but the possibility is not explored here.

Three categories of subject form are recognised: personal pronouns; nominals; and other. In 90% of the 'other' cases there is no overt subject, mainly due to elision, non-realisation of a relative pronoun, because the verb is an imperative or hortative form, or because the clause is non-finite. In the 10% of 'other' cases with an overt subject, the subject is either the indefinite pronoun *man*, a quantifier, an overt relative pronoun or a non-nominative subject in a non-finite clause.

The univariate results for the independent variable SUBJECT FORM, given in Table 3.18, confirm there is some basis for the intuition that motivated this variable: left-of-P placement occurs about half as frequently in the presence of a personal pronoun subject as it does elsewhere.

Table 3.18 Distribution of PPOPs by subject form

	Left-of-P	Right-of-P	Total
Personal pronoun	554 (17.8%)	2,564 (82.2%)	3,118
Nominal	1,081 (35.8%)	1,937 (64.2%)	3,018
Other	1,140 (32.0%)	2,427 (68.0%)	3,567
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

3.5.4   Extra-linguistic variables

3.5.4.1 Date

The linguistic evidence represented by the YCOE's text files derives from scholarly editions of manuscripts written in English over a period of around 300 years, with the majority of evidence belonging to late rather than early Old English. There is good reason to think that the frequency of left-of-P placement of PPOPs might not have been constant throughout these 300 years. At some point during the Middle English period, left-of-P placement ceased to be an option (van Kemenade 1987: 190) so we might expect to see evidence of a declining frequency from early to late Old English.

Almost all of the linguistic evidence represented in the YCOE ultimately derives from copies of texts rather than from original compositions. This introduces a significant problem for deciding what date a particular text file should be associated with. This problem is illustrated in its simplest form by the text file cogregdC. This text file represents the language of a copy of a translation from Latin of *Gregory's Dialogues*. The original translation was undertaken in the closing quarter of the ninth century by Bishop Wærferth of Worcester (Yerkes 1982: 9), while the copy represented by cogregdC was written in the second half of the eleventh century (Ker 1957). What we really want to know, of course, is what period of Old English the *language* of the copy represents. If the copy were a literatim copy, cogregdC could be included in the early Old English category. If, on the other hand, the copyist 'modernised' the language of his exemplar in accordance with his own grammar and norms, cogregdC could be included in the late Old English category. The problem is that Wærferth's translation has not survived in its original version so we cannot determine whether the copy is literatim or not. This problem can be further confounded where information about the period in which the text was originally composed or translated is lacking, as with cogenesiC, which represents a version of the second half of Genesis (Raith 1952). Lacking the necessary skills and time to resolve such problems, I must instead choose between two practical methods for dating the language represented by the YCOE's text files: dating by reference to manuscript date (where known), and dating by reference to the date the text was originally composed (again, where known). Neither method is ideal: what might be right for some text files might not be right for others, and in some cases it may not be clear which method is the right one, as with cogregdC.

Taylor (2008) dates her materials by reference to date of original composition, and found conflicting evidence of the diachrony of PPOP placement during the Old English period. Unexpectedly, PPOPs governed by *to* 'to' in her sample show a significant increase in probability of left-of-P placement between early Old English, treated as Taylor as pre-950, and late, i.e. post-950, Old English. This contrasts with her results for other PPOPs, which show a slight decrease in left-of-P probability over time. Koopman (1992: 74–5) also looks for evidence of a change in PPOP placement over time and neatly avoids dating

problems by comparing data from a small number of early manuscripts (whose texts necessarily have an early composition date) with data from late compositions (whose Old English manuscript versions are necessarily late). His results, summarised in Table 3.19, are also unexpected and conflicting: overall, there appears to be a 9% increase in left-of-P frequency from the early to late period but results for the early period are very mixed.

Table 3.19 Left-of-P frequency by date (Koopman 1992: 75, table 5)

Early texts	N	Left-of-P	Late texts	N	Left-of-P
Cura Pastoralis	291	56 (19%)	*Ælfric's Homilies I*	565	158 (28%)
Orosius	383	130 (34%)	*Ælfric's Homilies II*	527	180 (34%)
Bede	449	61 (14%)	*Wulfstan's Homilies*	101	33 (33%)
Total	1,123	247 (22%)	Total	1,193	371 (31%)

For reasons I will shortly explain, I have elected to measure diachronic trends by reference to manuscript date. The YCOE documentation provides, for each text file, the date of the associated manuscript where given in Ker (1957), and PPOPs occurring in these text files were coded on that basis. Dates for charters and wills, i.e. text files with 'codocu' in their title, were obtained from Sawyer (1968), and PPOPs in those text files were coded accordingly. A three-way dating system was used: early, for manuscripts pre-dating AD 925; late, for manuscripts post-dating AD 975; and unclassified. Unclassified data are those associated with manuscripts dated to within 25 years of AD 950 or whose date range straddles AD 950, and those associated with manuscripts for which a date is not supplied by Ker or Sawyer. The 50-year 'buffer' between my 'early' and 'late' categories is equivalent to one or two generations of language users and should help to crystallise any evidence of diachronic variation. Coorosiu (*Orosius*), whose associated manuscript is dated by Ker to the first half of the tenth century, is categorised as 'early'. Coverhom (which represents a large proportion of the *Vercelli Homilies*), whose associated manuscript is dated by Ker to the second half of the tenth century, is categorised 'late'. Several text files are based on composite editions, i.e. the base material for the text file comes from more than manuscript, although only three of these text files — coaelhom (*Ælfric's Supplementary Homilies*), cobede (*Bede*) and coboeth (*Boethius*) —

supply a reasonable number of PPOPs. All of the source manuscripts represented by coaelhom are dated to post-975 by Ker, so all PPOPs in this text file are dated 'late'. For cobede, PPOPs associated with the Tanner manuscript are coded 'early', all others are coded 'late', again in accordance with Ker. The base edition for coboeth represents two manuscripts: the main one dated to between AD 925 and AD 975 and the other falling into my 'late' category. As excerpts from the late manuscript occur frequently but irregularly, I have not attempted to date PPOPs in this text file: these pronouns are therefore included in the unclassified category. Data in cochronA (the *Parker Chronicle*) is associated with individual hands by the YCOE editors by means of text file label extensions, e.g. cochronaA-1 indicates scribe 1. PPOPs associated with each hand are dated in accordance with the dates assigned to the hands by Bately (1986: xxi-xliii). Material in each of the other three versions of the *Anglo-Saxon Chronicle* represented in the YCOE, i.e. cochronC, cochronD and cochronE, are associated with manuscripts that Ker dates to after AD 975, and so all PPOPs in these text files are coded 'late'.

The univariate results for the independent variable DATE are given in Table 3.20. These indicate a slight decrease in frequency of left-of-P placement from early to late Old English according to manuscript date.

Table 3.20 Distribution of PPOPs by manuscript date

	Left-of-P	Right-of-P	Total
Early	355 (31.0%)	792 (69.0%)	1,147
Late	2,327 (28.1%)	5,952 (71.9%)	8,279
Unclassified	93 (33.6%)	184 (66.4%)	277
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

For the majority of PPOPs for which both a manuscript date and a text composition date have been identified, the choice between the two dating methods is immaterial: 73% of PPOPs coded 'early' or 'late' according to manuscript date would have the same date value if composition date was used instead. There are, however, three factors favouring manuscript date as the means of measuring diachronic change in my sample. Firstly, about four times as many PPOPs end up

in the 'unclassified' category if composition date rather than manuscript date is used. Secondly, all of the Anglian-influenced data are associated with texts with an early composition date, which would make it impossible to investigate diachronic change for those data by means of that dating system, at least on an early ~ late scale. Thirdly, if diachronic change were to be measured by reference to composition rather than manuscript date, it would be impossible to differentiate between diachronic change and dialect effects. This can be seen by comparing the distribution of the sample in Table 3.21, which cross-tabulates the two major dialect categories introduced in the following section data by manuscript date, with the distribution in Table 3.22, which cross-tabulates dialect by date of composition. Data unclassified for data and dialect are excluded from each table for expository purposes. As soon as we move beyond univariate analyses, the empty cell in the second table creates a problem. Imagine, for example, a study of the effects of age (child *vs*. adult) and gender on, say, *th*-fronting. If the sample were to include data from girls, women and men but not from boys, we might be able to discern something about the effects of age and something about the effects of gender, but we could not be certain that our findings for one variable were entirely independent of our findings for the other.

Table 3.21 Date (manuscript) x Dialect

	Anglian-influenced	West Saxon	Total
Early	378	762	1,140
Late	1,657	4,180	5,837
Total	2,035	4,942	6,977

Table 3.22 Date (composition) x Dialect

	Anglian-influenced	West Saxon	Total
Early	1,578	1,101	2,679
Late	—	3,643	3,643
Total	1,578	4,744	6,322

Lacking access to further data, we could limit the study to the effects of age, comparing data on girls with data on boys, or else limit the study to the effects of gender, comparing data on women with data on men. The point is that we could not investigate the effects of *both* independent variables from such an unbalanced sample. The same type of problem would be encountered in the present multivariate study if date were measured by reference to composition date: we could discern something about the effects of date, but only at the expense of learning something about dialect effects, OR we could discern something about the effects of dialect, but only at the expense of learning something about change over time, but we could not learn about the effects of both variables from a sample which distributes as shown in Table 3.22. This problem is avoided entirely by measuring date by reference to manuscript date.

In short, manuscript date is far from an entirely reliable method for measuring diachronic change, but it is no less reliable, at least in principle, than text composition date. It also has the advantage of permitting a more inclusive analysis of diachronic change and allowing dialectal differences to be independently measured. For these reasons, I have elected to measure date by reference to manuscript date.

3.5.4.2 Dialect

The 'dialects' of Old English denote collections of linguistic features that distinguish one group of text languages from others, but the possibility of a correlation between frequency of left-of-P placement and dialect has yet to be explored. Wende (1915: 77–81) notes some differences in case forms between *Bede*, an Anglian-influenced text, and the West Saxon texts in his sample, but does not attempt to measure differences in left-of-P placement along dialectal lines. Taylor's (2008) sample includes both West Saxon and Anglian-influenced materials but as she measures date by reference to date of text composition, it is not surprising to find dialect absent from her inventory of variables.

Known syntactic (as opposed to morphosyntactic) differences between the dialects of Old English are few and far between, but this is largely a reflection of the nature of the extant material. Fischer *et al*. (2000: 37) suggest '[t]here is little scope for work on dialect syntax in Old English; almost all the texts are in the

West Saxon dialect, while those works of any length that were not written in West Saxon consist mostly of interlinear glosses on parts of the Vulgate bible'. This paints an overly pessimistic picture. Evidence of dialects other than West Saxon may be 'relatively meagre' (Toon 1992: 451), but has been found in a number of predominantly West Saxon materials with a diverse history of transmission. Since information about the dialects of many of the base manuscripts represented in the YCOE is provided in the corpus documentation (Taylor 2003, *Text information*), it would be a pity not to exploit it here.

The dialectal information in the YCOE documentation comes directly from the Helsinki Corpus of English Texts (HCET) documentation. The HCET editors recognise five Old English dialects: Anglian, Anglian Mercian, Anglian Northumbrian, Kentish and West Saxon (Kytö & Rissanen 1992: 17–18, Kahlas-Tarkka, Kilpiö & Österman 1993: 27–8). Some texts are classified for more than one dialect, indicating that their language shows some evidence of dialectal variety, and so some mixed dialect categories are also recognised, e.g. West Saxon+Anglian, West Saxon+Kentish. Neither the degree nor type(s) of dialectal variation evident in the language of a mixed dialect text is indicated by these labels and it is entirely possible that all PPOPs occurring in a text file classified as, say, West Saxon+Anglian happen to occur in stretches of the text where the language is distinctly West Saxon rather than Anglian. Still, it may be hoped that this system of categorisation is 'good enough' to allow linguistic differences between well established dialects to be detected, particularly where these differences are marked. Some texts are wholly unclassified for dialect, others only partly so, e.g., *Solomon and Saturn I* is classified 'West Saxon+unclassified'.

I use the same categories and, with a few exceptions, the same categorisations as the HCET editors. For some of the larger unclassified texts, I assume a particular dialect using information gleaned from the literature review in van Bergen's (2008) study of dialectal differences in negative contraction in Old English. I have classified the following text files as West Saxon+Anglian (see van Bergen 2008: 409, 415–17): coalcuin (Alcuin's *De virtutibus et vitiis*); comart1 and comart2 (the Old English *Martyrology*, mss. D and C respectively); conicodD (*Homily on the Harrowing of Hell*, ms. D), coverhom, coverhomE and

coverhomL (the *Vercelli Homilies*). I have classified comargaC (*Life of Saint Margaret*, ms. C) as West Saxon+Kentish (see van Bergen 2008: 414). On the basis of Napier (1894: lvii–lviii), I have classified corood (*History of the Holy Rood Tree*) as West Saxon. Classifications of all other texts follow those listed for each of the YCOE's text files in Taylor *et al.* (2003: *Text information*).

The univariate results for the independent variable DIALECT are given in Table 3.23.

Table 3.23 Distribution of PPOPs by dialect

	Left-of-P	Right-of-P	Total
West Saxon	1,801 (28.1%)	4,614 (71.9%)	6,415
West Saxon+unclassified	107 (36.4%)	187 (63.6%)	294
West Saxon+Anglian	335 (24.0%)	1,058 (76.0%)	1,393
West Saxon+Anglian Mercian	175 (26.4%)	488 (73.6%)	663
Anglian Mercian	6 (43.0%)	8 (57.0%)	14
Kentish	—	2 (100%)	2
West Saxon+Kentish	26 (37.0%)	44 (63.0%)	70
Unclassified	325 (38.1%)	527 (61.9%)	852
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

As the Kentish PPOPs show no variation in position, they are obviously unsuitable for the analysis of variation in pronoun placement. As there are only two of them, they were subsequently added to the West Saxon+Kentish category. Results for data in the 'unclassified' and 'West Saxon+unclassified' categories cannot be interpreted from a dialectal perspective, so they were subsequently conflated into a single 'other' category.

The difference between the three linguistically meaningful groups of data of which a reasonable number of examples is available, i.e. West Saxon, West Saxon+Anglian and West Saxon+Anglian Mercian is slight. This does not hold much promise for detecting a dialectal difference in PPOP placement, but multivariate analyses often show that raw frequencies can be misleading.

3.5.4.3 Latin interference

A sizeable proportion of data in the YCOE comes from Latin translations and Taylor (2008) shows that there is a much reduced tendency to place PPOPs to the preposition's left in texts translated from Latin to Old English in comparison to non-translated Old English texts. Unlike Old English PPs, Latin PPs are uniformly head-initial unless headed by *cum* 'with'.[40] Taylor's study shows that when translators produced an Old English PP in direct response to a PP in their Latin source text, they were much more likely to avoid left-of-P placement. This she interprets as evidence of a direct interference effect. In biblical translations in particular, Taylor also found that left-of-P placement tends to be dispreferred even when the Old English PP does not correspond to a PP in the Latin. This she interprets as evidence of an indirect interference effect.

In order to control for both types of interference effect, the data were coded for three variables. One variable identifies whether each PPOP comes from a Latin translation (N=5,146), a non-translation (N=3,396) or an unclassified text (N=1,161). 'Unclassified' means that neither the YCOE editors nor the editors of the Helsinki Corpus of English Texts know whether or not the text is a translation. Data were coded for this variable in accordance with information taken directly from the YCOE's documentation, with one exception. CogenesiC (the version of *Genesis* as found in Cambridge, Corpus Christi College, 201), which is unclassified for genre in the YCOE's documentation, is classified here as a biblical translation on the basis of information in Raith (1952). A second variable is introduced to distinguish between direct and indirect interference effects. This variable identifies whether each PPOP in a Latin translation comes from biblical translation (N=2,080) or a non-biblical translation (N=3,066). The third variable identifies whether PPOPs in translations belong to a PP that corresponds directly to a PP in the Latin source text. Of the 5,146 PPOPs occurring in translations, approximately 1,950 have already been coded for this third variable by Taylor for the purposes of her (2008) study and I am fortunate to have been granted full access to this information.

---

[40] Ablative personal pronoun objects of *cum* tend to be positioned immediately before the preposition, e.g. *mecum* 'with me', *tecum* 'with you' (Sinkovich 1984: 75).

The examples at (26), taken from Taylor (2008: 345, ex. 4), illustrate an Old English PP that corresponds directly to a PP in the Latin. I will refer to this type of Old English PP as 'matched'. The examples at example (27), taken from Taylor (2008: 345, ex. 7), illustrate an Old English PP that does not correspond to a PP in the Latin. I will refer to this type of Old English PP as 'unmatched'. Taylor further distinguished matched PPs according to whether the Latin PP is head-final, i.e. a *cum*-PP, or head-initial.

(26) a. Efne nu   eall seo eorðe lið **ætforan ðe**
        even now all   the earth lies before   you
        'Even now all the earth lies before you'

                                                      (cootest, Gen:13.9.506)

    b. Ecce   universa terra **coram te**   est
        behold all   earth before you is
        'Behold all the earth is before you'

                                                      (Genesis 13.9)

(27) a. Hwæt gesawe ðu   **mid us**, þæt ðu   swa don woldest
        what   saw   you with us that you so   do   would
        'What did you see among us, that you would do so?'

                                                      (cootest, Gen:20.10.846)

    b. Quid vidisti,   ut   hoc faceres
        what you-have-seen that this you-would-do
        'What have you seen, that you would do this?'

                                                      (Genesis 20.10)

To increase the size of the sample already coded for the third variable, I analysed a further 915 PPOP-governing PPs using exactly the same methods specified by Taylor (2008: 360–2). For biblical translations, I coded all examples in the *Heptateuch*'s book of *Joshua* (cootest,Josh), using Crawford (1922) for the Latin, and all those in *West Saxon Gospel*'s book of *Mark* (cowsgosp,Mk), plus the first 50% of PPOPs in *Luke* (cowsgosp,Lk) and *John* (cowsgosp,Jn), using Colunga & Turrado (1985) for the Latin. This increases the proportion of PPOP-

governing PPs in biblical translations coded for the third variable to 66%. For non-biblical translations, I coded all PPOPs in *Life of Saint Margaret* ms. C, using Clayton & Magennis (1994) for the Latin, and all PPOPs in books three and four of *Bede*, using Colgrave & Mynors (1969) for the Latin. This increases the proportion of PPOP-governing PPs in non-biblical translations coded for the third variable to 48%. The only non-biblical translation that contains more than 100 PPOPs but which is completely unsampled is *Boethius* (N=approximately 200).

Using the information provided by these three variables, the data were transformed into a fourth variable, LATIN INTERFERENCE, consisting of the categories relevant for recognising direct and indirect interference effects. The univariate results for this variable are given in Table 3.24. Unsampled data are those that have not been compared to the Latin.

Table 3.24 Distribution of PPOPs by Latin Interference factors

	Left-of-P	Right-of-P	Total
Non-translations	1,206 (35.5%)	2,190 (64.5%)	3,396
Non-biblical translations:			
matched, *cum*-PP	3 (5.0%)	60 (95.0%)	63
matched, other PP	49 (10.9%)	400 (89.1%)	449
unmatched	402 (41.4%)	568 (58.6%)	970
unsampled	476 (30.1%)	1,108 (69.9%)	1,584
Total	930 (30.3%)	2,136 (69.7%)	3,066
Biblical translations:			
matched, *cum*-PP	2 (2.0%)	82 (98.0%)	84
matched, other PP	63 (8.8%)	652 (91.2%)	715
unmatched	119 (20.4%)	463 (79.6%)	582
unsampled	57 (8.2%)	642 (91.8%)	699
Total	241 (11.6%)	1,839 (88.4%)	2,080
Unclassified	398 (34.3%)	763 (65.7%)	1,161
Total	2,775 (28.6%)	6,928 (71.4%)	9,703

These results are in line with Taylor's description of the two types of interference effect. Left-of-P placement occurs less frequently in translations than in non-translations when the PP is matched with a Latin PP (the direct interference effect), and left-of-P placement occurs less frequently in biblical translations than in non-translations when the PP is not matched with a Latin PP (the indirect interference effect).

Given Taylor's clear evidence of a tendency on the part of translators to conform to the word order of Latin PPs when translating them directly into English PPs, it is particularly striking that only 5 of the 147 PPOPs belonging to a translation of a head-final Latin *cum*-PP are in a left-of-P position. Taylor (2008: 361–2) notes this too, remarking that 'the OE translators are clearly not influenced by this inversion'. However, the fact that left-of-P placement is even *less* frequent in translations when the PP is a translation of a head-final *cum*-PP than when it translates some other PP is perhaps to be expected: 128 (87%) of the 147 Old English PPs matched with a *cum*-PP are headed by *mid* 'with'. As is evident from Table 3.10 of section 3.5.2, *mid* is one of the prepositions least likely to occur with a left-of-P PPOP, at least in absolute terms. Since the 147 PPOPs matched with a *cum*-PP exhibit minimal variation, they are not ideally suited for variation analysis. Rather than exclude them, however, I follow Taylor (2008) and reduce the four 'matching' categories to two: one for biblical translations and one for non-biblical translations.

3.5.5   Variables not encoded

3.5.5.1 Animacy of referent

Cardinaletti & Starke (1996, 1999) show that in many languages the syntax of simple personal pronouns differs from that of their nominal counterparts when the pronoun refers to a non-human entity. For example, personal pronouns with a human referent can be freely coordinated and contrastively stressed, as in (28), those with a non-human referent cannot, e.g. (29).

(28) a. {Janet / She} and John had a fight

    b. Who stole the cake? It was {JOHN / HIM}

(29) a. {The cat / *It} and the dog had a fight

   b. Who stole the cake? It was {the DOG / *IT}

Such observations led me to investigate the possibility of an association between left-of-P placement of Old English PPOPs and pronoun referent (Alcorn 2009: 438–42) with a view to exploring the contrast between third person PPOPs and first/second person PPOPs in terms of frequency of left-of-P placement. What I found instead is that Old English PPOPs very rarely refer to non-human entities regardless of their position. *Hit*, the object pronoun most likely to denote a non-human entity (Heltveit 1958: 361–2, 366, Mitchell 1985: §§55–71), occurs only seven times as a PPOP in the entire YCOE, and only four times with a non-human referent, and of all 500+ PPOPs occurring in *Lives of Saints* (coaelive), only fourteen have a non-human referent, some of which are clearly personified. In this respect, Old English appears to be broadly consistent with most modern West-Germanic varieties, whose pronominal objects of prepositions take the form of a locative or demonstrative pronoun, rather than a personal pronoun, when the pronoun's referent is non-human (e.g. van Riemsdijk 1982: 36–45, Toebosch 2003: 45–7, Zwart 2005: 920). Although my (2009) sample was not exhaustive, I remain firmly of the opinion — based on my translations of hundreds of individual examples for many different purposes — that PPOPs with non-human reference are fairly rare.

3.5.5.2 Preposition modification

Wende (1915) identifies two variables that correlate with PPOP positioning for which I have not encoded the data. One is discussed in the following section. The other is modification of the preposition by an adverb, which Wende claims co-occurs regularly with right-of-P placement — even when the adverb and preposition are not adjacent (*ibid*: 65–6). The types of example Wende includes in this category are parsed in three different ways in the YCOE, as illustrated in the following examples. Each of these examples is taken from the un-numbered list in Wende (1915: 65). Some of the prepositions Wende classifies as modified are parsed as a cliticised adverbial particle, as in (30);

(30)   þa   wæs sume dæge se Godes wer <u>in</u>gongende **to him**

      then was some day   the God's man in-coming   to him'

      'then one day the man of God came in to him'

<div align="right">(cobede,Bede_2:9.132.15.1271)</div>

some are parsed as an element of what the YCOE editors call a 'multi-word preposition', as in (31);

(31) a. ða   hwearf se   ana   eft   <u>in</u> **to him**, se   ðe þone heofenlican

      then came   the-one alone afterwards in to him who that the   heavenly

      song gehyrde

      song heard

      'then he who had heard the heavenly song afterwards came in alone to him'

<div align="right">(cobede,Bede_4:3.266.17.2711)</div>

   b. swa astrehte   he hine sylfne to eorðan **wið**   his <u>weard</u>

      so   prostrated he him self   to ground against him towards

      'so he prostrated himself on the ground towards him'

<div align="right">(cocathom2,ÆCHom_II,_11:99.247.2075)</div>

and some are parsed as a free-standing adverb, as in (32).

(32)   ... þætte swa æðele wer ... swa <u>feor</u> **fram him**   gewite

       that   so   noble man ... so   far   from them goes

      '… that so noble a man ... should go so far from them'

<div align="right">(cobede,Bede_2:1.98.5.912)</div>

I do not agree with Wende that examples like those in (30) and (31) involve a modified preposition. In examples like (30), the status of the underlined word is generally, if not also more naturally, interpretable as a verbal particle/prefix, and in examples like (31), I agree with the YCOE editors that we are dealing with a single, complex preposition. Some complex prepositions are sometimes written as one word, e.g. *into* 'into', *oninnan* 'within, into', *onuppan* 'upon', *toforan*

<div align="right">116</div>

'before', *ymbutan* 'around, outside', and I treat them as equivalent to simplex prepositions whether written as one word or two since, as with simplex prepositions, PPOPs are found on either side of the complex, whereas nominal objects do not occur to the left. As example (31b) shows, the object can also occur between the elements of some complex prepositions. Fifteen PPOPs occur between the elements of a complex preposition in the YCOE, but so too do 96 nominal objects (Alcorn 2009: 436, fn. 7) so clearly the medial position is not a special clitic position.

There certainly appears to be a modifying relationship between adverb and PP in examples like (32), although I am not entirely certain that the adverb is necessarily the modifier. I have looked at all cases involving *feor*, as well as *neah* and *gehende*, both 'near', and it is true that right-of-P placement is canonical. However as there are only 25 examples in total, no firm conclusions can be drawn. The YCOE parse of this particular example does not identify either constituent as dependent on the other, consequently there is no way to identify other examples except by identifying every clause in which an adverb or adverbial phrase co-occurs with a PPOP-governing PP and then sorting though the results to identify those potentially involving dependency. This is far too onerous a task to be tackled here.

3.5.5.3 PP semantics

The second variable identified by Wende (1915) as correlating with PPOP placement for which my data are not coded is PP semantics. According to Wende (*ibid*: 73–75), left-of-P placement occurs much more frequently when the preposition has a spatial meaning, be it literal or metaphoric. The problem is that most Old English prepositions can express some kind of spatial relation, and few — if any — can be excluded from a spatial categorisation purely on the basis of their form. In addition, prepositions can express different types of spatial relations, i.e. GOAL, e.g. *to* 'to', SOURCE, e.g. *fram* 'from', PATH, e.g. *þurh* 'through', and LOCATION, e.g. *æt* 'at', and many can express more than one type, e.g. *æt*, which can additionally express GOAL, as in (33), and SOURCE, as in (34).

(33)   &   ge   ne comon **æt me**

and you not came   at me

'and you did not come to me'

(cowsgosp,Mt_[WSCp]:25.43.1818)

(34)   Anymaþ þæt pund   **æt hym**

take   that pound at him

'Take that pound from him'

(cowsgosp,Mt_[WSCp]:25.28.1779)

In short, it is impossible to accurately classify the data for PP semantics without examining each example individually. The relationship between PP meaning and pronoun case, however, is considered in the following chapter.

3.6 Summary

The univariate analyses have revealed several contexts in which right-of-P placement occurs at least 95% of the time (allowing for rounding):

- where the PPOP is, or can be assumed to be, accusative or genitive (860/899 [95.7%] right-of-P, see Table 3.7);
- where the preposition is *be* (277, all right-of-P), *for* (290, all right-of-P) or *þurh* 'through' (N=182, all right-of-P) (see Table 3.10);
- where the PP is coordinated with another PP in the same clause (118/120 [98.3%] right-of-P, see Table 3.11);
- where the PP is embedded under a non-verbal constituent (181/184 [98.4%] right-of-P, see Table 3.12);
- where the PP occurs in a clause with an elided main verb (127/134 [94.8%] right-of-P, see Table 3.13).

Collectively, these factors account for 1,760 of the 9,703 PPOPs in the sample, which means that 18.1% of unmodified and uncoordinated PPOPs in the YCOE can be predicted, with a very high degree of confidence, to follow their governor

by reference to these factors.[41] The univariate analyses also revealed one context in which left-of-P placement occurs at least 95% of the time (allowing for rounding): that is where the preposition is *betweonum* (229/242 [94.6%] left-of-P, Table 3.10). The following chapter considers why PPOPs behave the way they do in the presence of each of these factors.

In order to better expose the underlying relationship between placement of the 80% or so of PPOPs that are not dealt with in Chapter 4 on the one hand and each of the other factors introduced in this chapter on the other, I use the multiple logistic regression function of Goldvarb X (Sankoff, Tagliamonte & Smith 2005). This procedure and the results it derives are introduced and discussed in Chapters 5 and 6.

---

[41] The number of pronouns associated with these factors is less than the sum of the numbers associated with each factor individually. This is because some pronouns are associated with more than one factor, e.g. pronouns belonging to a coordinated PP headed by *be* 'by'.

Chapter 4 (Near) categorical variables

4.1 Introduction

This chapter considers why PPOPs are (almost) never or, in one case, almost always realised as special clitics in the presence of certain factors. Section 4.2 considers right-of-P placement of pronouns belonging to a coordinated PP and section 4.3 considers right-of-P placement of those belonging to an embedded PP or to a clause with an elided main verb. The effect of pronoun case on PPOP placement is then considered in section 4.4 and the chapter concludes by looking closely at the near-invariable placement of pronouns governed by certain prepositions.

4.2 PP coordination

In Chapter 3, PP coordination was found to correlate almost always with right-of-P placement. Recall that Wende (1915: 66–8) identifies a number of different types of coordinate structures but that only those involving a PP coordinated with a PP in the same clause are classified as coordinated in the present study. Wende (*ibid*) claims that PPOPs belonging to a PP that corresponds to a PP in a parallel VP also occur regularly to the preposition's right but, as explained earlier, PP coordination is not easily recognised in the YCOE unless the PP conjuncts occur in the same clause. A total of 120 PPOPs were identified as belonging to a PP coordinated with another PP in the same clause, and just two of these pronouns are specially placed. However, in 70 of these 120 examples (58%, all right-of-P), the PPOP is clearly or potentially accusative and/or governed by one of the 'knockout' prepositions, i.e. *þurh*, *be* or *for* and right-of-P placement of these 70 pronouns is considered later in this chapter. With only 50 examples left, the correlation between PP coordination and right-of-P placement is not quite so remarkable. Nevertheless, there does appear to be a plausible explanation for right-of-P placement for the majority of the remaining cases.

In 38 of these 50 examples (76%), the PP is coordinated with a PP headed by the same preposition, as in (1).

(1) a.   ic eom swiðe bliðemod **to eow** &   **to eallum Godes monnum**

      I   am   very   friendly   to you and to all   God's men

      'I am very friendly to you and to all God's men'

<div align="right">(cobede,Bede_4:25.348.9.3503)</div>

   b.   ac **fram me** ge   beoð ascyrede, and **fram ælcere myrhþe**

      but from me you are   separated and from every   pleasure

      'but you will be separated from me and from every pleasure'

<div align="right">(cosevensl,LS_34_[SevenSleepers]:160.122)</div>

   c.   Ne spræc ic na **to ðe**, ac   **to minum þeowetlinge**

      not spoke I   not to you but to my   servant

      'I spoke not to you, but to my servant'

<div align="right">(cogregdC,GDPref_and_3_[C]:20.221.27.3014)</div>

The conjunction of two PPs headed by the same preposition and occurring in the same clause naturally establishes a relationship of equality between their objects' referents, as in (1a) and (1b), just as their disjunction naturally establishes a relationship of contrast, e.g. (1c). In Present Day English at least, this relationship of identity/contrast demands some degree of prosodic focus on the objects of the coordinated prepositions. For none of the translations given for these examples would it be natural to articulate the emboldened pronoun without stress and I assume this would be the case in spoken Old English also. On the assumption that the 38 PPOPs illustrated by the examples at (1) are prosodically strong pronouns, i.e. phonologically independent words, their status as non-clitic elements is predicted (Zwicky 1977, Cardinaletti 1994, Cardinaletti & Starke 1999: 172). Consequently so too is their placement to the preposition's right.

     The example at (2) stands as the only exception to right-of-P placement when the PP is coordinated with a clause-mate PP headed by the same preposition.

(2)   … for ðam wundum þe   **him on** wæron &   **on eallum Egypta lande**

     for the   sores   that them on were   and on all   Egypt's land

   '… because of the sores that were upon them and upon all the land of Egypt'

<div align="right">(cootest,Exod:9.11.2700)</div>

As (2) is a close translation from Latin, (3), left-of-P placement of *him* is especially surprising.[42] Taylor (2008) shows that where the Old English PP is a direct translation of a head-initial Latin PP, left-of-P placement tended to be avoided, and her findings are confirmed in Chapter 6.

(3)   ... propter   ulcera quæ **in illis**   erant, et   **in omni terra Ægypti**

        because-of sores that in those were and in all   land of-Egypt

    '... because of the sores that were upon them and upon all the land of Egypt'

(Crawford 1922: 235)

Eleven of the remaining twelve examples likewise involve a PP that is coordinated with another in the same clause, except that the PPs are headed by different prepositions. In six cases, their objects co-refer, e.g. (4), and in five cases their objects have different referents, e.g. (5). In all eleven examples the PPOP is right-of-P.

(4)   ... forðan þe se sunu is þæs   fæder wisdom, **of   him** and **mid him**

        because   the son   is of-the father wisdom from him and with him

    '... because the son is the father's wisdom, from him and with him'

(coaelive,ÆLS_[Christmas]:35.26)

(5)   Þa   þæt folc   of   þære ceastre þæt geherde þæt Ioseph wæs gecumen,

        when the people from the   town   that heard   that Joseph was come

        þa   comen heo ealle him togeanes and cwædon, La, fæder Ioseph, sibb

        then came   they all   him towards   and said   Lo father Joseph peace

        sy **mid þe**   and **on þine ingange**

        is with you and on your entry

    'When the people of the town heard that, (i.e.) that Joseph had come, then they all came to meet him and said: "Lo, Father Joseph! Peace be with you and on your entry"'

(conicodC,Nic_[C]: 130.136–7)

---

[42] The other Old English witnesses to (2) — from the *Heptateuch*'s version of *Exodus* — use the same lexis and syntax for the coordinated PPs (Crawford 1922: 235).

As (4) and (5) represent just eleven examples altogether, the lack of any left-of-P examples is very possibly accidental. In examples like (4) in particular, left-of-P placement is not obviously precluded by pronoun prosody: rather it is the prepositions, not their objects, that are likely to be (contrastively) stressed.

The last example, (6), was given at (17) of Chapter 3, where it was identified as the only example in my sample in which the PP is coordinated with an adverbial phrase. In this case the pronoun is left-of-P.

(6)　… ealle ða ricu　　þe　**him under** bioð oððe **awer**　　　**on neaweste**

　　　　all　the mighty that him　under　are　or　　somewhere in proximity

　　　'… all the mighty who are under him or somewhere in proximity'

(coboeth,Bo:16.34.20.628

Although in both examples with a left-of-P PPOP, i.e. (2) and (6), the coordinated phrases are non-adjacent, this does not appear to be significant for PPOP placement: in seven of the 38 examples illustrated in (1), the coordinated PPs are non-adjacent, e.g. (1b). Rather, the data suggest that right-of-P placement is generally the rule when the pronoun belongs to a PP coordinated with a clause-mate PP headed by the same preposition regardless of whether the PPs are adjacent or not. The example at (2) is an exception to this rule, but the one at (6) is not.

The explanation proposed for the examples illustrated by those at (1) could be extended to examples like those at (12) and (13) of Chapter 3, repeated here at (7). These examples, from Wende (1915: 67), likewise involve two PPs headed by the same preposition but here they belong to parallel VPs. Wende (*ibid*: 66–8) identifies this as a further context in which right-of-P placement is the rule, whether the verb is repeated, as in (7a), or not, as in (7b).

(7) a.　Þæt þæt ic **to eow** gecweðe. þæt ic cweðe **to eallum mannum**

　　　　that that I　to you　say　　that I say　to all　　men

　　　　'That which I say to you, that I say to all men'

(cocathom2,ÆCHom_II,_40:301.57.6852)

b. he wunað **on me** and ic **on him**

he dwells in me and I   in him

'he dwells in me and I (dwell) in him'

<div align="right">(cocathom2,ÆCHom_II,_15:152.71.3365)</div>

As these parallel PPs belong to different clauses, relevant examples are difficult to identify in the YCOE. I am therefore unable to say whether PPOPs are invariably placed to the preposition's right when the PP corresponds with another in a parallel VP, nor whether right-of-P placement is the rule when PPs belonging to parallel VPs are headed by different prepositions, as other examples supplied by Wende (*ibid*) suggest. I would not be surprised if the data were to show that right-of-P placement is indeed the rule where the parallel VPs involve the same preposition, as in (7). In such cases, the object of each preposition is likely to be stressed, in the same way that objects belonging to coordinated clause-mate PPs are likely to be stressed when the same preposition is involved, as in (1). Where the prepositions differ, as in (8), however, the prepositions' objects may be stressed, but there is no reason to suppose they are necessarily stressed. Consequently, I would not be surprised if the data were also to show that left-of-P placement is possible where the parallel VPs involve different prepositions, although I have yet to stumble upon an example to support this assumption.

(8)   and he wunað **betwux us**. and we **mid him**

and he dwells between us and we with him

'and he will dwell among us, and we with him'

<div align="right">(cocathom2,ÆCHom_II,_45:339.121.7604)</div>

4.3   PP embedding and main verb elision

A total of 184 PPOPs were identified as belonging to a PP that is parsed as a constituent of some non-verbal element in the YCOE (see Table 3.12). Just three of these PPOPs (1.6%) occur in a left-of-P position. A total of 134 PPOPs were identified as belonging to a PP that is parsed as a constituent of a verbless clause (see Table 3.13). Just seven of these PPOPs (5.2%) occur in a left-of-P position. For reasons I will come to, both contexts are dealt with together in this section.

In 34 (18.5%) of the 184 examples involving an embedded PP and in 23 (17.2%) of the 134 examples in a clause with an elided main verb, the PPOP is not a clear/likely dative pronoun or else is governed by one of the 'knockout' prepositions, i.e. *þurh*, *be* or *for*. The invariable right-of-P placement of these 57 pronouns is considered later in this chapter. Another 28 of these pronouns are very probably pragmatically focused due to their occurrence in a parallel structure, e.g. (9), or for reasons of contrast, as in (10). In the previous section, I concluded that such pronouns are very probably stressed and that their right-of-P placement can be predicted from their phonological independence.

(9) a. ... þæt ic come **to him** and na hi **to me**

       that I come to them and not they to me

   '... that I come to them, and not they to me'

<div align="right">(coaelive,ÆLS_[Basil]:423.748)</div>

  b. ... hwæþer heora sceolde on oþrum sige habban, þe he on

      which of-them must on other victory have, whether he on

  Romanum, þe Romane **on him**

  Romans, or Romans on him

  '... which of them would have victory against the other, whether he against the Romans, or the Romans against him'

<div align="right">(coorosiu,Or_4:1.84.8.1695)</div>

(10) a. Mycel and mære is se God Cristenra manna, and an soð God

    great and splendid is the God of-Christian men and one true God

  Hælende Crist, and nis nan oþer **buton him**

  Saviour Christ and not-is no other except him

  'Great and splendid is the God of Christian men and the one true God Saviour Christ, and there is no other except him'

<div align="right">(coeust,LS_8_[Eust]:462.483–4)</div>

  b. He cwæð: blissiað mid me. for þan ðe ic gemette min forlorene sceap.

    he said rejoice with me because I found my lost sheep

  Ne cwæð he blissiað mid þam sceape: ac **mid me**

  not said he rejoice with the sheep but with me

'He said, "Rejoice with me, because I have found my lost sheep." He did not say "Rejoice with the sheep", but "with me"'

<div style="text-align: right">(cocathom1,ÆCHom_I,_24:372.41.4694–5)</div>

Of the remaining 233 PPOPs (144 in an embedded PP and 89 in a clause with an elided main verb), 157 (67%) are of the type illustrated in (11) and (12).

(11) a. ... þæt Nicanor feol and eall his folc   **mid him**
           that Nicanor fell and all his people with him
        '... that Nicanor fell, and all his people with him'

<div style="text-align: right">(coaelive,ÆLS_[Maccabees]:647.5259)</div>

   b. Her Oswiu ofsloh Pendan on Winwidfelda. &   xxx   cynebearna
       here Oswiu killed Penda   in Winwidfeld   and thirty of-royal-offspring
       mid him
       with him
       'In this year Oswiu killed Penda in Winwidfeld, and thirty royal offspring with him'

<div style="text-align: right">(cochronE,ChronE_[Plummer]:654.1.368)</div>

(12) a. þonne mannes sunu cymð on hys mægenþrymme &   ealle englas **mid**
         when man's   son comes in his majesty   and all   angels with
         him, þonne ...
         him then
         'when the son of man comes in his majesty, and all the angels with him, then ...'

<div style="text-align: right">(cowsgosp,Mt_[WSCp]:25.31.1785)</div>

   b. &   þy ilcan geare man ofsloh Ecgferð cining be norðan sæ. &
   and the same year   one killed Ecgferth king   by north   sea and
   mycelne here **mid him**
       great   army with him
       'and in the same year King Ecgferth was killed by the north sea, and a great army with him'

<div style="text-align: right">(cochronE,ChronE_[Plummer]:685.4.593)</div>

Each of these 157 examples exhibits the following features: (i) the PP is headed by *mid* 'with'; (ii) the PP is part of a coordinated structure; and (iii) the PPOP co-refers with an NP in the same clause or sentence. The difference between the examples represented by (11) and those represented by (12) lies in the way they are parsed in the YCOE. The syntactic tree notation used in the YCOE is much simpler than that found in most theoretical models and, lacking a VP node, representations are multiply-branching and consequently quite flat. A simplified version of part of the YCOE parse of (11b) is given at (13). I omit features such as annotations of punctuation and I simplify word- and phrase-level labels that are not crucial to point at hand. In this example, the *mid*-PP is parsed as a constituent of a post-positioned conjunct NP.

(13)
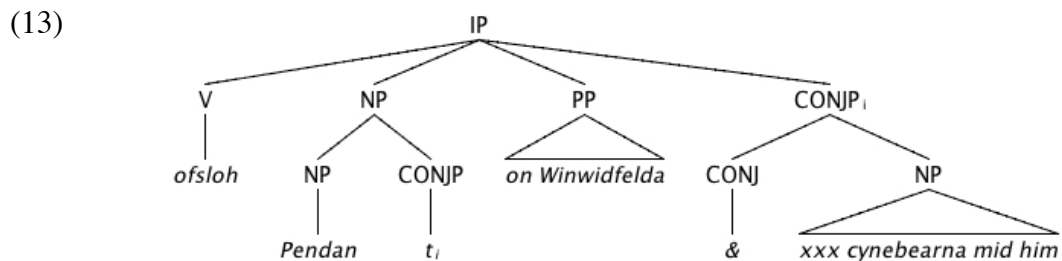
A simplified version of part of the YCOE parse of (12b) is given at (14). In this case the *mid*-PP is parsed as a constituent of a conjunct clause whose main verb is elided under identity with that of the previous clause.

(14)
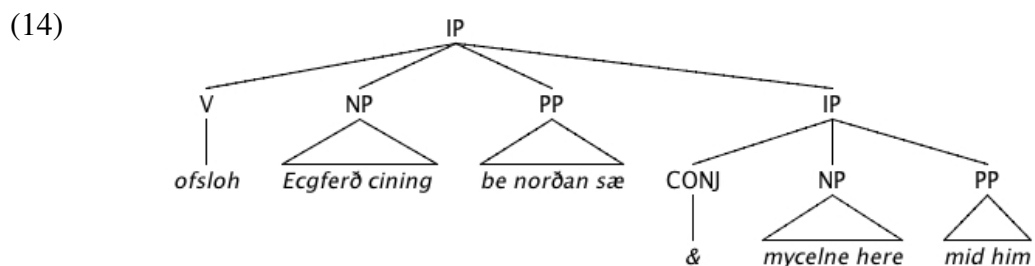
For most of the 157 examples represented by (11) and (12), either parse would appear to be entirely reasonable. The single-clause analysis given by the YCOE editors to the examples at (11) would be just as feasible for those at (12) given that coordinated NPs, especially those coordinated by *and* 'and', can be separated in Old English (Mitchell 1985: §§1464–72) and since heavy NPs are

often placed clause-finally (Pintzuk & Kroch 1985, van Kemenade 1987: 39–41). And although the overt verb agrees syntactically with its singular subject in (12a), this does not mean that *ealle englas* necessarily belongs to a separate clause: when a semantically singular subject NP is coordinated with, and separated from, a semantically plural subject NP, the verb may show singular agreement with the first conjunct (Mitchell 1985: §31), e.g. (15).

(15)　　... þa gastlican drohtnunga þe　Crist　syððan gesette.　　　& his
　　　　　the spiritual　reputation　that Christ later　established-SG and his
　　　　apostoli
　　　　apostles
　　　　'... the spiritual reputation that Christ and his apostles later established'
　　　　　　　　　　　　　　　　　　　　(cocathom1,ÆCHom_I,_25:384.153.4916)

On the other hand, the bi-clausal analysis given by the YCOE editors to the examples at (12) would be just as feasible for those at (11) given that verbs are not always repeated in subsequent clauses, even where a different morphological form of the verb would be required (Mitchell 1985: §1413). There is, however, no obvious reason why examples like those in (11) and (12) are not given the *same* analysis in the YCOE.

　　　　Parsing differences aside, one property all these 157 *mid*-PPs share is that right-of-P placement of the PPOP is invariable. The low frequency of left-of-P placement with *mid*-PPs in general has already been noted in the previous chapter, where an overall frequency of 8.6% was quantified (see Table 3.10). By this estimation, we might expect to find a dozen or so examples with a left-of-P object pronoun among the 157 *mid*-PPs represented by (11) and (12), so it is somewhat surprising to find none. The almost parenthetical nature of these 157 PPs makes it highly improbable that their PPOPs are stressed, so a phonological explanation for right-of-P placement seems unlikely. It also seems unlikely that right-of-P placement follows from the adjunct-like status of these PPs since *mid*-PPs occur with a left-of-P PPOP when functioning as an adjunct in other contexts, e.g. (16).

(16) a. Noman hi   eac swylce **him** wealhstodas of   Franclande **mid**

took   they also likewise them interpreters from France   with

'They also likewise took interpreters from France with them'

<div align="right">(cobede,Bede_1:14.58.3.537)</div>

b. and hæfdon **him mid** twegen ormæte   dracan

and had   them with two   enormous dragons

'and had two enormous dragons with them'

<div align="right">(cocathom2,ÆCHom_II,_37:275.106.6197)</div>

With a small number of these 157 *mid*-PPs, the PP does seem to be a clause-level constituent rather than an NP modifier. In (17), for example, *forð* is difficult to parse without assuming that it modifies an elided instance of *faran* 'to go'. Similarly in (18), the presence of the two *to*-PPs indicates that each belongs to different realisations of *biegan* 'to convert'. But there are less than a handful of examples like these among the 157-*mid* PPs in question, so the fact that their PPOPs are invariably right-of-P is not necessarily significant.

(17)   Her   on þysum geare for   Swegn   eorl into Wealan, and Griffin se

Here in this   year went Sweyne earl into Wales   and Griffin the

norþerna cyng forð **mid him**

northern king forth with him

'In this year Sweyne went into Wales, and Griffin the northern king (went) forth with him'

<div align="right">(cochronC,ChronC_[Rostizke]:1046.1.1834)</div>

(18)   and gebigde   þone cynincg Kyneglys to Gode, and ealle his leode   to

and converted the   king   Cynegils to God   and all   his people to

geleafan **mid him**

faith   with him

'and converted King Cynegils to God, and (converted) all his people to faith with him'

<div align="right">(coaelive,ÆLS_[Oswald]:128.5454)</div>

There are, however, another 23 examples that fit the non-lexical criteria by which the 157 *mid*-PPs were identified. As with the *mid* examples, some are parsed as the constituent of a verbless clause, e.g. (19), and some as a constituent of a coordinated NP, e.g. (20). In some cases, the PP expresses accompaniment, as in (19a) and (20a), but some express serial ordering, as in (19b) and (20b). Again, none of these examples involves a left-of-P PPOP.

(19) a. &   þær   wearð Sigulf ealdormon ofslægen, &   Sigelm ealdormon, &
          and there was   Sigulf chief   slain   and Sigelm chief   and
          Eadwold cynges ðegen, &   Cenulf abbod, &   Sigebreht Sigulfes sunu,
          Eadwold king's thain   and Cenulf abbot   and Sigebreht Sigulf's son
          &   Eadwald Accan sunu, &   monige **eac   him**
          and Eadwald Acca's son   and many   besides them
          'and there was killed chief Sigulf, and chief Sigelm, and Eadwold the
          king's thain, and Abbot Cenulf, and Sigulf's son Sigebreht, and Acca's
          son Eadwald, and many besides them'
                                        (cochronA-2b,ChronA_[Plummer]:905.11.1187)
     b. ... þæt eow sy well &   eowrum bearnum **æfter eow**
            that you be well and your   children after   you
          '... that it shall be well with you, and your children after you'
                                                       (cootest,Deut:4.40.4564)

(20) a. &   on mæigðhade wunode   Iohannes se   fulluhtere. þe   embe Crist
          and in chastity   continued John   the baptist   that about Christ
          cydde.   &   manega oðre   **toeacan him**
          testified and many   others besides him
          'and John the baptist, who testified about Christ, continued in chastity,
          and many others besides him'
                                        (cocathom1,ÆCHom_I,_9:255.198.1751)
     b. And þu   healtst min wed   &   ðin ofsprinc **æfter ðe** on heora
          and you keep   my covenant and your offspring after you in their
          mægðum
          generations

'And you will keep my covenant, and your offspring after you in their generations'

<div align="right">(cootest,Gen:17.9.646)</div>

I suggest that the crucial difference between the examples at (11), (12), (19) and (20), on the one hand, and those at (16), on the other, is that in the former examples the PP is an NP modifier but in the latter it is a VP constituent. Support for this treatment is found in Mitchell (1985: §1413), who recognises that at least some PPs expressing accompaniment 'must be construed with nouns or pronouns, and not with verbs', contrasting examples like those in (11) and (12) with examples like those in (21), which Mitchell construes with the verb.

(21) a. ... þæt he geseah Ceaddan sawle his broðor　**mid engla weorude** of
　　　　that he saw　Chad's　soul　his brother's with angels' throng　from

　　heofonum astigan
　　heaven　descend

　　'... that he saw the soul of his brother Chad descend with a throng of
　　angels　from heaven'

<div align="right">(cobede,Bede_4:3.270.21.2749 [Mitchell 1985: §1412])</div>

　b. ... mid his þegnum, þe　**him mid** wæron
　　　with his servants　that him with were

　'... with his servants, who were with him'

<div align="right">(cobede,Bede_3:2.158.6.1520 [Mitchell 1985: §1412])</div>

This distinction is potentially significant as it is generally accepted that in Present Day English preposition stranding is generally impossible when the PP functions as an NP modifier, as in (22), although is generally possible when the PP functions as an NP complement, as in (23) (e.g. Takami 1992: 51–88, Huddleston, Pullum & Peterson 2002: 1092–3).[43]

(22) a. *Which city$_i$ did you meet a man from __$_i$?
　b. *How many credits$_i$ are you a student with __$_i$?

---

[43] The examples at (22) and (23) are my own.

(23) a. <u>Which city</u>ᵢ did you witness the destruction of __ᵢ?

　　b. <u>How many subjects</u>ᵢ is he an expert in __ᵢ?

　　　　The constraint on stranding when the preposition modifies a noun has been accounted for in terms of Subjacency (e.g. Chomsky 1986). Where PP is L(exically)-marked by a lexical category, e.g. N, PP is not a blocking category and hence not a barrier, and so *wh*-movement of the preposition's object does not violate the Subjacency condition. This is the case in (23a), for example, where the PP is L-marked by *destruction*. Where the PP is sister to a non-lexical category, e.g. N', PP <u>is</u> a blocking category, and hence a barrier, and so *wh*-movement of the preposition's object violates Subjacency, as in (22).

　　　　The Subjacency condition could certainly provide a principled account of why none of the 180 PPs represented by the examples at (11), (12), (19) and (20) occurs with a clearly PP-external PPOP, i.e. why there are no examples like (24). On the assumption that the PP is not L-marked, e.g. by *folc*, movement of the pronoun out of its PP would violate Subjacency.

(24)　　*... þæt Nicanor feol and **him** eall his folc **mid**

Subjacency could also provide a principled account of the absence of any examples in which the pronoun appears immediately to the preposition's left, as in (25), but only at a rather significant cost.

(25)　　*... þæt Nicanor feol and eall his folc **him mid**

In order for Subjacency to be able to account for the (assumed) ungrammaticality of (25), it must be assumed that placement to the immediate left of a governing preposition also involves movement out of PP. As discussed in Chapter 2, such an assumption has little support. Two-thirds of PPOPs appear immediately to the preposition's left despite variation in PP placement, and this is difficult to explain other than by assuming that left-of-P PPOPs are clitics and that there is a PP-internal clitic position to the left of P^0 available to clitic PPOPs. While treatments of Old English clitics differ as to whether clitics can attach syntactically to phrasal

heads, as proposed by van Kemenade (1987: 126–33), or whether they attach to phrases, as claimed by Pintzuk (1991: 234–6, 1996: 383–5), those that specifically consider the placement of PPOPs agree that, when adjacent to their governor, clitic PPOPs remain within PP (van Kemenade 1987: 133, Pintzuk 1991: 276–7, 1996: 384). The Subjacency condition cannot therefore be held to account for the absence of examples like (25) without abandoning the analysis that allows the placement of two-thirds of special clitic PPOPs to be straightforwardly accounted for. And if Subjacency does not explain why there are no examples like (25), then it cannot be held to for the absence of examples like (24) without assuming that the absence of examples like (25) is accidental while the absence of examples like (24) is due to ungrammaticality. Examples involving R-pronouns do not help us out here. There are few examples in total in which the PP could potentially be analysed as the modifier of a noun. In most cases, the R-pronoun is placed immediately before the preposition, e.g. (26). If these PPs do in fact function as NP-modifiers then, under a clitic analysis of *þær*, these examples would suggest that clitics can indeed appear left-adjacent to a governing preposition, which would support treating the absence of examples like (25) as accidental.

(26) a. and þa   geworhte weal mid turfum, and bredweal **þæronufon**
        and then made   wall with turves and palisade thereupon
        'and then made a rampart with turves, and a palisade thereupon'
                                    (cochronD,ChronD_[Classen-Harm]:189.1.78)

   b. &   heo Ierusalem þa burh æft   aræardan, & þet tempel **ðærbinnæn**
        and she Jerusalem the city likewise raised   and the temple there-within
        'and she likewise built Jerusalem the city, and the temple within it'
                                    (colsigewB,ÆLet_4_[SigeweardB]:726.188)

   c. ONGYTAÐ ðæt God sette to dæg beforan eow lif &   god, &   **ðær**
        know   that God set to day before you life and good and there
        ongean deað &   yfel
        against death and evil
        'Be aware that God set before you this day life and good, and conversely death and evil'
                                    (cootest, Deut:30.15.4958)

There is, however, one example in which the R-pronoun is clearly outside its PP, (27). Again, on the assumption that this PP modifies a noun — which certainly seems a possibility — then, under a clitic analysis of *þær*, this example is in clear violation of Subjacency. However, as there is only example, which could conceivably involve an elided main verb instead, it does not tell us anything much about the significance the absence of examples like (24).

(27)   &   he þa   þam biscope gesealde in æht   þreo hund   hida, &   **þær**
        and he then the   bishop   gave   in goods three hundred hides and there

        eahta **to**
        cattle to

        'and he then gave the bishop in goods three hundred hides, and cattle besides'

                                      (cobede,Bede_4:18.306.29.3112)

Overall, then, it seems that when the pronoun belongs to an NP-modifying PP, left-of-P placement is possible with R-pronouns but not with personal pronouns. Although this finding for PPOPs looks like very like an effect of the Subjacency condition, the placement of special clitic PPOPs in general does not support the idea they are always occur PP-externally, which is a necessary assumption for the Subjacency argument to work here.

The examples not yet examined do not include any that run counter to the view that left-of-P placement is precluded when the PP modifies a noun. Of the three left-of-P PPOPs belonging to a PP that is parsed as the constituent of a non-verbal element in the YCOE, none is unambiguously the modifier of a noun. One is parsed as a constituent of an element tagged 'X', where X denotes a clause that exhibits textual problems, such as missing words or lines. For this example, (28), there is insufficient material to determine how the PP should be parsed.

(28)   ... þæt ic me sylf onfand þæt Iudeas hie   sylfe   þurh   æfeste **him**
           that I   me self found   that Jews   them selves through rivalry them

        betweonan [text missing]
        between

                         (coblick,LS_32_[PeterandPaul[BlHom_15]]:177.100.2243)

In the second example, (29), the PP is parsed as a constituent of an ADJP. Note that the YCOE editors do not treat *wið earme men* as a constituent of *ardæde* in this example, and I think it is just as reasonable to treat the emboldened PP as independent of *eaþmode*.

(29) Uton beon ælmesgeorne & ardæde wið earme men, & eaþmode
      let-us be   charitable   and merciful with poor   men and benevolent
      us betweonan
      us between
      'Let us be charitable and merciful with poor men, and benevolent among ourselves'

                          (coblick,HomS_46_[BlHom_11]:131259.1600)

In the third example, (30), the YCOE editors parsed the PP as dependent on the adverb *nean*, which would translate as '(came) near to him'. However, *nean* can also mean 'from nearby' and this reading contrasts nicely with *feorran*, which occurs a few clauses later and which is parsed in the YCOE — correctly, in my opinion — as independent of the *to*-PP with which it co-occurs.

(30) Eac ða ðe of iudeiscum folce on Crist gelyfdon. comon **him**
      also those that of Jewish   people in Christ believed came   him
      nean   **to**. for ðan ðe hi wæron be   him gelærede ... Þa
      from-nearby to because   they were concerning him taught   ... those
      soðlice ðe gelyfdon on Crist of hæðenum folce. ða   comon him
      truly   that believed in Christ of heathen   people those came   him
      feorran   to
      from-afar to
      'Also those of the Jewish people who believed in Christ came to him from nearby, because they had been taught about him ... Truly those of the heathen people who believed in Christ, they came to him from afar'

                          (cocathom2,ÆCHom_II,_29:232.58.5156–7)

None of the seven PPs parsed as a constituent of a verbless clause and which occur with a left-of-P PPOP is evidently an NP-modifier either. In two examples, the PP is most naturally interpreted as belonging to a different realisation of the overt main verb, e.g. (31).[44]

(31) Efne   ða   on middre nihte com   sum harwencge mann into þam
      exactly then on middle night came some hoary   man   into the
      cwearterne, and his cnapa   **him ætforan**
      prison   and his servant him before
      'Then exactly at midnight a hoary man came into the prison, and his
      servant (came) in front of him'
      ??'Then exactly at midnight a hoary man and his servant before him came
      into the prison'

                                        (coaelive,ÆLS[Agatha]:131.2093)

In another three examples, the only lexical item on which the PP could possibly depend is a predicative adjective, e.g. (32), although there is no reason to think that is necessarily the case.[45]

(32) ... þæt hi   æþele cempan wæron, and on ælcum gefeohte fæstræde
       that they noble soldiers were   and in each   battle   steadfast
      him betwynan
      them between
      '... that they were noble soldiers, and steadfast between themselves in
      every battle'

                                        (coaelive,ÆLS[Forty_Soldiers]:19.2479)

The last two examples with a left-of-P PPOP involve *to gamene*(*s*) in an incomplete comparative clause:

---

[44] The other example is at coaelhom,ÆHom_10:38.1428.

[45] The other two examples are at cocathom1,ÆCHom_I,_17_[App]:540.161.3298 and coaelive, ÆLS_[Maurice]:132.5760.

(33)   and dwollice leofast, swylce þe   togamenes

          and foolishly live   as   you to sport

          'and live foolishly, as if for your amusement'

<div align="right">(colwgeat,ÆLet_6_[Wulfgeat]:140.57)</div>

(34)   Hi   scuton þa mid gafelucum swilce **him** to gamenes **to**

          they shot   then with javelins   as   them to sport   to

          'they then shot (at him) with javelins, as if for their amusement'

<div align="right">(coaelive,ÆLS_[Edmund]:116.7031)</div>

In (33), *togamenes* is parsed as the prepositional governor of *þe*, although no such preposition is listed in any of the Old English dictionaries. In the entry for *gamen* n., the *Dictionary of Old English* (Cameron, Amos, Healey *et al.* 2007) (the '*DOE*') lists *to gamene*(*s*) 'for the amusement of (someone)', so it appears the YCOE editors have mistaken a PP for a preposition on this occasion. The *DOE*'s definition additionally indicates that *to gamene*(*s*) requires an NP complement, so in (33) we appear to have a PP, *to gamenes*, with its NP complement, *þe*. In (34), however, the second instance of *to* is unexpected under the *DOE*'s analysis, but even rejecting the idea of scribal error, it may be argued that the second *to*-PP is the complement rather than modifier of *gamenes*.

      Overall, the data indicate that right-of-P placement is invariable when the PP modifies a noun. The data further indicate that left-of-P placement is in fact possible when the PP belongs to a clause with an elided main verb, e.g. (31) and (32). The reason why the univariate analysis revealed this to be a near knockout factor is because a large proportion of PPOPs appearing in this context more probably belong to an NP-modifying PP.

4.4 Case

4.4.1 Introduction

This section considers the relationship between pronoun case and pronoun placement. Although ultimately I am unable to explain why PPOPs are rarely realised as special clitics unless dative, I do rule out some of the more likely explanations. I begin by looking closely at the placement of non-dative PPOPs

and argue that the number, and consequently the proportion, placed to the preposition's left is probably even lower than quantified in the previous chapter (section 4.4.2). I then examine the behaviour of PPOPs governed by prepositions that do not strongly favour one case over another and show that the correlation between pronoun placement and pronoun case cannot be attributed to lexical effects of the preposition nor probably to PP semantics either (section 4.4.3). In section 4.4.4 I consider and reject arguments made by Colman (1991) that seek to explain why left-of-P objects in Old English tend to be both pronominal and dative. Finally, I look for — and find — evidence of a correlation between pronoun placement and case among simple personal pronouns governed by verbs (section 4.4.5). The findings of this section are summarised in section 4.4.6.

4.4.2 Non-dative PPOPs

The univariate analyses have shown that the most frequently occurring context in which right-of-P placement can be correctly predicted at least 95% of time is defined by pronoun case. These results, given in Table 3.7, show that PPOPs rarely occur in a left-of-P position unless they are clearly or very probably dative.

There are very few genitive PPOPs in total (N=31). Most (N=23) are governed by *wið*, e.g. (35a), seven are governed by *tomiddes*, e.g. (35b), and one is governed by *toweard*, (35c).

(35) a. and efne   þær   swam an næddre **wið**   **heora**

      and behold there swam a   serpent towards them

      'and behold there swam a serpent towards them'

                           (coaelive,ÆLS_[Martin]:1259.6794)

   b. &   setton hig **tomiddes hyra**

      and set   her amongst   them

      'and set her amongst them'

                           (cowsgosp,Jn_[WSCp]:8.3.6367)

   c. Ða   þe cyng undergeat ealle þas   þing. &   hwilcne swicdom

      when the king understood all   these things and what   treachery

      hi   dydon **toweard his**, þa ...

      they did   against   him then

> 'When the king understood all these things and what treachery they were
>
> employing against him, then ...'

<div align="right">(cochronE,ChronE_[Plummer]:1087.37.3003)</div>

Given such a small number of genitive PPOPs, it is impossible to draw any firm conclusions about their invariable right-of-P placement. A different sample might produce some counter-examples, although there are none in the York Poetry Corpus (Pintzuk & Plug 2001).[46] On the other hand, the lack of a single left-of-P example among the 31 genitive PPOPs is clearly out of step with dative PPOPs, one-third of which are situated to the preposition's left.

Although accusative PPOPs are sometimes situated to the left of their governor, the proportion placed left-of-P (4.5%) is very small in comparison to dative PPOPs. The text files with the largest share of accusative special clitic PPOPs are: cobede (*Bede*) and cogregdC (*Gregory's Dialogues*, ms. C), x5 each; coorosiu (*Orosius*) x4; and coboeth (*Boethius*), cocathom2 (*Catholic Homilies II*) and cocura (*Cura Pastoralis*) x3 each. The other sixteen examples come from thirteen different text files. Having closely examined all 39 examples and having checked them against their particular base editions, I find reason to be less than confident about their status as accusative special clitic PPOPs in 24 cases (62%). I give three questionable examples here for the purposes of illustration: the remainder are given in Appendix E. If my suggested analyses for these 24 examples are accepted, then the proportion of genuine accusative special clitic PPOPs would reduce from the already low figure of 4.5% (i.e. 39/868, see Table 3.7) to just 1.8% (i.e. 15/844).

There are various types of ambiguity that could cause the number of accusative special clitic PPOPs to be overstated. I give three examples here. In the first example, (36), the accusative pronoun could be understood as a verbal object and the preposition as stranded in a relative clause.

---

[46] The York Poetry Corpus is considerably smaller than the YCOE. Its component texts, which represent a range of authors and dates of composition, amount to some 71,500 words.

(36) and him geedniwode þa ylcan lare   þe he ær   **hi   mid**

and them restored   the same doctrine that he previously them with

lærde

taught

'and restored to them the same doctrine with which he had previously

taught them'

<div align="right">(cocathom2,ÆCHom_II,_18:170.27.3761)</div>

That *mid* is stranded rather than the governor of *hi* is especially likely since
mid+acc. is an Anglian feature (Mitchell 1985: §1195) and so is unlikely to be
found in this West Saxon text (*Ælfric's Homilies II*). Indeed, out of some 1,100
mid-PPs in this text file, the only other example in which *mid* is parsed with an
acc. object involves an NP headed by a fem. noun, which could be (sg.) dat. rather
than (pl.) acc.:

(37) and he ealle gefæstnode heora fet to eorðan ... **mid his strangan**

and he all   fastened   their feet to ground ... with his powerful

bene

prayer(s)

'and he fully fastened their feet to the ground ... with his powerful

prayer(s)'

<div align="right">(cocathom2,ÆCHom_II,_39.1:292.154.6633)</div>

In another three examples, represented by the one at (38), the word parsed
as the prepositional governor of the accusative pronoun is *on*, which is situated
immediately to the left of a form of *becuman*.[47] *Becuman* has a range of senses
including 'to come, approach, meet with, happen, befall', but Clark Hall (1960)
and the *DOE* list *onbecuman* as a derivative, although neither provide a definition.
I give two translations for (38): my own and that of Orchard (2003).

---

[47] The other two examples are at coalex,Alex:14.6.123 and cobede,Bede_1:9.46.6.393. One of the
other Old English witnesses to the latter example has dative <him> (Miller 1898: 21) where the
base manuscript has accusative <hi> (Miller 1890: 46, ll. 6–8).

(38) Ða cwædon men þætte hie wendon þæt þæt wære goda eorre þæt

then said   men that   they thought that that was   gods' anger that

usic þær **on** becwome

us   there on happened

'Then men said that they thought that it was the anger of the gods which

had befallen us'

(coalex,Alex:30.18.385)

'Then men said that they thought it was the anger of the gods which had

fallen upon us'

(Orchard 2003: 245, §30)

In support of my own translation, which treats *on* as a verbal prefix, I give the
example at (39). This example has two collocations of *on* and *becuman*, and
placement of the verbal negator before *on* in the second instance provides the best
possible evidence for a prefixed verb (Elenbaas 2006: 122). Although the object
of *onbecuman* in (39) is dative rather than accusative in both instances, that does
not preclude treating *on* as a verbal prefix in the three examples represented by
(38) since verbal rection is not entirely consistent in Old English (Mitchell 1985:
§1081).

(39) Wenstu   nu þæt þe anum þellecu hwearfung & þillecu unrotnes

think-you now that you alone such   change   and such   sorrow

on becume &   nanum oðrum mode swelc **ne on become**, ne   ær

befall   and no   other mind such not befall   neither before

þe ne æfter þe?

you nor after you

'Do you now think that such change and such sorrow should befall you
alone, and (that) such should not befall any other mind, either before you
or after you?'

(coboeth,Bo:8.20.17.332)

141

I agree that in the third example, (40), the pronoun in question is probably a PPOP, but as a masc. sg. acc. form in particular, I see no possible antecedent. Scribal error is therefore a distinct possibility in this case.

(40)   us is eac to witanne þæt þæt wæs þearfendra manna asægdnesse in
        us is also to know   that that was poor   mens' offerings   in
        þære ealdan æ   þæt hie   sceoldon þy dæge bringan twegen turturas
        the   old   law that they must   by day bring   two   turtle-doves
        oððe twegen culfran briddas Gode to asægdnesse. Swylce asægdnesse
        or   two   culver birds   God   to offering.   such   offering(s)-FEM
        Cristes   aldoras **hine mid** brohton to þam Godes temple
        Christ's elders   him with brought to the   God's temple
        'It is also known to us that that was the offerings of poor men (i.e.) that
        they had to bring two turtle doves or two culver birds to God as an
        offering. Such offering(s) Christ's elders brought with him(?) to the
        temple of God'
                                 (coverhom,LS_19_[PurifMaryVerc_17]:67.2183–4)

Two of the fifteen examples which almost certainly involve a special clitic accusative PPOP are given below. Another involving the same combination of preposition and verb as in (41) occurs in the same text (coorosiu, Or_2:8.51.19.982), and another involving the same combination of preposition and verb as in (42) occurs in that same text (cobede,Bede_1:18.92.14.848). The other eleven examples with a seemingly genuine special clitic accusative PPOP are given at Appendix F.

(41)   &   þa   nihtes   on ungearwe **hi   on** bestæl
        and then by-night unwares   them on stole
        'and then stole upon them in the night unawares'
                                 (coorosiu,Or_1:10.30.16.594)

(42) Þa   teah Penda **hine**[48] fyrd **on** &   here, &   hine his rices   benom

      then led   Penda him   army on and host and him his kingdom took

      'Then Penda led his army and host against him, and deprived him of his

      kingdom'

                                                    (cobede,Bede_3:5.168.20.1638–9)

In contrast to the prose, there is no statistical evidence that dative and accusative PPOPs behave differently in the York Poetry Corpus, although the total number of clear accusative examples in that corpus is very small. Of the 88 simple, i.e. unmodified and uncoordinated, clear dative PPOPs in this corpus of Old English poetry, 59 (67%) are in a left-of-P position, compared to 8 (47%) of the 17 simple clear accusative PPOPs.[49] None of the 8 clear accusative pronouns parsed as left-of-P PPOPs in the poetry is amenable to any obvious alternative analysis. Recall, however, that left-of-P placement is much more frequent with full NP objects in the poetry than in the prose (see Chapter 1) so it is perhaps not surprising to find that left-of-P placement of accusative PPOPs is also relatively more frequent in the poetry than in the prose. Nevertheless, the trend in the poetry for left-of-P placement by pronoun case is not out of step with the trend in the prose, although the difference in the poetry is not statistically significant ($\chi^2 = 2.46, p = 0.117$).[50]

4.4.3 Case-alternating prepositions

Table 4.1 identifies four prepositions that (a) alternate between dative and accusative government and (b) govern at least ten clear dative and at least ten clear accusative PPOPs in contexts where PPOP placement is not constrained by other factors discussed in this chapter. Data for each of these prepositions tell much the same story: left-of-P placement occurs frequently when the pronoun is

---

[48] Ms. B has dative <him> (Miller 1898: 173). Likewise for the example at cobede,Bede_1:18.92.14.848, the base ms. has <hine> but ms. B. has <him> (Miller *ibid*: 77).

[49] The 8 left-of-P accusative PPOPs in the York Poetry Corpus can be found as follows: cogenesi, 34.1040.278 and 34.1044.280; coriddle, 22.3:181.10.39, 22.27.194.9.415, 22.55:208.14.705 and 34.20:235.3.788; cobeowul,23.688.580; and cochrist:CHRIST_III,31.1007.690.

[50] All chi square values are calculated using Lowry's (2010) online resource. For significance at the 0.05 level, a chi square value ≥ 3.84 is required.

dative but comparatively infrequently, if at all, when accusative. Note that the data for *on* exclude the three accusative examples given an alternative analysis in the previous section (see the discussion of the example at (38)) as well as the six accusative examples given an alternative analysis in Appendix E.

Table 4.1 Left-of-P frequency by preposition and case

Preposition	Clear dative		Clear accusative	
	N	Left-of-P	N	Left-of-P
ongean 'towards, against'	105	101 (96%)	58	0
wið 'against, with'	91	67 (74%)	165	0
ofer 'over'	11	8 (73%)	65	1 (2%)
on 'on, in'	457	186 (41%)	166	7 (4%)

Also omitted from Table 4.1 are the 1,300 or so PPOPs governed by *mid* 'with' (dat. x1,252, acc. x46). Left-of-P placement is rare with *mid* in any event (see Table 3.10). Further, *mid*+acc. is an Anglian feature (Mitchell 1985: §1195) and so alternates with *mid*+dat. in very few text files. Cobede (*Bede*) alone supplies more than a handful of both dative and accusative *mid*-PPOPs. The placement of *mid*-PPOPs in *Bede* is consistent with the trend evident in Table 4.1: of the 50 dative pronouns, 3 are left-of-P (6%), while all 32 accusative pronouns are right-of-P. The data in Table 4.1 and the data for *mid* in *Bede* clearly indicate that the constraint against left-of-P placement of accusative PPOPs cannot be attributed to a lexical effect of the particular preposition involved.

Like many Old English prepositions, many Modern German prepositions vary between dative and accusative government. In German, this alternation tends to mark a distinction between a locative reading, as in (43), and a directional reading, as in (44).

(43)   Diana schwamm im   See
        Diana swam   in-the-DAT lake
        'Diana swam in the lake'

(Gehrke 2008: 96, ex. 23a)

144

(44)   Diana schwamm in den   See

        Diana swam   in the-ACC lake

        'Diana swam into the lake'

Alternations between locative and directional semantics are also found with certain Modern Dutch prepositions, except that the different readings are derived from the order of the PP's constituents:

(45)   Willemijn zwom in het meer

        Willemijn swam in the lake

        'Willemijn swam in the lake'

(46)   Willemijn zwom het meer in

        Willemijn swam the lake in

        'Willemijn swam into the lake'

A number of scholars (e.g. Gehrke 2008, Koopman 2010, den Dikken 2010) propose that such alternations reflect a difference in PP structure, with directional PPs claimed to be structurally more complex than locative PPs. Could this have something to do with the apparent constraint against special placement of non-dative PPOPs in Old English?

While there is a tendency for accusative PPs to denote 'motion towards' and for dative PPs to denote 'location at which' in Old English, these tendencies are not consistently observed (Mitchell 1968: 294, 1985: §1177(4), Traugott 1992: 202–3). Moreover, this tendency does not describe the semantics of a large share of the PPs quantified in Table 4.1. For example: both *ongean* and *wið* tend to denote opposition regardless of case; about half of the *on*+acc. examples are complements of *liefan*, giving non-spatial 'to believe in'; and while most of the *ofer* + dat. examples do seem to denote a locative relationship, as in (47a), a locative reading is possible with *ofer* + acc. too, e.g. (47b).

145

(47) a. & worhton mid stanum ænne steapne beorh him ofer

and made with stones a high mound him over

'and made a high mound with stones over him'

(cootest,Josh:7.26.5366)

b. Ða wæs his ofergewrit ofer hine awriten greciscum stafum & ebreiscum

then was his inscription over him written Greek letters and Hebrew

'Then his inscription was written above him in Greek and Hebrew letters'

(cowsgosp,Lk_[WSCp]:23.38.5609)

In short, I have found no evidence of any regular semantic distinction between the dative examples and the corresponding accusative examples reported in Table 4.1. Instead, I found many pairs of examples that vary by case but not, apparently, by meaning, e.g. (48)–(57). Each of these pairs of examples is drawn from the same text file and, in each, the combination of verb and preposition is held constant. None of these case alternations correlates with any obvious difference in meaning to a modern reader, but all correlate with a difference in PPOP placement.

(48) a. and se casere eode **ongen hine** and cyste hine

and the emperor went towards him-ACC and kissed him

'and the emperor went towards him and kissed him'

(coeust,LS_8_[Eust]:287.306–7)

b. þa eode se casere **him ongean** swa hit þeaw is mid Romanum

then went the emperor him-DAT towards as it custom is with Romans

'then the emperor went towards him, as it is custom with the Romans

(coeust,LS_8_[Eust]:394.422)

(49) a. and his scypu wendon ut abutan Legceaster and sceoldan cuman **ongean**

and his ships turned out about Chester and should come towards

hine, ac hi ne meahton

him-ACC but they not could

'and his ships turned out around Chester and should have come towards him, but they could not'

(cochronC,ChronC_[Rositzke]:1000.1.1311–12)

b. þa com **him** swilc wind **ongean** swilc nan man ær ne

then came them-DAT such wind towards as no man before not

gemunde

remembered

'then came towards them such a wind as no man remembered before'

(cochronC,ChronC_[Rositzke]:1009.15.1438)

(50) a. Martinus ferde hwilon to Ualentiniane þam casere, wolde for sumere

Martin travelled once to Valentinian the emperor wished for some

neode **wið hine** spræcan

business with him-ACC speak

'Martin was travelling once to Valentinian the emperor, wishing to speak

with him about some business'

(coaelive,ÆLS_[Martin]:650.6389–90)

b. Maximus, se casere þe wæs on Martinus dæge, gelaðede foroft

Maximus the emperor that was in Martin's day asked frequently

þone arwurðan wer þa ða he **him** **wið** spræc þæt he wære his

the venerable man when he him-DAT with spoke that he was his

gemetta

guest

'Maximus, who was the emperor in Martin's day, frequently asked the

venerable man whenever he spoke with him that he be his guest'

(coaelive,ÆLS_[Martin]:610.6365)

(51) a. Her on þissum geare se cyng gerædde & his witan. þæt man sceolde

here in this year the king decided and his council that one should

gafol gyldon þam flotan. & frið **wið hi** geniman

tribute pay the fleet and peace with them-ACC make

'In this year the king and his council decided that tribute should be paid to

the fleet and peace made with them'

(cochronE,ChronE_[Plummer]:1002.1.1624)

b. And þy ilcan geare com mycel hæðen here on Angelcynnes land. &
and the same year came large heathen army into England and

wintersetle namon æt East Englum. & þær gehorsade wurdon &
winter-quarters took at East Anglia and there horsed were and

hi **heom** **wið** frið genamon
they them-DAT with peace made

'And the same year a large heathen army came into England and
appropriated winter-quarters at East Anglia and there were provided with
horses, and they made peace with them'

(cochronE,ChronE_[Plummer]:866.1.1079–82)

(52) a. & he raðe þæs **wið** **hie** gefeaht mid sciphere, &
and he soon afterwards against them-ACC fought with fleet and

ofslagen wearð
killed was

'and soon afterwards he fought against them with a fleet and was killed'

(coorosiu,Or_4:6.92.28.1881-2)

b. & þa nihstan landleode on ægþere healfe him on fultum geteah, oþ
and the nearest natives on either half him in support drew until

Somnite **him** gefuhton **wið**, & þone cyning ofslogon
Samnites them-DAT fought against and the king killed

'and drew the nearest natives on either side to him in support, until the
Samnites fought against them and killed the king'

(coorosiu,Or_3:7.60.27.1166)

(53) a. & spætton **on hyne**
and spat on him-ACC

'and spat on him'

(cowsgosp,Mt_[WSCp]:27.30.2051)

b. & spætton **him** **on**
and spat him-DAT on

'and spat on him'

(cowsgosp,Mk_[WSCp]:15.19.3460)

(54) a. ... for þon þe he eac wolde **on hine** winnan

because he also wished on him-ACC make-war

'... because he wished to make war on him'

(coorosiu,Or_6:15.142.10.2983)

b. ... for þon þe Dorus Thracea cyning **him** eac **an** wann

because Dorus Thrace's king him-DAT also on made-war

'... because Dorus, king of Thrace, also made war on him'

(coorosiu,Or_3:11.82.6.1638)

(55) a. Ac seo sunne scynð þeah **on hi**

but the sun shines still on them-ACC

'But the sun still shines on them'

(cosolilo,Solil_1:31.20.415)

b. þonne seo sunne **hym** **on** scynð, hi lyhtað ongean

when the sun them-DAT on shines they shine back

'when the sun shines on them, they shine back'

(cosolilo,Solil_1:31.17.413)

(56) a. Ac se deofol færinga eac **on hine** gefor

but the devil suddenly nevertheless in him-ACC went

'But nevertheless the devil suddenly entered into him'

(cogregdH,GD_1_[H]:10.73.1.711)

b. ... se deofol, þe **hyre** ær **on** gefor

the devil that her-DAT previously in went

'... the devil, who had previously entered into her'

(cogregdH,GD_1_[H]:10.73.22.718)

(57) a. Þa ongunnan ða geongan biddan þone bisscop, þa ðe **mid hiene**

then began the youths ask the bishop who that with him-ACC

wæron

were

'Then the youths, who were with him, began to ask the bishop'

(cobede,Bede_5:6.398.30.3999)

149

b. ... & from þæm he fulwihtes geryno onfeng mid his þegnum, þe

and from whom he baptism's sacrament received with his servants that

him mid wæron

him-DAT with were

'… and from whom he received the sacrament of baptism with his

servants, who were with him'

(cobede,Bede_3:2.158.6.1250)

Wende (1915: 77–81) supplies a number of further minimal pairs and remarks that dative PPOPs sometimes occur in a left-of-P position 'wo die syntaktische Gestaltung des Satzes die Wahl des Akkusativs eigentlich begünstigen sollte' ('where the syntactic formation of the sentence should actually favour the choice of the accusative' *ibid*: 81). This statement seems to imply, perhaps unintentionally, that accusative pronouns are sometimes replaced by dative pronouns when the pronoun is left-of-P. Wende does not mention, however, that right-of-P PPOPs are sometimes unexpectedly dative too. The examples in (58), from *Catholic Homilies I*, show variation in PPOP case for a frequently occurring verb + preposition combination that usually takes accusative in the YCOE.

(58) a. Se ðe **on hine** gelyfð he gesyhð hine nu mid his geleafan

he that in him-ACC believes he sees him now with his faith

'He who believes in him, he sees him now with his faith'

(cocathom1,ÆCHom_I,_9:253.136.1684)

b. His nama wæs Hiesus. þæt is hælend. for ðan þe he gehælð ealle ða þe

his name was Jesus that is saviour because he saves all those that

on him rihtlice gelyfað

in him-DAT rightly believe

'His name was Jesus, that is 'saviour', because he saves all those who

rightly believe in him'

(cocathom1,ÆCHom_I,_13:285.113.2455)

The examples in (59), from *West Saxon Gospels*, *John*, involve the same combination and, additionally, involve PPs which are both direct translations of Latin *in eum* 'in him-ACC'.

(59) a. Manega of   ðære menigeo gelyfdon **on hine**

     many   from the   multitude believed in him-ACC

     'Many among the multitude believed in him'

<div align="right">(cowsgosp,Jn_[WSCp]:7.31.6320)</div>

  b. Ða he ðas   ðing spræc manega gelyfdon **on him**

     as he these things said   many   believed in him-DAT

     'As he said these things, many believed in him'

<div align="right">(cowsgosp,Jn_[WSCp]:8.30.6429)</div>

So Wende's statement is factually correct but as it describes unsystematic variability in Old English case assignment only as it found among left-of-P PPOPs, it is potentially misleading.

If there is a semantic basis to the difference in probability of left-of-P placement between dative and non-dative PPOPs, it is not apparent from an examination of the data described in Table 4.1, yet these are the best data for a controlled comparison of PP semantics according to case. Examples (48)–(59) show that variation between P+acc. and P+dat. is not always systematic, which means that theories of PP syntax that associate syntactic differences with semantic differences are unlikely to prove helpful in accounting for the apparent constraint against specially placed non-dative PPOPs.

4.4.4   Colman (1991)

The frequency with which left-of-P prepositional objects are both pronominal and dative in Old English has been variously observed (see references in section 3.5.1.2). Colman (1991) alone suggests a possible explanation. The thrust of her argument is simple: she suggests that many of these dative pronouns are not prepositional objects at all. Her argument exploits the fact, discussed in section 3.2.2, that many Old English prepositions have the same form as a verbal prefix and/or adverb. Colman suggests that what appear to be left-of-P prepositional

objects may instead be one of two things: either the object of a prefixed verb, or an 'ethic' dative, which she describes as an ungoverned circumstantial locative expressing the location of interest. Colman makes no claim about what proportion of dative pronouns that would be regarded as left-of-P PPOPs under a traditional analysis might be regarded as ethic datives, suggesting only that it is a possibility 'in many instances' (*ibid*: 98). Given that some 2,662 dative pronouns are parsed as a left-of-P PPOP in the YCOE (see Table 3.7), we would have serious concerns about these data should Colman's suggestions be sustainable, so I consider each of her proposals in turn.

The first suggestion is that some left-of-P dative PPOPs can be explained by reanalysing the preposition as a verbal prefix and the pronoun as the object of the prefixed verb. The problem of distinguishing prepositions from prefixes and adverbs was discussed in section 3.2.2, where I offered direct evidence to support my conclusion that the proportion of elements wrongly tagged as prefix or adverb instead of preposition in the YCOE is likely to be negligible. I was unable, however, to offer direct evidence to support my belief that the proportion of elements wrongly tagged as preposition instead of prefix or adverb is also likely to be negligible because: (i) only a very small proportion of words labelled as prepositions do not have the same form as a prefix or adverb (or both), and (ii) there is no position in which a preposition can appear relative to its object and to other clausal constituents in which adverbs, if not also prefixes, may not also be found. Short of examining every preposition with a left-of-P PPOP in the YCOE, it is impossible to gauge how many examples are or might be incorrectly parsed. With some 2,662 left-of-P PPOP dative PPOPs to check in order to give full consideration to Colman's idea, this would be a long, laborious task with every likelihood that the editors' analysis would remain a plausible option in any event.

There is, however, a fundamental problem with Colman's suggestion. Her idea presupposes that verbal prefixes are mistaken for prepositions and that objects of prefixed verbs are mistaken for objects of prepositions. Let us accept that such mistakes are possible. Let us also accept Colman's assumption that verbs and their prefixes are syntactically inseparable (*ibid*: 56).[51] Under these

---

[51] Following Denison (1981: 57), Colman's view is that separable prefixes are best regarded either as prepositions or adverbs.

assumptions, the only sequence of prefixed verb and verbal object that *could* be mistaken for verb, preposition and prepositional object would be [OBJ (...) P V], where P represents a word classifiable as a preposition or inseparable prefix: in [P V (...) OBJ] configurations, P could not be analyzed as a preposition because right-of-P prepositional objects are always adjacent to P. This means that the hypothesised mistake would result in a falsely inflated number of prepositions with left- rather than right-of-P objects, and this point is central to Colman's argument. Further, of the 2,662 examples containing a word parsed in the YCOE as the prepositional governor of a left-of-P dative pronoun, the word in question immediately precedes the main verb in 1,421 cases. So, according to Colman's argument, just over half (53%) of what appear to be special clitic dative PPOPs could actually be the object of a prefixed verb. The problem, however, is that Colman offers no reason why the hypothesised mistake would result in a falsely inflated number of left-of-P objects that are (a) dative and (b) personal pronouns in particular. After all, there is no reason why nominal objects and non-dative object pronouns should appear to the left of a prefixed verb any less regularly than dative object personal pronouns. In short, Colman's first suggestion, i.e. that some left-of-P dative PPOPs can be explained by reanalysing the preposition as a verbal prefix and the pronoun as the object of the prefixed verb, presupposes a particular type of parsing error which, if made consistently, would surely produce greater numbers of nominal objects and non-dative personal pronoun objects to the left of a governing preposition than are found in the YCOE.

Colman's second suggestion, i.e. that what appears to be a left-of-P object of a preposition may instead be interpreted as an ungoverned circumstantial, seems more promising in that it offers an alternative analysis for dative objects in particular. Her suggestion builds on the fact that Old English has other circumstantial, i.e. ungoverned, datives, which express a range of circumstances, including manner, accompaniment, degree, measure, place, time and cause (Mitchell 1985: §§1408-27). Because of their denotations, however, these circumstantials are very often — if not typically, judging from Mitchell's examples — headed by a nominal. Nevertheless, other languages exhibit a type of circumstantial that expresses 'some sort of "interest"' and which, at least in German and Latin, is invariably realised as a dative pronoun (Colman 1991:

78–9). Colman suggests that a similar class of circumstantial can be recognised in Old English, and cites five examples as evidence (*ibid*: 84, b). In three of her examples, the dative pronoun is, as Colman herself acknowledges, amenable to various interpretations including indirect object, so I do not reproduce these examples here. The other two examples are given in (60). Both are from *Cædmon's metrical paraphrase of parts of the Scripture in Anglo-Saxon* (Thorpe 1832), i.e. a poetic text, and both involve *gewitan*, an intransitive verb, hence Colman's claim that the dative pronouns are ungoverned.[52] She suggests that these pronouns are interpretable as ungoverned circumstantial locatives (*ibid*: 86), expressing the location of interest. Adopting a term generally applied to such circumstantials, Colman refers to them as 'ethic' datives.

(60) a. Gewat him   ham   siðian

     went   him-DAT home go

     'went off home'

                        (Thorpe 1832: 130, ll.17–18 [Colman 1991: 84, ex. b])

    b. Him   Noe   gewat eaforan   lædan

     him-DAT Noah went   offspring lead

     'Noah went leading his offspring'

                        (Thorpe 1832: 82, ll.2–5 [Colman 1991: 84, ex. b])

Colman then examines around thirty examples involving what appears to be a dative object of P to see if they may be interpreted as ethic datives instead. Three of the examples that Colman judges to allow such an interpretation are given at (61)–(63). Having 'de-coupled' these dative pronouns from what appears to be their prepositional governor, Colman supplies an alternative analysis of the

---

[52] Colman references both examples to Bosworth & Toller (1898: *gewítan*), in which the original sentences are truncated. The full sentences are given below, with '/' indicating line breaks.

ii.   Gewat him þa   se healdend. / ham   siðian.
        went   him then the ruler   home go
        'The ruler then left to go home'

                        (Thorpe 1832: 130, ll. 17–18)

iii.   Him þa   Noe   gewat. /   swa hine nergend het. /   under earce-bord. / eaforan lædan.
        him   then Noah departed as   him saviour ordered under ark-   board leading offspring
        'Noah then left, as the preserver had ordered, under the ark-board, leading his offspring'

                        (Thorpe 1832: 82, ll.2–5)

preposition. For (61), Colman suggests that *fram* may be regarded as adverbial, more specifically a PP with an unspecified object, an analysis for which some independent evidence is provided (see also my discussion of prepositions with null objects in Section 1.5.2). Colman does not supply translations for her interpretations and I do not find her intended readings to be obvious. Accordingly, I simply gloss her examples as neutrally as possible and, following Colman, leave the reader to decide on the appropriate translation.

(61)   Se engel **him** gewat   fram

        the angel him departed from

    (cocathom2,ÆCHom_II,_28:221.20.4890 [Colman 1991: 59, ex. 8, *ibid*: 89])

For (62), Colman suggests *æfter* as the prefix of *ridan*, offering some independent evidence for *æfterridan* as an accusative-governing verb. Although this evidence supports her analysis of *him* in (62) as ungoverned, Colman does not comment on the absence of the object of the prefixed verb.

(62)   and **him** æfter rad

       and them after rode

  (cochronC,ChronC_[Rositzke]:879.15.746 [Colman 1991: 63, ex. 16, *ibid*: 93])

For (63), it is suggested that *to* could be either an adverb or prefix without any further discussion.

(63)   ... ðæt **him** ne magon to cuman ða speru

       that him not could to come the spears

        (cocura,CP:35.245.7.1601 [Colman 1991: 67, ex. 24, *ibid*: 95])

    Two problems with Colman's proposed analysis of the dative pronouns in (61)–(63) (and in her other examples) as ungoverned circumstantials are immediately obvious, both in principle and as a potential explanation for the difference in frequency of left-of-P placement according to pronoun case. Firstly, as Colman's aim is to offer an alternative analysis for the disproportionate number

of left-of-P prepositional objects that are (i) dative and (ii) personal pronouns, pronominality would seem to be a *desideratum* for her definition of ethic datives, but there is no part of her definition or analysis which predicts their pronominality. In her discussion of the referential properties of ethic datives (*ibid*: 78–9), Colman concludes that they may be co-referential with the subject in Old English. Such co-referentiality would naturally predict pronominality through reflexivity, but Colman expressly rejects the possibility that the object pronouns in examples at (60) are reflexive (*ibid*: 84–6), although she appears to have overlooked Mitchell (1985: §273), where the pronoun in *him gewitan* constructions is classed as a 'pleonastic' dative reflexive. Unlike 'necessary' dative reflexives, pleonastic dative reflexives occur with intransitive verbs, especially those implying motion or rest, and with transitive verbs that do not normally take a dative object (Mitchell 1985: §§271–4). In addition, three of the dative left-of-P objects she considers are full NPs, which suggests Colman accepts that her purported circumstantials may be nominal. Her rejection of these nominals as ethic datives for reasons unconnected to their nominality simply reinforces the implication that her proposed class of ethic datives need not be pronominal.

The second problem, which is independent of the first, is that no part of her account predicts that this type of circumstantial can be realised to the left but not the right of the word that could be mistaken for its prepositional governor. Admittedly, there is a tendency for Old English pronouns to come early in the clause, but if we allow that (61), for example, involves an adverb and ethic dative rather than a preposition and its object, then surely we must allow the same analysis for examples like (64)?

(64)　　&　he gewat　fram him
　　　　and he departed from him-DAT
　　　　'and he departed from him'

<div align="right">(cowsgosp,Lk_[WSCp]:24.31.5697)</div>

It seems to me that by accepting the principle of Colman's proposed analysis for (61), we must also accept that a considerable proportion of what appear to be

clear-cut examples of Old English dative PPs may not be PPs at all. Furthermore, we would also have to accept that there is no reliable way to distinguish between prepositions and their dative objects on the one hand and adverbs + ethic datives on the other.

The alternative to Colman's second suggestion, i.e. that some sequences of object and governing preposition could be reanalysed as ethic dative and verbal prefix, is more constrained. Colman limits its relevance to examples in which the preposition immediately precedes the verb, thus ruling it out for head-initial PPs. Nevertheless, there are some 1,400 PPs in the YCOE that would have to be individually examined in light of this suggestion, but I have little doubt that the YCOE's editors' analysis would remain a plausible option in all but a very few cases. For almost every P+V combination involved in these 1,400 examples, and certainly for all frequently occurring combinations, the YCOE supplies at least one other example involving the same combination in which the preposition precedes a dative object. Although this does not preclude Colman's analysis in principle, examples like (65) and (66) provide some indirect support for the YCOE editors' decision to treat examples like (62) and (63) as involving a preposition rather than a verbal prefix.

(65)   &   se cyning Ælfred **æfter þam   gehorsudan here**   mid fierde
        and the king   Alfred after the-DAT horsed-DAT troop-DAT with army
        rad   oþ   Exanceaster
        rode until Exeter
        'and Alfred the king rode after the mounted troop with an army as far as Exeter'
                                            (cochronA-1,ChronA_[Plummer]:877.3.866)

(66)   Þa   come **to ðam   bedde**   boc fram þam hælende
        then came to the-DAT bed-DAT book from the   saviour
        'Then a book came to the bed from the saviour'
                                            (coaelive,ÆLS[Julian_and_Basilissa]:64.976)

By exploiting the formal ambiguity of many Old English prepositions, Colman seeks to diminish the number of specially placed dative PPOPs, although as she herself acknowledges, each of her examples involving a dative PPOP is capable of more than one interpretation, including the very one she seeks to counter. Her proposed reanalyses are therefore 'possible' rather than 'better' alternatives at best. I am not at all convinced, however, that her suggestions offer an explanation for the fact that the vast majority of left-of-P objects of P are both dative and pronominal. Her suggestion that some prepositions and their dative objects could be reanalysed as verbal prefixes and verbal objects rests on the assumption of a parsing error which, if made consistently, *could* produce an overabundance of 'postpositions' but not an overabundance of dative PPOP-governing 'postpositions' in particular. That is not to say that all PPs in the YCOE are necessarily immune to such a reanalysis: my point is that I find it wholly implausible that the problem which Colman seeks to address, and which is clearly manifest from the YCOE editors' parsing of the texts, can be explained as a *consequence* of such parsing errors. Her second suggestion, which rests on a (somewhat poorly defined) class of 'ethic' datives, is no more compelling: it casts doubt on the status of constituents which Colman presumably has no quibble with, and leaves other examples — including some proportion of the dative left-of-P pronouns in my sample and all of the accusative left-of-P pronouns — unexplained.

4.4.5 Pronominal objects of verbs

Like PPOPs, simple personal pronoun objects of verbs ('PPOVs') often appear in positions from which functionally equivalent nominals are generally excluded. Also like PPOPs, PPOVs may be dative, accusative or genitive. Given that PPOPs are rarely specially placed unless dative, we should like to know whether the same is true of PPOVs.

Whereas special clitic PPOPs can easily be recognised from their position relative to their governor, the same is not true for PPOVs. As we have seen, objects of P almost always occur immediately to the preposition's right unless the object is a (dative) personal pronoun. Nominal as well as pronominal objects of verbs, however, may be found on either side of their governor due to the

possibility of movement of finite verbs from V^0 to I^0 or C^0 and to variation between underlying OV and VO word order (e.g. Pintzuk 1991: 177–84, 2002b, 2005, Koopman 2005, Pintzuk & Taylor 2006). As neither verb movement nor the internal structure of VP can be readily detected without reference to the position of other clausal constituents which may be absent or may themselves be ambiguously positioned, it is often impossible to determine whether a PPOV is in a special clitic position or not. For example, it is generally agreed that there is a special clitic position between topic and second-position finite verb in main clauses. This position, termed 'position A' by Koopman (1992), is one of five positions identified by van Kemenade (1987: 126–35) in which special clitic PPOVs can occur, as discussed in Chapter 2. Placement of a clitic object pronoun in position A produces an XOV, or verb third, word order, as in (67).

(67)　　Se Hælend hi　gefrefrode mid his fægerum wordum

　　　　the Saviour them cheered　with his beautiful words

　　　　'The Saviour cheered them with his beautiful words'

　　　　　　　　　　　　　　(cocathom2,ÆCHom_II,_23:200.34.4439)

One problem with recognising position A clitics, however, is that XOV word order can also transpire with a nominal object when the finite verb fails to materialise in second position, in as in (68).

(68)　　Ealle gesceafta scyppend ænne sunu gestrynde

　　　　all　creatures' creator　a　　son　begot

　　　　'The creator of all creatures begot a son'

　　　　　　　　　　　　　　(coaelive,ÆLS_[Cecilia]:156.7206)

Since finite verbs do not always materialise in second position in main clauses, especially in coordinated main clauses (e.g. Mitchell 1985: §§904–5, Traugott 1992: 277, Koopman 1995, Pintzuk & Haeberli 2008), we cannot tell whether the pronoun in (69), for example, is in position A or not as the verb's position is ambiguous between second and final position.

(69)   Se man him andswerode

        the man him answered

        'The man answered him'

<div align="right">(coapollo,ApT:2.4.15)</div>

Placement of a PPOV between topic and finite verb is therefore not necessarily diagnostic of the PPOV's status as a special clitic.

Of the other four special clitic positions identified by van Kemenade (*ibid*) for PPOVs, positions D (VP-initial) and E (to the left of V in VP) are not generally accepted because neither position excludes nominal objects of verbs (e.g. Pintzuk 1991: 222–3, 1996: 389–91, Koopman 1992: 53–4, 1997a: 79–80). Position B (immediately after the finite verb in clauses with inversion of all subject types, i.e. in operator-initial clauses) and position C (immediately after the complementiser in subordinate clauses) are, however, generally accepted as positions for clitic object pronouns in general.

There is already some evidence that placement of Old English PPOVs is sensitive to pronoun case. Morgan's (2004) study of the placement of verbal objects includes a detailed analysis of PPOV placement in position C and distinguishes PPOV case (dative *vs*. accusative) as well as subject form. Her results are summarised in Table 4.2.[53]

Table 4.2 Frequency of PPOV placement in position C (Morgan 2004)

Subject form	Dative PPOVs	Accusative PPOVs
Demonstrative pronoun	81% (of 31)	29% (of 17)
Nominal	67% (of 773)	39% (of 831)
man	77% (of 118)	86% (of 362)
Total	69% (of 922)	53% (of 1,210)

Morgan's data show that dative PPOVs occur in position C significantly more frequently in comparison to accusative PPOVs when the subject is a demonstrative pronoun ($\chi^2 = 12.3$, $p < 0.001$) or full NP ($\chi^2 = 128.03$, $p < 0.001$).

---

[53] Data in row one of Table 4.2 comes from Morgan (2004: 123, Table 4.5), data in row two from (*ibid*: 126, Table 4.7) and data in row three from (*ibid*: 137, Table 4.12).

When the subject is the indefinite pronoun *man*, the situation is reversed ($\chi^2 = 5.43$, $p = 0.02$). Overall, however, dative PPOVs occur more often in this particular clitic position.

In order to verify Morgan's findings about the effect of pronoun case on PPOV placement in clitic position C, I extracted from the YCOE all *that*-clauses with: (i) a case-unambiguous third person PPOV; (ii) an overt subject of any form other than a simple personal pronoun; and (iii) either [COMP PPOV Subject …] word order, e.g. (70), in which the PPOV is unambiguously in position C, or [COMP Subject (…) PPOV (...)] word order, e.g. (71), in which the PPOV is unambiguously <u>not</u> in position C.[54]

(70)   … þæt him   hiera Godas gehulpan
            that them their Gods   helped
        '… that their Gods helped them'

                                                    (coorosiu,Or_5:2.115.14.2409)

(71)   … þæt se   deað him genealæhte
            that the death him approached
        '… that death approached him'

                                                    (cocathom2,ÆCHom_II,_6:58.184.1173)

Clauses with a personal pronoun subject are ignored on the basis that personal pronoun objects invariably follow a personal pronoun subject unless the object pronoun is topicalised (see the discussion in Chapter 2). I also ignore clauses in which the PPOV is a first or second person form due to their dative/accusative ambiguity. Of course, the availability of position C is not confined to subordinate clauses introduced by *þæt*. However, as I am interested here in the effect of pronoun case on, rather than the overall frequency of, PPOV placement in position C, I assume data from *that*-clauses to provide a suitable sample.

As Morgan's results differ according to whether the subject is *man*, I make the same distinction in my analyses. The results for the 478 third person PPOVs occurring in a qualifying subordinate clause are given in Table 4.3 for those co-

---

[54] The search queries are given at Appendix C to show exactly how my data were derived.

occurring with a demonstrative pronoun or nominal subject and in Table 4.4 for those co-occurring with *man*. Percentages in parentheses indicate the proportion of the row total.

Table 4.3 Frequency of third p. PPOV placement in pos. C (subject = demonstrative/NP)

	COMP – PPOV – Subject	COMP – Subject – (…) PPOV	Total
Dative	44 (37%)	76 (63%)	120
Accusative	31 (17%)	151 (83%)	182
Genitive	1 (8%)	11 (92%)	12
Total	76 (24%)	238 (76%)	314

Table 4.4 Frequency of third p. PPOV placement in pos. C (subject = *man*)

	COMP – PPOV – *man*	COMP – *man* – (…) PPOV	Total
Dative	30 (77%)	9 (23%)	39
Accusative	108 (88%)	15 (12%)	123
Genitive	1 (50%)	1 (50%)	2
Total	139 (85%)	25 (15%)	164

The results for my sample of dative and accusative PPOVs are broadly in line with Morgan's finding. Dative PPOVs occur in position C significantly more frequently in comparison to accusative PPOVs when the subject is a demonstrative pronoun or full NP (χ^2 = 14.93, p < 0.001). When the subject is *man*, dative PPOVs occur in position C <u>less</u> frequently in comparison to accusative PPOVs, although the difference by case is not significant in my (smaller) sample (χ^2 = 2.78, p = 0.10). Also like Morgan's results, the results in Table 4.3 and Table 4.4 indicate that placement in position C is significantly more frequent when the subject is *man* than when the subject is a demonstrative pronoun or full NP, whether the PPOV is dative (χ^2 = 19.17, p < 0.001) or accusative (χ^2 = 148.21, p < 0.001). There are too few genitive PPOVs in Table 4.4 to draw any firm conclusions, but when the subject is nominal or a demonstrative pronoun, genitive PPOVs, like accusative PPOVs, appear in position C much less frequently than dative PPOVs.

In order to investigate the effect of pronoun case on PPOV placement in position B, I separately extracted all main clauses with: (i) *þa*, *þonne* or a (positive or negative) finite verb in first position (ignoring sentential conjunction); (ii) a case-unambiguous third person PPOV; (iii) an overt subject of any form other than a simple personal pronoun; and (iv) either [(*þa*/*þonne*) V$_{FIN}$ PPOV Subject (…)] word order, e.g. (72), in which the PPOV is unambiguously in position B, or [(*þa*/*þonne*) V$_{FIN}$ Subject (…) PPOV (...)] word order as in (73), in which the PPOV is unambiguously <u>not</u> in position B.[55] I assume that the delimited sample is suitably representative of clitic placement in Position B in main clauses in general.

(72) Þa   andwyrde hire se halga mid twylicere   spræce
      then answered her the holy   with ambiguous statement
      'Then the holy one answered her with an ambiguous statement'

                                    (cocathom2,ÆCHom_II,_10:87.217.1770)

(73) Ða   nolde   se hælend him þæs forwyrnan
      then not-wished the saviour him this deny
      'Then the saviour did not wish to deny him this'

                                    (coaelhom,ÆHom_28:22.4022)

The results for the 478 third person PPOVs occurring in the main clause sample are given in Table 4.5 for those co-occurring with a demonstrative pronoun or nominal subject and in Table 4.6 for those co-occurring with *man*. Again, percentages in parentheses indicate the proportion of the row total.

---

[55] The search queries are given at Appendix D to show exactly how my data were derived.

Table 4.5 Frequency of third p. PPOV placement in pos. B (subject = demonstrative/NP)

	($\flat a/\flat onne$ –) V_{FIN} – PPOV – Subject	($\flat a/\flat onne$ –) V_{FIN} – Subject (…) PPOV	Total
Dative	127 (55%)	102 (45%)	229
Accusative	89 (33%)	177 (67%)	266
Genitive	1 (33%)	2 (67%)	3
Total	217 (44%)	281 (56%)	498

Table 4.6 Frequency of third p. PPOV placement in pos. B (subject = *man*)

	($\flat a/\flat onne$ –) V_{FIN} – PPOV – *man*	($\flat a/\flat onne$ –) V_{FIN} – *man* (…) PPOV	Total
Dative	14 (78%)	4 (22%)	18
Accusative	31 (79%)	8 (21%)	39
Genitive	—	—	—
Total	45 (79%)	12 (21%)	57

Data on placement in position B are broadly in line with data on placement in position C, i.e. dative PPOVs occur in position B significantly more frequently in comparison to accusative PPOVs when the subject is a demonstrative pronoun or a full NP ($\chi^2 = 24.22$, $p < 0.001$), but not when the subject is *man*. Placement in position B, again like placement in position C, is more frequent when the subject is *man* than when the subject is a demonstrative pronoun or full NP, whether the PPOV is dative ($\chi^2 = 3.39$, $p = 0.07$, not significant) or accusative ($\chi^2 = 30.2$, $p < 0.001$). It is not possible to draw any conclusions about genitive PPOVs.

As these analyses of PPOV placement are not controlled for the possible effects of other variables, the results must be treated with some caution. In addition, some of the pronouns counted as position B clitics may be amenable to an alternative analysis where the subject is (a) phonologically heavy and (b) the only constituent to follow the PPOV, as in (74).

(74) Ða andwyrde **him** an his   leorningcnihta Andreas, Simones broþur

then answered him one of-his disciples   Andrew   Simon's brother

Petres

Peter's

'Then one of his disciples, Andrew, Simon Peter's brother, answered him'

(cowsgosp,Jn,_[WSCp]:6.8.6152)

In such examples, the position of the subject could perhaps result from extraposition (van Kemenade 1987: 39–41) — or, more accurately, 'heavy NP shift' (Pintzuk & Kroch 1985) — rather than inversion, in which case the PPOV could be anywhere between its VP-internal base position and clitic position B. As there is no precise definition of a phonologically heavy NP, I assume all examples in Table 4.6 and Table 4.5 in which the PPOV is followed by a clause-final subject involve subject-verb inversion with a position B PPOV.

In summary, the data reviewed in this section do indeed suggest that special placement of PPOVs is sensitive to pronoun case — unless the subject is *man*. *Man* appears to strongly favour special placement of PPOVs regardless of the object pronoun's case. When the subject is a demonstrative pronoun or a full NP, dative PPOVs are specially placed proportionately more frequently than non-dative special clitics. But, whereas special placement of PPOPs is rare <u>unless</u> the pronoun is dative, the same is not true for PPOVs.

4.4.6 Summary

Genitive PPOPs are invariably placed to the preposition's right. The proportion of special clitic accusative PPOPs is probably no higher than 4.5% and could be as low as 2.8%. Whatever the 'true' proportion may be for accusative PPOPs, it is clearly considerably smaller than the proportion of dative PPOPs that are special clitics (33.9%). A similar trend is evident in the poetry, but the number of accusative PPOPs is very small and the difference by case is not statistically significant.

Although Old English PPs are said to show a tendency to use accusative to denote 'motion towards' and dative to denote 'place at', this does not describe the semantics of many of the PPs for which a difference in meaning might be

expected to correlate with case. A number of minimal pairs were presented to show that PP case is not always systematic in Old English in any event. These examples further suggest that the apparent constraint against non-dative special clitics transcends whatever factor or factors mediate the choice of PP case. Data on PPOVs further indicate that 'being non-dative' seems to inhibit their special placement, although not nearly to the extent that is apparent with PPOPs and not at all when the subject is *man*. Short of assuming that non-dative PPOPs are regularly stressed — and I see no basis for such an assumption — it would appear that the near-invariable right-of-P placement of non-dative PPOPs follows directly from pronoun case. There is, however, no obvious explanation for why this effect should be almost categorical with PPOPs in the prose but not with PPOPs in the poetry or with PPOVs in the prose.

4.5   Lexis

4.5.1   THROUGH

We have seen that PPOPs governed by *þurh* 'through' are invariably situated to the preposition's right in the YCOE (see Table 3.10). According to data in Lapidge (2006: 173), the same is true for those governed by *þurh* in the concordance to *The Anglo-Saxon Poetic Records* (Bessinger 1978). As *þurh* is the only Old English preposition that almost always governs accusative (see Table 3.6), it seems reasonable to assume that this is why *þurh* is not attested with a special clitic object. It is, however, quite rare for *þurh* to occur without its object to the right in any event. Out of 96 examples in which its object is relativised, *þurh* is stranded in just seven (7%), e.g. (75).

(75)   ... for ðan þe he is se wisdom &   miht   þe se fæder ealle gesceafta
              because   he is the wisdom and power that the father all   creatures

    þurh   gesceop

    through shaped

    '... for he is the wisdom and power through which the father created all
    creatures'

<div align="right">(cocathom2,ÆCHom_I,_15:306.183.2906)</div>

No other Old English preposition shows nearly such a strong preference for pied piping except for *be* and *for*, which are always pied piped as we will see in the following section. The next closest preposition is *in* 'in', although with 25/128 (20%) examples stranded rather than pied piped, *in* stands at some distance from *þurh*.[56]

In addition, neither the YCOE nor the York Poetry Corpus provides any examples of *þurh* as the governor of an R-pronoun, i.e. elements invariably situated to the left of a governing preposition,[57] and according to the *OED*, 'therethrough' and 'herethrough' are not attested in English before *c*. 1175 and *c*. 1200 respectively. As *þær* and *her* are indeclinable, the non-attestation of, for example, **þærþurh* in Old English cannot be for reasons of case. It seems highly unlikely that the absence of *[R-pronoun(...)*þurh*] and *[PPOP(...)*þurh*] is sheer coincidence, but at present I see no obvious alternative explanation. Only one other frequently occurring preposition is not attested as the governor of *þær* or *her* in Old English. According to the *OED*, *therefrom* is first attested about 1250 and *herefrom* considerably later, in 1596, and neither the YCOE nor the York Poetry Corpus provide any earlier examples.[58] This too is rather mysterious.

While a case-based constraint would explain why pronouns governed by *þurh* are never specially placed, it would not explain why *þurh* rarely tolerates stranding and why it does not occur with either of the indeclinable locative pronouns that invariably occur to the left of a governing preposition in Old English. The existence of some general constraint against left-of-P objects of Old English *þurh* in particular does seem likely, but it is not at all obvious how this constraint could be formulated other than by stipulation.

---

[56] At the other end of the scale is *ymb(e)* 'about', which is stranded in 121 examples out of 129 (94%). No other preposition shows nearly such a strong preference for stranding in the YCOE.

[57] Lapidge (2006) does not include R-pronouns in his larger sample of Old English poetry. His comment that 'many apparent cases of postpositioned prepositions turn out to be adverbs' (*ibid*: 154) is in line with Bosworth & Toller (1898), Clark Hall (1960) and the *OED*, for example, each of which classify *þæræfter*, *þæron* and *þærto*, etc. as adverbs.

[58] Interestingly, when its semantics is locative, Dutch *van* 'from' is unable to govern *er* (van Riemsdijk 1982: 202), e.g. (i).

(i)   *Hij is ook **er**   **van**
            he is also there from
            'He is also from there'

<div align="right">(van Riemsdijk 1982: 202, ex. 46a)</div>

4.5.2 BY and FOR

Chapter 3 identified three prepositions that occur with a PPOP reasonably frequently and always with the pronoun to the right (see Table 3.10). I have concluded that the invariable right-of-P placement of objects of *þurh* 'through' could be due to the fact that this preposition almost always governs accusative, although there are other aspects of *þurh* that suggest this explanation be too simplistic. *Be* 'by, concerning', however, has been identified as a dative-favouring preposition (see Table 3.5) and *for* 'before, because of' varies in its assignment of case, so a different type of explanation must be sought for the invariable right-of-P placement of objects of these two prepositions.

The invariable right-of-P placement of objects of *be* and *for* has already been noted by Wende (1915: 14), who notes that the same is not true for *bi*(*g*) 'by, concerning' and *fore* 'before, because of'. Wende takes for granted the lexical identity of *be*, *bi* and *big* and of *for* and *fore*, but relationships between these elements are somewhat complex and, at least for *for* and *fore*, subject to some disagreement. Let us begin by considering what has been said about forms of 'by'.

According to the *Oxford English Dictionary* (*OED*) (1989: *by*), Old English *bi* and *be* go back to Old Teutonic *bi*. The relationship between 'by' forms in Old English are described in Campbell (1959) as follows: *bi* and *be* are accented and unaccented forms of the same element (*ibid*: §125), and *big* is a spelling variant of *bi* (*ibid*: §271). The prosodic alternation between *be* and *bi* shows up most clearly in compounds (Campbell *ibid*: §§71–4). Old English inherited from Proto-Germanic the assignment of stress to the first syllable of simple words and most compound words. The main exceptions, common to West Germanic languages, are compound verbs with a prepositional adverb as prefix, which have stress on the first syllable of the verb stem, cf. *began* 'to go over, traverse (to by-go)', with stress on *gan*, and *bicwede* 'proverb (by-word)', with stress on *bi*. The same prosodic analysis of *bi*(-) and *be*(-) is given in the *OED* (1989: *by*). Clark Hall (1960) also treats prepositional *bi* and *be* as lexically equivalent and *big* as a spelling variant of *bi*, although prosodic conditioning is not mentioned. Bosworth & Toller (1898) lists *be* as an 'abbreviated' form of *bi*(*g*), but the type of abbreviation is not specified. In the index of words and phrases, Mitchell (1985) lists '*be*/*bi*' and treats them as lexical equivalents

throughout (e.g. *ibid*: §§810–1, 926, 1183–4), as does the *DOE* (Cameron *et al*. 2007) for *be*, *bi* and *big*. In sum, the literature suggests the following: (i) prepositional *be*, *bi* and *big* represent the same preposition; (ii) *be* and *bi* are prosodically conditioned variants of this preposition; and (iii) *big* is a spelling variant of *bi*. Let us see if this is consistent with the data. I will henceforth refer to the assumed prepositional lexeme as BY.

Firstly, I associated *be* with <be> and *bi* with <bi~bie~big~bii~by>. There is only one other spelling of BY in the YCOE, i.e. <beo> (x62), which the *DOE* alone lists as a spelling variant of BY. Lacking independent evidence to show whether <beo> should be classified as a *bi* variant or a *be* variant, I focus on *be* and *bi* for the moment. The numbers and positions of simple personal pronoun objects of *be* and *bi* are given in Table 4.7. These figures exclude accusative pronouns, as well as those associated with other factors discussed in this chapter.

Table 4.7 Distribution of PPOPs by form of BY

	Left-of-P	Right-of-P	Total
be	—	271 (100%)	271
bi	5 (38%)	8 (62%)	13

These data indicate a syntactic difference between PPOPs governed by *be* and those governed *bi*: only the latter occur in a left-of-P position. The invariable right-of-P placement of objects of *be* does, however, make sense if we assume, firstly, that *be* is a phonologically deficient form, as Campbell (1959) and the *OED* suggest, and secondly, that *be* is phonologically dependent on its following object. The latter assumption has some independent support. *Be* is not consistently adjacent to any particular constituent other than its object, e.g. (76), and, moreover, need not be adjacent to anything else, e.g. (77).

(76) a. Se witega   cwæð **be**   **him** þæt ...

   the prophet said   concerning them that

   'The prophet said of them that ...'

<div align="right">(coaelhom,ÆHom_7:98.1108)</div>

b. Þa   cwæð se   Hælend **be**   **hyre** þæt ...

then said   the Saviour concerning her   that

'Then the Saviour said of her that ...'

<div align="right">(coaelhom,ÆHom_6:311.1022)</div>

(77)   **Be**   **hire** is awrytan þæt ...

concerning her is written that

'Of her it is written that ...'

<div align="right">(coaelive,ÆLS_[Æthelthryth]:41.4166)</div>

The dependency of *be* on a following object is also apparent in two other contexts. Although numbers of examples are small, the pattern is consistent. Firstly, in each of the four examples in the YCOE in which BY governs an R-pronoun, the preposition is realised as *bi* (<big> x4) rather than *be*, e.g. (78).

(78)   &   se   king **þærbig** sæt hleowwinde hine beo þan fyre

and the king thereby sat warming   him by the fire

'and the king sat thereby, warming himself by the fire'

<div align="right">(coneot,LS_28_[Neot]:128.118)</div>

Secondly, when stranded in a relative clause, BY is invariably realised as *bi*. There are 20 such examples (<big> x15, <bi> x5), e.g. (79).

(79)   ... up to þæm cnolle, þe   ic ær   **big**   sægde

up to the   knoll   that I   previously concerning spoke

'… up to the knoll, which I previously spoke about'

<div align="right">(coblick,LS_25_[MichaelMor[BlHom_17]]:197.20.2528)</div>

Be occurs frequently with a relativised object, but always with pied piping (N=91). There are no examples of BY with any type of left-of-P object in the York Poetry Corpus (Pintzuk & Plug 2001). According to Lapidge (2006: 155), however, there are two examples in the larger concordance to *The Anglo-Saxon Poetic Records* (Bessinger 1978), and in both instances the form is <big>.

Table 4.7 shows some variation between *be* and *bi* with right-of-P PPOPs, but *be* is clearly the preferred variant when the object is to the right, occurring in 97% (271/279) of examples. This is also the case with nominal objects: 4,149 are governed by *be* in the YCOE compared to 183 governed by *bi*, a ratio of about 23:1.[59] Unlike *be*, however, *bi* is clearly not phonologically dependent on a following object, e.g. (78) and (79). The examples at (80) further show that *bi* does not require to be adjacent to any particular type of constituent, which is consistent with an analysis of *bi* as the prosodically strong, i.e. phonologically independent, form of BY.

(80) a. þa   stodan **him** twegen weras **big** on hwitum hræglum

      then stood him two   men by in white   vestments

      'Then stood by him two men in white vestments'

                              (coblick,HomS_46_[BlHom_11]:121.99.1531)

   b. þa   stod **hyre big** iong   man fæger mid gyldenum hræglum gegyred

      then stood her   by young man fair   with golden   vestments adorned

      'Then stood by her a young, fair man, adorned with golden vestments'

                              (comart3,Mart_5_[Kotzor]:Au2,B.8.1358)

It therefore appears that *be* and *bi* are phonologically conditioned allomorphs of BY. In contexts where the preposition's object is not available for the preposition to 'lean on', the spellings consistently represent the strong allomorph {bi}. In contexts where the preposition's object <u>is</u> available for the preposition to 'lean on', spellings represent the weak allomorph {be} much more frequently than its strong counterpart.

   Two examples of *bi* with a right-of-P object are given at (81). In neither case is there reason to think the (emboldened) preposition is necessarily accented.

---

[59] These data exclude complementiser examples such as *be þæm þe* 'because'.

(81) a.   &   cwæð, þæt he þær eac gesawe Caiphan þone ealdorman þara

         and said   that he there also saw   Caiphan the   chief   of-the

         sacerda mid þam oþrum, þa   þe   Drihten Crist   ofslogon, **bi him**

         priests with the others who that Lord   Christ killed   by him

         wrecendum ligum   gesealdne   beon

         avenging   flames surrendered be

         'and said that there he also saw Caiphan, the chief of the priests, with the

         others who had killed the Lord Christ, be surrendered to the avenging

         flames by him (i.e. Satan)'

                                          (cobede,Bede_5:15.442.29.4456)

    b. Forðy   wæs **bi**   **him** gecweden

       therefore was concerning him said

       'Therefore it was said of him'

                                          (cocura,CP:47.357.17.2420)

Only *Bede* (cobede) and *Cura Pastoralis* (cocura) frequently use *bi* rather than *be* with a right-of-P (pro)nominal object, i.e. where {be} is expected: 57 times out of 124 in cobede (46%) and 46 times out of 279 in cocura (16%).[60] This would indicate that, in these texts in particular, either {be} is not consistently realised as *be* or else the unaccented form is not consistently used where it could be. The base manuscript for each of these two text files dates to early Old English, but as coorosiu (*Orosius*) is also associated with an early base manuscript but uses *bi* rather then *be* with a right-of-P object in just two out of 38 examples (5%), text date does not appear to be relevant to the use of *bi* where *be* is expected. {Bi}, on the other hand, is never realised as *be* in any of the prose or poetry text files examined.

The data thus confirm that there is good reason for treating *be* as representative of the unaccented variant of BY, i.e. {be}, and for treating *bi* as representative of the accented variant, i.e. {bi}, at least when BY governs a left-of-P object. With a right-of-P object, it seems that *bi* may represent either {be} or {bi}, but this happens with any regularity only in (Anglian-influenced) *Bede* and

---

[60] *Bede* and *Cura Pastoralis* account for 75% of the 138 instances of *bi* with a right-of-P object in the YCOE.

(West Saxon) *Cura Pastoralis*. As noted earlier, <beo> occurs 62 times in the YCOE, and in each case the object is right-of-P (PPOP x3, NP x59).[61] <Beo> thus appears to behave like *be* in avoiding special clitic objects altogether, but there is insufficient data to form an opinion about how this form fits into the system proposed for *be* and *bi*: just five of the 62 <beo> examples come from a dialectally classified text file — coeluc1 (*Honorius of Autun, Elucidarium*), Kentish — so a dialectal component to the distribution of <beo> cannot be ruled out.

The relationship between *for* and *fore* is a little unclear. According to the *OED*, Old English *fore* derives from Old Teutonic *fora, and Old English *for* is probably an apocopated form of this same protoform. Bosworth & Toller (1898) lists four forms: *for* and *fore*, both 'for, on account of', and *fōr* and *fōre*, both 'before, fore'. As vowel length is not represented in Old English spellings, *for* and *fōr* are indistinguishable, as are *fore* and *fōre*. Clark Hall (1960) and the *DOE* conflate these senses and list two, rather than four, forms, i.e. *for* and *fore*, both 'before, because of'. The possibility that they are all variants of the same preposition is suggested by the *OED*'s (1989: *for*) observation that in Old English '*for* and *fore* seem to have been used indiscriminately as preps.' This comment is discussed further in Mitchell (1985: §1185), who mentions two relevant studies. The first is that of Wülfing (1901), who found no difference in the meaning or use of prepositional *for* and *fore* in Alfredian texts. The second is that of Belden (1897), who found evidence of the lexical identity of these forms in non-Alfredian texts and who additionally notes that left-of-P placement occurs only with *fore*. Mitchell (*ibid*) apparently takes no issue with the *OED*'s treatment of prepositional *for* and *fore* as one preposition. However, on the basis of their different behaviour with respect to object placement, he concludes that 'there is no doubt that the word 'indiscriminately' in the *OED* observation [...] should be dropped, for OE at least.' Campbell (1959: §73) alone provides a prosodic analysis, but only for prefixal *for-*, claiming that its accented and unaccented forms are identical. Campbell (1959) does not mention how *for*(-) is related to *fore*(-), but Elenbaas (2006: 117), for example, treats prefixal *for-* and *fore-* as

---

[61] <beo> occurs in six text files: coalcuin (*Alcuin's De Virtutibus et Vitiis*) x22; conicodC (*The Gospel of Nichodemus*, ms. C) x11; cojames (*James the Greater*) x11; coneot (*Saint Neot*) x6; coeluc1 and coeluc 2 (both *Honorius of Autun, Elucidarium*), x5 and x1, respectively.

equivalent. In sum, the literature is divided as to whether the semantically equivalent prepositional forms *for* and *fore* should be treated as different prepositions or not. I will now argue that the relationship between *for* and *fore* is systemically equivalent to that shown to hold between *be* and *bi*. I refer to the underlying lexeme in this case as FOR(E).

There is little variation in the vowel of prepositional *for* in the YCOE: there are more than 7,000 instances of <for> compared to three of <fær>, two of <fur> and one each of <far> and <fer>. *Fore* is invariably spelled <fore>. The numbers and positions of simple personal pronoun objects of *for* and *fore* are given in Table 4.8. These figures exclude accusative pronouns, as well as those associated with other factors discussed in this chapter.

Table 4.8 Distribution of PPOPs by form of FOR(E)

	Left-of-P	Right-of-P	Total
for	—	191 (100%)	191
fore	25 (86%)	4 (14%)	29

The parallels between the data distribution in Table 4.8 and Table 4.7 are obvious: the more frequently occurring form occurs only with right-of-P objects, while the less frequently occurring form occurs with objects to the left as well as the right.

As with *be*-PPs, *for*-PPs occur in a variety of positions, e.g. (82), which is consistent with treating *for* as phonologically dependent on its object.

(82) a. & hi   ne dorston ut faran ne in faran **for him**

and they not dared   out go   nor in go   for them

'and they did not dare to go out nor in because of them'

(cootest,Josh:6.1.5288)

b. &   Abram   underfeng fela   sceatta **for hyre**

and Abraham received   much wealth for her

'and Abraham received much wealth on account of her'

(cootest,Gen:12.16.486)

c. **For ðe**, Geori, ic begeat   þisne dry

for you George I acquired this   magician

'For you, George, I have acquired this magician'

<div align="right">(coaelive,ÆLS_[George]:59.3098)</div>

Examples involving an R-pronoun or a relativised object also support a prosodically weak analysis of *for*. Although there are only two examples of FOR(E) as governor of an R-pronoun in the YCOE, in both cases the form of the preposition is *fore*, e.g. (83). Likewise, in each of the 30 examples in which FOR(E) is stranded in a relative clause, the preposition is realised as *fore*, e.g. (84).

(83)   he do   swa micel to Godes lacum   **þærfore**
         he gives so   much to God's offerings therefore
         'he shall contribute as much to God's offerings instead of that'

<div align="right">(coaelhom,ÆHom_31:103.4180)</div>

(84)   ... þæt he se man wære þe   Martinus **fore** gebæd
            that he the man was   that Martin   for   prayed
         '… that he was the man who Martin had prayed for'

<div align="right">(coaelive,ÆLS_[Martin]:231.6113</div>

For occurs frequently with a relativised object, but always with pied piping (N=178). There are six examples of FOR(E) in the same vein as those at (83) and (84) in the York Poetry Corpus (Pintzuk & Plug 2001), and Lapidge (2006: 168–9, 173) found a further two examples in a larger sample (the concordance to *The Anglo-Saxon Poetic Records*, Bessinger 1978). In all eight cases, the preposition's form is *fore*.

The examples at (83) and (84) also indicate that, like *bi*, *fore* is not phonologically dependent on a following object, although Table 4.8 shows that *fore*, again like *bi*, is found with right- as well as left-of-P PPOPs. *For* and *fore* both occur with NP objects also, although there are 45 times as many instances of *for* as there are of *fore* with a right-of-P (pro)nominal object in the YCOE.[62] Interestingly, most (70%) of the examples of *fore* + right-of-P object occur in two

---

[62] These data exclude complementiser examples such as *for þæm þe* 'because'.

texts with recognised Anglian features: cobede (*Bede*) x45 and cogregdC (*Gregory's Dialogues* ms. C) x21. Although examples of *fore* + right-of-P object also occur in West Saxon texts, only one — *Cura Pastoralis* (cocura) — provides more than one example. Regardless of dialect, all texts that provide an example of *fore* + right-of-P object use *for* more frequently with such objects, but the difference in proportions in cobede (45:178, or 20.2% *fore*) and cogregdC (21:300, or 6.5% *fore*), compared to West Saxon cocura (6:348, or 1.7% *fore*) and, for example, coaelive (*Lives of Saints*, 1:346, or 0.3% *fore*), rather suggests that *fore* + right-of-P object might be an Anglian feature.

Given their similar semantics, I think a strong case can be made for treating forms of *for* as representative of the unaccented variant of a single prepositional lexeme, FOR(E) 'before, because of', and for treating *fore* as representative of the accented variant, at least when FOR(E) governs a left-of-P object. With a right-of-P object, it seems that *fore* may represent either variant, although there is scant evidence of its use as the unaccented variant in West Saxon texts.

4.5.3   BETWEEN

Table 3.10 of Chapter 3 revealed a marked difference in frequency of left-of-P placement between PPOPs governed by *betweonum* and those governed by other forms of BETWEEN, as already noted by Wende (1915: 71, 73) and Kitson (1996: 28–32). These data are repeated in Table 4.9, from which PPOPs associated with a non-lexical knockout factor have been excluded.[63]

Table 4.9 Distribution of PPOPs by form of BETWEEN (*betweonum vs*. 'other')

	Left-of-P	Right-of-P	Total
betweonum	224 (95%)	12 (5%)	236
'between' (other)	20 (14%)	119 (86%)	139
Total	244	132	375

---

[63] Six pronouns governed by *betweonum* and ten governed by 'between' (other) are associated with a non-lexical knockout factor. Only two — both governed by 'between' (other) — are excluded for reasons of case.

Kitson (*ibid*) draws attention to another contrast, namely a tendency to avoid *betweonum* forms unless the object is a simple personal pronoun. This tendency is also evident in the YCOE as shown in Table 4.10. The data for simple PPOPs in this table is for the full sample, i.e. the data correspond to data in Table 3.10 rather than Table 4.9.

Table 4.10 Form types of BETWEEN by object type (*betweonum* vs. 'other')

	Simple PPOP	Other object type	Total
betweonum	242 (89%)	30 (11%)	272
'between' (other)	149 (19%)	624 (81%)	773
Total	391	654	1,045

Kitson (1996: 31) suggests that the strong association between *betweonum* and simple personal pronouns (as opposed to other object types) could be due 'to its origin as two words with the word governed in between, i.e. *be...tweonum*, which would tend to select for short words.' There are no examples of BETWEEN with a medial object in the YCOE, but some 47 objects are situated between the elements of other compound prepositions, e.g. *on...uppan* 'upon, on', *wið...weard* 'towards' and *on...ufan* 'upon'. Just four of these objects are simple personal pronouns, the rest are full NPs, demonstrative pronouns and modified or coordinated personal pronouns (Alcorn 2007: 57). This suggests that Old English compound prepositions do not tend to select for a short medial word. Even if Kitson were right, however, this would not explain why *betweonum* forms tend to occur with special clitic objects in particular when other forms rarely do since all forms of BETWEEN originate as two words. In this section, I argue that both tendencies are connected, but firstly I consider the various types of relationships that exist, or appear to exist, between different form types of Old English BETWEEN.

Throughout this section, I adopt the five major form types of Old English BETWEEN identified by Kitson (1993: 11–12), using bold face to indicate their status as types. The types, i.e. **betwēonum**, **betwēo(h)n**, **betwēoh**, **betwēox** and **betwēoxn**, are distinguished on etymological grounds, as we will see. Kitson's typology is not reflected in all historical English dictionaries. Bosworth & Toller

(1898) and Toller (1921) list multiple variants of BETWEEN but do not organise them according to any typology. Clark Hall (1960) identifies **betwēonum** (listed as *betwēonan*) and **betwēox** (listed as *betwux*) as well as some of their variants, but variants given for *betwux* include variants of **betwēoh** and **betwēoxn**. The *DOE* lists **betwēoh** (listed as *be-twēoh*), **betwēonum** (listed as *be-twēonan*) and **betwēox** (listed as *be-twux*), but **betwēoxn** is listed as a sub-type of *be-twux*, and variants of **betwēo(h)n** are divided between *be-twēoh* and *be-twēonan* (Kitson 1996: 28, fn. 29). The *OED* most closely mirrors Kitson's typology: **betwēonum** and **betwēo(h)n** and their variants are dealt with in the entry for *between*; **betwēoh** and its variants in the entry for *bitwih*; **betwēox** and its variants in the entry for *bitwixt*; and **betwēoxn** and its variants in the entry for *bitwixen*.

Despite differences in organisation of variants, each of these dictionaries treats (variants of) the major form types as synonymous, with senses including at least 'between', if not also 'among', '(a)mid', 'in the midst' or 'betwixt'. Clark Hall and the *DOE* also include a temporal sense, i.e. 'during' (Clark Hall does so only for *betwux*), although BETWEEN is unlikely to govern a personal pronoun with this meaning. Examples involving the two most frequently occurring form types, i.e. **betwēox** and **betwēonum**, are given in (85), where the meaning is sociative, and (86), where the meaning is locative.

(85) a. Ða cwædon hi **betwux him** þæt hi woldon wircan   ane burh
　　　　then said   they between them that they would construct a   fort
　　　　'Then said they among themselves that they would make a fort'

　　　　　　　　　　　　　　　　　　　(cocathom1,ÆCHom_I,_1:185.204.212)

　　 b. hi   cwædon **him betweonan** þæt hi   woldon bugan to þæra apostola
　　　　they said   them between   that they would bow   to the   apostles'
　　　　geferrædene
　　　　fellowship
　　　　'They said among themselves that they would bow to the fellowship of the apostles'

　　　　　　　　　　　　　　　　　　　(cocathom1,ÆCHom_I,_22:357.88.4391)

(86) a. Þa læg Petrus on ðære nihte þe Herodes wolde hine on merigen forð

     then lay Peter on the night that Herod would him on morning forth

     lædan **betwux twam cempum slapende**. mid twam racenteagum getiged.

     lead between two soldiers sleeping with two chains tied

     'Then Peter, on the night that Herod would lead him forth in the morning,

     lay sleeping between two soldiers, bound with two chains'

                                  (cocathom1,ÆCHom_II,_28:221.10.4875)

b. … þæt ða Iudei læddon Crist æt sumum sæle to anum clife, and woldon

     that the Jews led Christ at some time to a cliff and would

     hine niðer ascufan. ac he eode **betweonan heora handum**

     him down shove but he went between their hands

     '... that on one occasion the Jews led Christ to a cliff, and wished to shove

     him down, but he went between their hands'

                                (cocathom1,ÆCHom_II,_13:134.231.2966-7)

Two of the five major form types occur relatively infrequently and have a very limited distribution in the YCOE. ***Betwēo(h)n*** occurs in just three text files (cobede [*Bede*] x13, comargaC [*Saint Margaret*, ms. C] x3 and cosolilo [*St. Augustine's Soliloquies*] x3), and ***betwēoxn*** occurs in just two (cocura [*Cura Pastoralis* Hatton ms.] x16 and cocuraC [*Cura Pastoralis* Cotton Tiberius ms.] x1). I therefore postpone further discussion of these two types until later in this section.

In terms of preferred object type and in terms of PPOP placement, ***betwēox*** and ***betwēoh*** show similar behaviour. This is apparent from the data in Table 4.11, in which their objects are analysed by type, and from the data in Table 4.12, in which their simple personal pronoun objects are analysed by position. Each table gives corresponding data for ***betwēonum*** for purposes of comparison.[64]

---

[64] Table 4.11 gives numbers of simple PPOPs in all contexts, whereas Table 4.12 excludes PPOPs whose right-of-P placement can be predicted with a high degree of accuracy by reference to something other than the preposition's form, i.e. accusative PPOPs, PPOPs belonging to a coordinated or embedded PP and PPOPs occurring in a verbless clause.

Table 4.11 Form types of BETWEEN by object type (***betwēox*** *vs.* ***betwēoh***)

	Simple PPOP	Other object type	Total
betwēox	89 (18%)	397 (82%)	486
betwēoh	47 (19%)	204 (81%)	251
betwēonum	242 (89%)	30 (11%)	272
Total	378	631	1,009

Table 4.12 Distribution of PPOPs by form of BETWEEN (***betwēox*** *vs.* ***betwēoh***)

	Left-of-P	Right-of-P	Total
betwēox	5 (6%)	80 (94%)	85
betwēoh	12 (30%)	28 (70%)	40
betwēonum	224 (95%)	12 (5%)	236
Total	241	120	361

Although Table 4.12 shows that left-of-P placement is significantly more frequent with ***betwēoh*** than with ***betwēox*** ($\chi^2 = 13.46$, $p < 0.001$), right-of-P placement of simple personal pronouns is evidently the preferred option with both.

The syntactic evidence therefore suggests that ***betwēoh*** and ***betwēox*** belong to one category: both oppose ***betwēonum*** with respect to preferred object type, and both oppose ***betwēonum*** with respect to frequency of left-of-P placement of PPOPs. A likely explanation for these similarities between ***betwēoh*** and ***betwēox*** and for their differences from ***betwēonum*** is already provided by Kitson (1993: 11–16) in terms of dialect. Kitson's analysis of the distribution of the major form types of BETWEEN in Old English charter boundaries reveals: ***betwēox*** to be the exclusive form in north and west Wessex, i.e. 'the heartland of literary "West Saxon"' (Kitson 1996: 16–7); ***betwēoh*** to be mainly evidenced in the south-east, where it is largely preferred over other variants; and ***betwēonum*** to predominate in most of the West Midlands and in the middle Thames Valley, occurring sporadically elsewhere. Kitson (1993: 14) points out that the charter samples 'are not so large as to preclude the possibility that the situation in any particular area was more complicated than appears from them, but are large enough for it to be improbable that the predominance [of ***betwēox, betwēoh*** **and** ***betwēonum*** – RA] in the areas indicated is not real.'

When data belonging to each of the two main dialect categories identified in Chapter 3 are extracted from Table 4.11, we see evidence of the dialectal patterns that Kitson describes. This analysis by dialect is given in Table 4.13, in which percentages marked '↓' indicate proportions of the column's total and percentages marked '→' indicate proportions of the row's total.

Table 4.13 Form types of BETWEEN by dialect

	West Saxon	Anglian-influenced	Total
betwēox	408 (62%↓, 94%→)	25 (11%↓, 6%→)	433
betwēoh	66 (10%↓, 27%→)	175 (74%↓, 73%→)	241
betwēonum	184 (28%↓, 84%→)	36 (15%↓, 16%→)	220
Total	658	236	894

These data show that: *betwēox* predominates in West Saxon texts, where it is the majority form; *betwēoh* predominates in Anglian-influenced texts, where it is the majority form; and *betwēonum* predominates in West Saxon texts, and is the main alternative to the majority form in both dialect groups. These data suggest that *betwēoh* and *betwēox* are very probably dialectal variants of the same category of BETWEEN.

There is further support for this conclusion. Firstly, none of the West Saxon texts in the YCOE use *betwēoh* to the complete exclusion of *betwēox*. Secondly, the 408 West Saxon *betwēox* examples come from 28 text files and do not cluster in any particular subset. Thirdly, although eight of these 28 text files provide at least one example of *betwēoh*, only three use *betwēoh* more frequently than *betwēox*. One of these three text files, cogregdH ('H'), represents (a copy of) a revised version of Bishop Wærferth's translation of *Gregory's Dialogues*, a copy of which is represented in the YCOE by cogregdC ('C').[65] The language of C is generally agreed to reflect Wærferth's Anglian origins, so it is possible that the twelve instances of (Anglian) *betwēoh* compared to the single instance of

---

[65] The other two West Saxon texts that use *betwēoh* more frequently than *betwēox* are: coorosiu (*Orosius*) 19x *betwēoh vs*. 17x *betwēox*; and cobenrul (*Benedictine Rule*) 7x *betwēoh vs*. 1x *betwēox*.

betwēox in (West Saxon) H is due to the presence of *betwēoh* in H's exemplar.[66] By comparing forms of 'between' in parallel examples in the base edition of C and H (Hecht 1900–7), I found that ten of the twelve *betwēoh* PPs in H correspond directly to a *betwēoh* PP in C.[67] In one of the other two cases there is no 'between' PP in the corresponding section of C,[68] but as one of the ten examples that do correspond occurs just thirteen sentences earlier, an indirect priming effect is a possibility.[69] The twelfth example is given at (87a) along with its counterpart in C at (87b).

(87)   a. ... þæt we hwilon   ure mod geliðian &   gebigean to þam godcundum
          that we sometimes our mind soften   and bend   to the   divine

&   gastlicum rihte **betweoh þas   eorðlican carfulnysse**
and spiritual   law between these earthly   anxieties

'… that we should sometimes soften and bend our minds to the divine and spiritual law amid these earthly anxieties'

(cogregdH,GDPref_1_[H]:1.1.2)

   b. … þæt we hwilon   ure mod **betwix   þas   eorþlican ymbhigdo**
       that we sometimes our mind between these earthly   anxieties

geleoðigen &   gebigen to ðam godcundan & þam gastlican rihte
soften   and bend   to the   divine   and the spiritual law

'… that we should sometimes amid these earthly anxieties soften and bend our minds to the divine and spiritual law'

(cogregdC,GDPref_1_[C]:1.1.2)

Given H has *betwēoh* in this sentence, it is especially surprising to find West Saxon *betwēox* in the corresponding example from C. Perhaps the form in H indicates that C was not the exemplar for H, although there is no 'between' PP in the corresponding section of the other extant copy of Wærferth's translation represented in Hecht (1900) to help us out here. However, as H tends to use

---

[66] Cf. cogregdH,GD_1_[H]:11.126.6.1198 (<betwux>) and cogregdC,GD_1_[C]:11.126.7.1514 (<betweoh>).

[67] For example, cf. cogregdH,GD_1_[H]:9.65.13.633 and cogregdC,GD_1_[C]:9.65.13.732.

[68] Cf. cogregdH,GD_1_[H]:5.46.13.467 and cogregdC,GD_1_[C]:5.46.18.507.

[69] Cf. cogregdH,GD_1_[H]:5.44.29.454 and cogregdC,GD_1_[C]:5.44.30.496.

betwēoh rather than *betwēox* for no obvious reason other than priming, a priming effect seems the most likely explanation for the form of 'between' in (87a).

The distribution of ***betwēoh*** forms by text file also supports treating it as a dialectal type. The 175 ***betwēoh*** examples in Anglian-influenced texts come from eleven text files, the majority (78%) from cobede (*Bede*, x88) and cogregdC (*Gregory's Dialogues* ms. C, x48). Six of these eleven text files also provide at least one example of ***betwēox***, but ***betwēoh*** is always used at least as frequently, if not more so, than ***betwēox*** in each. Just two other Anglian-influenced texts provide evidence of ***betwēox*** but not of ***betwēoh***, but BETWEEN occurs rarely in both overall.[70]

The similarities between ***betwēox*** and ***betwēoh*** in terms of object type and PPOP placement argue for the identification of these particular form types as one syntactic category. The dialectal evidence suggests that ***betwēox*** and ***betwēoh*** are dialectal variants of this category, with ***betwēox*** being the predominant form in West Saxon texts and ***betwēoh*** the predominant form in Anglian-influenced texts. This is the analysis I henceforth adopt. For ease of reference, I identify this category as ***betwēox/h***.

The contrasting syntax of ***betwēonum*** and ***betwēox/h*** is consistent across all text files in the YCOE that provide examples of both categories. That is to say there are no text files in which full NPs are more often governed by ***betwēonum*** than by ***betwēox/h*** apart from those in which ***betwēox/h*** is unattested. Twelve text files use ***betwēonum*** but not ***betwēox/h***, but as 28 (72%) of the 39 instances of ***betwēonum*** in these twelve text files govern a simple personal pronoun object, the absence of ***betwēox/h*** is not especially remarkable. There are, however, two groups of text files in which simple personal pronouns are more often governed by ***betwēox/h*** than by ***betwēonum***. In the first group, ***betwēonum*** is unattested. 26 text files use ***betwēox/h*** but not ***betwēonum***, but as only 24 (21%) of the 117 ***betwēox/h*** PPs in these 26 text files have a simple personal pronoun object, the absence of ***betwēonum*** is again not especially remarkable. The six text files in the second group use ***betwēonum*** as well as ***betwēox/h***. In four of these text files, the tendency to use ***betwēox/h*** rather than ***betwēonum*** reflects the tendency to use

---

[70] Coalcuin (*Alcuin's De Virtutibus et Vitiis*) has ***betwēox*** x5 and ***betwēonum*** x1. Comart2 (*Martyrology*, Corpus Christi College 196) has ***betwēox*** x1.

betwēonum only when the object pronoun is a special clitic. For example, in cocathom1 (*Catholic Homilies I*) all 12 pronominal objects of ***betwēonum*** are special clitics and all 18 pronominal objects of ***betwēox/h*** are not. The only real oddities, then, are cogregdC (*Gregory's Dialogues* ms. C) and cobenrul (*Benedictine Rule*), in which ***betwēox/h*** more often governs a special clitic pronoun than a right-of-P pronoun, although in each text the number of examples is very small: 5x left-of-P *vs*. 3x right-of-P in cogregdC, and 7x left-of-P *vs*. 0x right-of-P in cobenrul. The use ***betwēonum*** in each of these texts is, however, in line with the general trend: each uses ***betwēonum*** with simple personal pronouns only — cogregdC x6, cobenrul x1 — and with special clitics in particular — 5/6 in cogregdC, 1/1 in cobenrul.

The apparent compulsion to use ***betwēonum*** only with special clitics is most strikingly illustrated by data from cowsgosp,Jn (*West Saxon Gospels, John*). This text file shows the lowest frequency of left-of-P placement (4.1%) not only in comparison to its three sister texts (*Matthew* 13.3%, *Mark* 17.8%, *Luke* 24.8%) but also in comparison to all other text files in the YCOE.[71] In fact, just thirteen (out of 316) PPOPs in *John* are in a left-of-P position, which suggests special clitics are generally avoided in this text. And of these thirteen left-of-P PPOPs, ten are governed by ***betwēonum***. As ***betwēox*** was also part of the linguistic repertoire of *John*'s scribe, it would seem the combination of ***betwēonum*** + special clitic pronoun was sometimes unavoidable.[72]

To recap: I have shown that form types of Old English BETWEEN appear to be semantically equivalent, that ***betwēo(h)n*** and ***betwēoxn*** are minor types and that ***betwēoh*** and ***betwēox*** show syntactic behaviour that is similar to each other yet systemically different from ***betwēonum***. I have also offered evidence that suggests that ***betwēoh*** and ***betwēox*** may be regarded as dialectal variants of one category, which I identify henceforth as ***betwēox/h***.

I now turn to the origins of the major types of Old English BETWEEN, as I believe this may hold a clue to the trends evident in Table 4.11 and Table 4.12. All forms go back to an original construction of the type *bi sǣm tweonum* 'by seas

---

[71] These frequencies exclude pronouns associated with non-lexical knockout factors.

[72] ***Betwēox*** occurs four times in *John*, twice with an NP and twice with a modified pronoun. For a discussion of the authorship of each book of the *Gospels*, see Liuzza (2000: 102-19).

twain' (*OED* 1989: *between*), i.e. a BY-PP with a semantically plural object which was modified by 'twain', with 'twain' showing grammatical concord with the object. For the moment, I focus on the original relationship between **betwēonum** and **betwēoh**, but I return to the other form types presently. According to Kitson (1993: 12), the 'twain' element of **betwēonum** and **betwēoh** goes back to *twīh* + the collective suffix *-n-* + case inflection. **Betwēonum** forms derive from the dat. pl. form of *by...twain*, and **betwēoh** forms from the acc. pl. form.[73] Given what we now know about differences in the placement of simple PPOPs according to case (see section 4.4), this original case distinction can hardly be ignored.

In Chapter 3, I gave an analysis of prepositions for which dative is the norm (see Table 3.5). This shows dative is the usual form of BETWEEN objects (so far as case can be distinguished), although there was something of a difference in the dative proportions for PPOPs (99.3%) as opposed to nominal objects (82.2%). When the data for nominal objects of **betwēonum** and **betwēoh** are teased apart, as in Table 4.14, evidence of a case distinction becomes apparent. In this table, the percentages in parentheses indicate the proportion of the row's total.

Table 4.14 Distribution of **betwēonum** and **betwēoh** by case (nominal objects)

	Dat.	Acc.	Total
betwēonum	27 (96%)	1 (4%)	28
betwēoh	113 (59%)	78 (41%)	191

These data lend support to the first assumption I make with respect to the story I will propose for the distribution of **betwēonum** and **betwēoh** in the YCOE: namely that as these two types became lexicalised, the former took dative case and the latter took accusative, in line with the original distribution of the *by* NP *twain* variants from which these new prepositions emerged. Although the majority (51, or 65%) of the 78 accusative objects of **betwēoh** in Table 4.14 are from cobede (*Bede*), another nine text files provide between one and five examples each, so **betwēoh** + acc. is certainly not unique to *Bede*.

---

[73] The absence of *-n* in **betwēoh** is explained by Kitson (1993: 12) as (*?)betwiohn > *betwionh > betwioh*.

My second assumption is that the constraint against non-dative special clitics was operative at the time of these form types' lexicalisation. The historical record of English does not go back far enough for this assumption to be tested, so there is no independent support for this. The left-of-P PPOPs in the examples from Old Frisian presented in Chapter 1 happen to be dative, but they hardly provide a representative sample. Nevertheless, these two assumptions together predict that originally **betwēonum** but not **betwēoh** could govern special clitic objects.

While the accusative origin of **betwēoh** could potentially provide an explanation for the tendency of this form type to occur with right-of-P objects (see Table 4.12), the proposed original grammar does not predict the tendency of **betwēonum** to occur mainly with left-of-P objects nor does it predict the very frequent occurrence of **betwēoh**+dat. by the time the manuscripts represented in the YCOE were written, especially with simple person pronoun objects.[74] My third assumption, then, is that at some point and for some reason, **betwēoh** ceased to govern accusative exclusively. If **betwēoh**+dat. started to surface, it would be impossible for learners to discern the original association between the form of BETWEEN and object case and, consequently, the association between the form of BETWEEN and the possibility of special clitic government would no longer be learnable. Given this loss of transparency in the proposed original distribution of **betwēonum** and **betwēoh**, one of two outcomes could be expected. One would be conflation of these two form types. The YCOE data speak against this option. Both types were still in use when the YCOE's material was written, and their distribution remained contrastive, as we have seen in Table 4.11 and Table 4.12. The other possibility, then, is reanalysis.

Linguistic reanalysis is typically characterised by three features as shown, for example, by Andersen (1973) for phonological reanalysis, McMahon (1994: 92–7) for morphological reanalysis and Langacker (1977), Timberlake (1977) and Lightfoot (1979) for syntactic reanalysis. Firstly, it is made possible by ambiguity in the primary linguistic data ('PLD'). Secondly, it establishes a new productive systemic principle. And thirdly, this newly established principle generates output

---

[74] **Betwēoh** occurs 42 times with a dative PPOP and twice with an accusative PPOP. Comparative figures for other (pro)nominal objects are given in Table 4.14.

that approximates the PLD. The ambiguous nature of the PLD following the emergence of **betwēoh**+dat. satisfies the first descriptive criterion, so what sort of productive systemic principle could be established on the basis of the ambiguous PLD? Assuming that the most stable data in the PLD was generated by the original grammar, the PLD would provide evidence of two entailments: (i) +special clitic → **betwēonum** and (ii) **betwēoh** → –special clitic.[75] By 'learning' these entailments as bi-directional, i.e. (i) **betwēonum** ⇄ +special clitic and (ii) **betwēoh** ⇄ –special clitic, learners would establish a new grammar for **betwēonum** and **betwēoh** that would afford each form type a distinct identity, in line with the second feature of reanalysis, whilst generating output that would be consistent with a subset of the PLD, in line with the third.

Whereas reanalysis involves the reformulation of some component of a language's grammar, 'actualisation' describes the manifestation of its consequences. According to Timberlake (1977: 168), actualisation is characterised by 'the elimination of rules or subrules in the norm that are evaluated as unmotivated with respect to the (new — RA) productive systemic principle.' Under the reanalysis I have just proposed for **betwēonum** and **betwēoh**, the 'old' rule that generated **betwēonum** PPs with anything other than special clitic objects would no longer be motivated. Over time, output norms would then be expected to increasingly resemble the situation depicted in Table 4.15, which combines data from Table 4.11 and Table 4.12.[76]

Table 4.15 Form types of BETWEEN by object type (**betwēonum** *vs*. **betwēoh**)

	+special clitic	–special clitic	Total
betwēonum	229 (84%)	43 (16%)	272
betwēoh	13 (5%)	238 (95%)	251

---

[75] As **betwēoh** with a (dat.) special clitic object would be an innovation, I assume that evidence for, e.g. *him betweoh*, would be sporadic at best.

[76] There are five more special clitic objects of **betwēonum** and one more of **betwēoh** in Table 4.15 than in Table 4.12 as Table 4.15 includes PPOPs associated with non-lexical knockout factors. There are, however, no unambiguously acc. special clitic objects of either **betwēonum** or **betwēoh**.

The small proportion of [–special clitic] objects of ***betwēonum*** and the even smaller proportion of [+special clitic] objects of ***betwēoh*** in Table 4.15 could then be seen as evidence that actualisation had not quite reached completion by the end of the Old English period.[77] Moreover, the comparatively small proportion of 'rogue' ***betwēoh*** tokens (5%) in relation to that of ***betwēonum*** (16%) is also expected under my proposal. I have argued that special placement with ***betwēonum*** goes back to the dative origins of ***betwēonum*** and that special placement with ***betwēoh*** was not possible until ***betwēoh***+dat. began to appear, and was perhaps never a stable option (see fn. 75). It is therefore unsurprising that the proportion of ***betwēoh*** examples with a [+special clitic] object is smaller than the proportion of ***betwēonum*** examples with a [-special clitic] object.

I now return to developments with the other three form types. According to Kitson (1993: 12), the 'twain' element in ***betwēo(h)n*** has the same accusative origin as that in ***betwēoh***. There are only nineteen instances of ***betwēo(h)n*** in the YCOE but ten have an accusative object, which lends weight to identifying ***betwēo(h)n*** with ***betwēoh***. There are only two instances with a simple personal pronoun object, both in *Bede*. Both pronouns are dative and, contrary to the prediction of my proposal, both are left-of-P. *Bede*, however, is unusual in using ***betwēox*** with special clitics also (see fn. 77).

Had ***betwēox*** and ***betwēoxn*** derived from an accusative original, I could simply let my arguments for the grammar of ***betwēoh*** stand for all non-***betwēonum*** forms of Old English BETWEEN, and thereby also account for the data trend in Table 4.16, which combines ***betwēox*** data from tables Table 4.11 and Table 4.12.[78]

---

[77] The thirteen ***betwēoh*** PPs with a special clitic object occur in three texts, each of which use ***betwēonum*** with a special clitic object at least once: *Benedictine Rule* (cobenrul), with ***betwēoh*** x6 and ***betwēonum*** x1; *Gregory's Dialogues* ms. C (cogregdC), with ***betwēoh*** x4 and ***betwēonum*** x5; and *Bede* (cobede), with ***betwēoh*** x3 and ***betwēonum*** x2. The use of ***betwēoh*** rather than ***betwēox*** in the West Saxon *Benedictine Rule* is itself surprising, as already noted (see fn. 65 for data).

[78] The data for ***betwēox*** is taken from tables Table 4.11 and Table 4.12. There are are five more special clitic objects of ***betwēonum*** in Table 4.16 than in Table 4.12 because Table 4.16 includes PPOPs associated with non-lexical knockout factors. There are no unambiguously acc. special clitic objects of either ***betwēonum*** or ***betwēox***.

Table 4.16 Form types of BETWEEN by object type (***betwēonum*** *vs.* ***betwēox*** *vs.* ***betwēoxn***)

	+ special clitic	– special clitic	Total
betwēonum	229 (84%)	43 (16%)	272
betwēox	5 (1%)	402 (99%)	407
betwēoxn	1 (6%)	16 (94%)	17

The origin of ***betwēoxn***, however, is firmly dative (Kitson 1993: 12, *OED* 1989: *betwixen*), and ***betwēox*** probably shares the same dative origin (*OED* 1989: *betwixt*). The *OED* records *be* prep. + **twiscu*, acc. pl. of **twisc* adj. as an alternative possibility for ***betwēox***, but there is no clear evidence of an accusative origin for ***betwēox*** in the YCOE as there was for ***betwēoh***:

Table 4.17 Distribution of ***betwēonum*** and ***betwēox*** by case (nominal objects)

	Dat.	Acc.	Total
betwēonum	27 (96%)	1 (4%)	28
betwēox	344 (95%)	17 (5%)	361

Nevertheless, it seems highly improbable that ***betwēox*** and ***betwēoxn*** should exhibit the same syntactic opposition to ***betwēonum*** as ***betwēoh*** unless ***betwēox(n)*** and ***betwēoh*** were systemically equivalent. Quite how this systemic equivalence came about, however, I cannot explain.

In summary, the story I have proposed for the distribution of forms of BETWEEN in the YCOE assumes that each of the five major form types represents a grammatically conditioned variant. I have argued that ***betwēonum*** on the one hand, and ***betwēoh*** and ***betwēo(h)n*** on the other, have distributed in a systemic fashion since their lexicalisation, and that their original distribution was determined by the case properties of the constructional variants from which they emerged. The historical record does not go far enough back to test this, but case frequencies for ***betwēonum***, ***betwēoh*** and ***betwēo(h)n*** in the YCOE show that an original case-based distribution is plausible. On the assumption that the constraint against non-dative special clitic PPOPs that is evident in the YCOE was also operative at the time the BETWEEN form types lexicalised, ***betwēonum*** would have been the only one of these three types to govern special clitics, although it

would have been free to occur with objects that were not special clitics too. The subsequent innovations of **betwēoh**+dat. and **betwēo(h)n** +dat. would have created a situation in which the original grammar was no longer learnable. I have argued that this lack of systemic transparency triggered a reanalysis of the form types' distribution, with object type replacing case as the salient factor. Data generated by this new grammar would be compatible with a subset of data generated by the old grammar, while data that led to systemic ambiguity would no longer be generated. Consequently, I predict **betwēonum** to have steadily lost its ability to occur with anything other than special clitic objects. Actualisation of the proposed reanalysis would lead naturally over time to the polarised distribution of forms of BETWEEN according to object type that is evident in the YCOE, with exceptional tokens indicating that actualisation was not yet quite complete as the Old English period came to a close. Although **betwēox** and **betwēoxn** show the same opposition to **betwēonum** as do **betwēoh** and **betwēo(h)n**, I am unable to say why. I have, however, drawn attention to a near-complementary distribution of **betwēoh** and **betwēox** according to dialect, which supports their treatment as systemically equivalent.

My final comments concern the linguistic status of the major form types of Old English BETWEEN. Do they all represent the same preposition, or do they represent more than one lexeme? We have already seen evidence of a strong association between preposition form and object position with the Old English prepositions BY and FOR(E) in the previous section, where I concluded that each is represented by a prosodically strong and weak pair of forms, but there is no reason to suppose that **betwēonum** on the one hand and other BETWEEN form types are related through prosody. Prosodically weak elements are typically (if not necessarily) monosyllabic: this is true of *be* and *for*, but not of non-**betwēonum** forms. In addition, *be* and *for* never occur with a special clitic object, which is also not true of non-**betwēonum** forms. On the other hand, there clearly exists a systemic relationship between **betwēonum** and other BETWEEN form types that cannot be reduced to either prosody or semantics. Consequently, it seems to me that the best way to characterise this relationship is to assume one prepositional lexeme, i.e. BETWEEN, with two syntactically conditioned variants, **betwēonum** being one and the other being represented by four major form types whose

distribution is partly determined by dialect. This then raises a further interesting question. The proposed characterisation of the relationship between **betwēonum** and, say, **betwēox** in Old English bears the same hallmarks as that between, for example, Present Day English *go* and *went*, except that the distribution of the PDE GO pair is conditioned by tense, i.e. [± past], whereas I have argued that the distribution of the OE BETWEEN pair is conditioned by type of object, i.e. [± special clitic]. Whether Old English BETWEEN could, or indeed should, be characterised as an inflecting preposition is a question I leave for future research.

4.6 Summary

In this chapter I have argued that the correlation between right-of-P placement and PP coordination extends only to those cases in which the coordinated PPs are headed by the same preposition and that, in those cases, right-of-P placement follows from contrastive or emphatic stress on the pronoun. I have also identified an effect that looks very like an effect of the Subjacency condition but which cannot be explained through Subjacency without losing the ability to account for the placement of the largest share of special clitic PPOPs immediately to the left of their governor. I have proposed that the number of clearly non-dative special clitic PPOPs possibly amounts to no more than twenty in the entire YCOE and have shown that this finding cannot be characterised as a lexical effect of the preposition, as an effect of PP semantics or as a consequence of parsing errors in the YCOE. The possibility that there is a more general association between pronoun placement and pronoun case in Old English is indicated by data from the poetry and by data on pronominal objects of verbs. The section on lexis tackled findings for four prepositions. I concluded that the invariable right-of-P placement of objects of *þurh* 'through' could be explained by reference to pronoun case but noted that this fails to capture two other intuitively related observations, namely the non-attestation of *þurh* as the governor of R-pronouns and its strong aversion to stranding in relative clauses. I demonstrated that the invariable right-of-P placement of objects of *be* 'by, concerning' and *for* 'before, because of' can be accounted for by recognising them as the prosodically weak counteparts of *bi* and *fore* respectively. Finally, I weaved an intricate proposal to account for the remarkable contrast in the data on Old English 'between'.

Chapter 5 Non-substantive variables

5.1 Introduction

The results of the multivariate analysis are presented in this chapter and the next. The variables dealt with in this chapter are those that do not appear to correlate in a linguistically meaningful way — if at all — with PPOP placement once other variables in the model are controlled for, namely: the pronoun-related variables NUMBER (section 5.4) and REFLEXIVITY (section 5.5); the extra-linguistic variable DATE (section 5.6); the clause-related variables VERB (section 5.7), SUBJECT FORM (section 5.8) and CLAUSE TYPE (section 5.9). Firstly, however, I provide some background to the generation (section 5.2) and interpretation (section 5.3) of the results presented in this chapter and the next.

5.2 Balancing the sample

The initial sample consisted of 9,703 PPOPs, of which 1,996 (20.6%) were excluded from the analysis of variation on the grounds of their occurrence in a (near) categorical context. The linguistic analysis of the excluded data was discussed in the previous chapter. In Chapter 3 I elected to additionally exclude the 953 first and second person PPOPs that could be dative or accusative (see Table 3.7). One quarter (N=237) of these 953 pronouns occur in at least one of the (near) categorical contexts, so the number to be excluded solely for reasons of case ambiguity is 713 (or 7.3% of the remaining sample). Just over 10% (N=74) of these of these 713 pronouns are positioned to the preposition's left. While I assume that most, if not all, of these 74 pronouns are dative, I do not include them in the analysis of variation as I am unable to identify their corresponding right-of-P dative counterparts. This leaves 6,994 PPOPs to be analysed for variation between right- and left-of-P placement.

A common symptom of non-designed data samples is imbalance, or the uneven distribution of data across the contexts defined by the analyst's model of variation. We have already seen some evidence of the limitations of an unbalanced sample in Chapter 3, when two ways of measuring diachronic change were explored. That earlier discussion focused on an imbalance revealed by the

cross-tabulation of two independent variables, but the more complex the model, the greater the number of unique contexts it defines. For example, a model that consists of one two-way variable and three three-way variables defines 54 unique contexts, since there are 2 x 3 x 3 x 3 unique combinations of the four variables' categories. Naturally, the more complex the model, the more thinly the data will be spread, and thinly spread data can potentially lead to an increased likelihood of Type II errors, i.e. the failure to detect significant effects that would be detected in a larger or more balanced sample (Hoffmann 2005a: 288, Gorman to appear: §2.2.2). An imbalanced sample is also very likely to give rise to an unreliable estimation of model fit (Sigley 1997: 246, Hoffmann 2005b: 295). Goldvarb (Sankoff, Tagliamonte & Smith 2005), a widely used tool in the analysis of linguistic variation, is well equipped to handle a certain degree of imbalance in data, but is not configured to identify the likelihood of Type II errors as part of its procedures (Sigley 1997: 248–53).[79] Goldvarb users can, however, take certain steps to maximise the reliability of the programme's independent effect estimations. In this section, I outline three fairly simple procedures I have used to improve estimations for my data.

The first procedure — which I term 'eliminate microfactors' — dissolves categories associated with a small number of observations by conflating them with a linguistically appropriate category in the same group, as recommended by Paolillo (2002: 29–30) and Tagliamonte (2006: 170–1, 200–1). For example, the univariate analysis of PREPOSITION in Chapter 3 (Table 3.10) shows the full sample includes 407 PPOPs governed by *wið* 'against' and 148 governed by *ofer* 'over'. After exclusion of data occurring in a (near) categorical context, these numbers reduce to 91 and 11 respectively, primarily due to the exclusion of data that are not, or cannot be assumed to be, dative. Having already decided that any preposition which governs a PPOP fewer than 100 times in the YCOE should be included in the miscellaneous category (see section 3.5.2.1), I have reassigned these 102 PPOPs to that category. Likewise, I have reassigned the small numbers of non-excluded PPOPs that co-occur with *biddan* 'to ask' (N=70), *feohtan* 'to fight' (N=67), *don* 'to do' (N=47) and *liefan* 'to allow' (N=15) to the miscellaneous category of MAIN VERB. There are no hard and fast rules about how

---

[79] For a synopsis of Goldvarb, see Carrera-Sabaté (2002).

many observations are 'enough' for a given model, let alone for a given factor, but my decision to recognise only those prepositions and verbs that are associated with at least 100 PPOPs is not completely arbitrary. A higher threshold of, say, 200 PPOPs would send all but seven prepositions and all but five verbs to the miscellaneous category, while a lower threshold would almost certainly not be conducive to a meaningful analysis of lexical effects.

The five other 'microfactors' are: in NUMBER, dual (N=12); in DIALECT, Anglian Mercian (N=11) and West Saxon+Kentish (N=54); and in CLAUSE TYPE, participle phrase (N=55) and small clause (N=17). The numbers given represent the number of PPOPs that do not occur in a (near) categorical context. Dual pronouns refer to more than one person and their left-of-P frequency is almost identical to that of plural PPOPs (see Table 3.8), so I have reclassified dual PPOPs as plural. PPOPs in the Anglian Mercian category show the highest frequency of left-of-P placement of all the dialect groups (see Table 3.23), but as they are so few I have simply added them to the much larger West Saxon+Anglian Mercian category. Likewise, little if anything can be concluded about the influence of Kentish on PPOP placement on the basis of 54 examples. The question is how best to reclassify these pronouns. As noted in section 3.5.4.2, results for data in the 'unclassified' and 'West Saxon+unclassified' categories cannot be interpreted from a dialectal perspective. It is thus both sensible and expedient to conflate these two categories into a single 'other' category. This new 'other' category then provides a suitable solution for the West Saxon+Kentish PPOPs. The best treatment of the small numbers of PPOPs in participle phrases and small clauses is debatable. Being non-finite constructions, there is an argument for combining these categories with the infinitival clause category. However, this would presuppose that finiteness is the right way to interpret clause type effects, and there is no basis for such an assumption. Taylor (2008: 351) solved this problem by establishing a single 'other' category for PPOPs in participle phrases and small clauses. Taylor acknowledges there is no real linguistic justification for this treatment, although she did find that the probability of left-of-P placement is much the same in these two contexts. I therefore follow suit and combine these two microfactors into a single 'other' category.

The second procedure — which I refer to as 'eliminate empty cells' — involves cross-tabulating every independent variable with every other independent variable to see whether there are any two-way combinations of categories for which there are no observations. This procedure is recommended by Tagliamonte (2006: 182) and is an easy way to identify the potential for inter-variable confounds when using Goldvarb. As the likelihood of encountering an empty cell in a cross-tabulation is reduced by the first procedure, the cross-tabulations should be done after microfactors have been eliminated. Consider the two empty cells in Table 5.1, which provides a simplified version of the cross-tabulation of LATIN INTERFERENCE with DATE and, separately, with DIALECT.

Table 5.1 Latin interference*Dialect and Latin interference*Date

	Date			Dialect		
	Late	Other	Total	West Saxon	Other	Total
Biblical transl.	1,528	—	1,528	1,528	—	1,528
Other	4,462	1,004	5,466	3,125	2,341	5,466
Total	5,990	1,004	6,994	4,653	2,341	6,994

Table 5.1 shows that all data derived from biblical translations also derive from (i) late and (ii) West Saxon manuscripts. In order to illustrate the problem this creates for estimations of the effects of the three extra-linguistic variables in my model, let us assume a model consisting of these three variables alone, i.e. DATE (late *vs.* other), DIALECT (West Saxon *vs.* other) and, for the sake of simplicity, GENRE (biblical translation *vs.* other). The contingency table defined by these variables consists of eight cells (2 x 2 x 2), as shown in Table 5.2. To this table I have added the data from Table 5.1 plus a reference for each cell.

Table 5.2 Date*Dialect*Genre

	Late		Other		Total
	West Saxon	Other	West Saxon	Other	
Biblical transl.	A 1,528 B	—	C — D	—	1,528
Other	E 2,423 F	2,039	G 702 H	302	5,466
Total	3,951	2,039	702	302	6,994

Table 5.2 makes clear that if GENRE effects were to be estimated from all data in this table, i.e. by comparing data in cells A to D with data in cells E to H, we would not be comparing like with like. While we could conclude that any difference in left-of-P frequency between data in cell A and data in cell E is not due to effects of either DATE or DIALECT, any difference between data in cell A and data in cells F, G and H could be wholly or partly due to the effects, respectively, of DIALECT, DATE or both. A similar problem would be encountered for the estimation of DATE and DIALECT effects. If DATE effects, for example, were estimated from all data in Table 5.2, i.e. by comparing 'late' data in cells A, B, E and F with 'other' data in cells C, D, G and H, we again would not be comparing like with like. While we could conclude that any difference in left-of-P frequency between data in cells E plus F on the one hand and data in cells G plus H on the other is not due to the effects of either GENRE or DIALECT, any difference between data in cell A and data in cells G and H could be wholly or partly due to the effects of GENRE and/or DIALECT.

There is, however, a way around this problem. For an independent measure of GENRE, or rather LATIN INTERFERENCE, effects, i.e. one that is free of the effects of DATE and DIALECT, we may simply ignore data in cells F, G and H. In other words, we could measure LATIN INTERFERENCE solely from late West Saxon data. Similarly, for a reliable measure of DATE and DIALECT effects, we may simply ignore data in cell A, i.e. data from biblical translations. The exclusion of particular (sets of) data from the estimation of a variable's effects is a regular procedure in variation analyses (e.g. Paolillo 2002: 61–2, Tagliamonte 2006: 180–1) and is the only way to obtain independent measures for the three extra-linguistic variables in my model. It must be recognised, however, that LATIN INTERFERENCE effects will be interpretable only for late West Saxon data and that

DATE and DIALECT effects will not be interpretable for data in biblical translations, but I believe it is better to obtain a more reliable estimation for some of the data than a less reliable estimation for all of the data.[80]

Although there are empty cells in other cross-tabulations, none has the potential for such serious consequences as those identified in Table 5.1. Nevertheless, I have also corrected for prepositions for which no first or second person PPOPs remain, i.e. *on* 'on, in', *ongean* 'towards, against' and *wið* 'against'. The lack of first or second person PPOPs for these three prepositions is due to the exclusion of first and second person PPOPs unless governed by a dative-favouring preposition (see Table 3.5). The corrections I have made to resolve these particular empty cells are partial rather than full. A full correction for the lack of first and second PPOPs governed by *on*, for example, would require the corresponding third person PPOPs to be excluded from estimations of both PERSON and PREPOSITION effects. While this would leave ample data from which the effects of PERSON could be estimated, it would leave no data for estimating the effects of government by *on*. Consequently, I have excluded third person PPOPs governed by *on*, *ongean* and *wið* from the estimation of PERSON, but not of PREPOSITION, effects. My view is that is better to have a potentially less reliable estimation of the effect of these prepositions than no estimation at all.

The only potentially problematic empty cells that are not corrected for in any way occur in the cross-tabulation of PREPOSITION and MAIN VERB. There are 156 cells in this cross-tabulation, 41 (26%) of which have zero observations. Consequently, the effect estimations for PREPOSITION may be confounded by effects of VERB, and *vice versa*.

5.3   Interpreting results

5.3.1   Factor weights

My analysis of variation in PPOP placement in Old English uses the multiple logistic regression function of Goldvarb X (Sankoff, Tagliamonte & Smith 2005)

---

[80] See Gupta (2008) for an alternative illustration and explanation—by a statistician rather than a linguist—of why empty cells, as in Table 5.2, are potentially a problem for estimations of independent effects. If the effects of GENRE, DIALECT and DATE could be shown to be wholly independent of one another there would be no need to make these adjustments, but the data are too badly distributed for this to be determinable.

to quantify the magnitude and direction of the correlation between each component factor of each variable and left-of-P placement. This information is represented by *factor weights*, whose values range between 0 and 1. A factor weight value of 0.5 indicates there is no correlation with left-of-P placement, i.e. the model estimates that left-of-P placement is no more or less probable when that factor is present than when it is absent. The value of a factor weight relative to the neutral value of 0.5 indicates the *direction* of the factor's effect. A factor weight greater than 0.5 indicates a positive correlation with left-of-P placement, i.e. the model estimates that left-of-P placement is more likely to occur when the factor is present. A value below 0.5 indicates a negative correlation with left-of-P placement, i.e. left-of-P placement is estimated to be less likely to occur when the factor is present. The further a factor weight is from the neutral value of 0.5, the stronger the correlation.

The neutral value of 0.5 is interpretable by reference to the model's *input value*. The input value (or 'corrected mean'), which also ranges between 0 and 1, represents the predicted probability of left-of-P placement regardless of the presence or absence of any of the factors included in the model. Where the data sample provides an equal number of observations for all contexts defined by the model, the input value would equal the overall frequency of the variant of interest in that sample (Johnson 2009: 360, fn. 4). Such a perfectly balanced data sample is the exception rather than the rule however, especially when the sample is undesigned. An input value of, say, 0.10 indicates that the likelihood of the linguistic variable being realised as the variable of interest is around 10%, and factor weight values express a favouring or disfavouring effect relative to that (Paolillo 2002: 34).

Goldvarb offers two options for input value and factor weight estimations: *weighted* and *centred*. The difference between these two options essentially lies in whether the neutral value of 0.5 is weighted towards the category which accounts for the larger share of the data (the weighted option) or not (the centered option), but factor weight ranges will be the same whichever option is used (Johnson 2009: §2.2). Paolillo (2002: 167–8) is somewhat sceptical of the weighted option due to its poorly documented methods. I therefore follow his advice — and that of Johnson (2009: §§2.2, 2.4) — and use the centred option.

Goldvarb also offers two options for the method of regression: *one-level* and *step-up/step-down*. I have elected to use values estimated by the second method since this procedure additionally identifies which variables correlate significantly with the dependent variable. In the *stepping-up* phase, the 'best' independent variable, i.e. the one that correlates most significantly with PPOP placement, is identified and added to the model. The data are then controlled for this variable, allowing the second 'best' variable to be identified and added to the model. This procedure is repeated sequentially until all statistically significant variables have been identified. The *stepping-down* phase works in reverse: the 'worst' independent variables, i.e. those which do not correlate significantly with the dependent variable, are sequentially deleted from the model in order of insignificance until only statistically significant variables remain.

5.3.2 Statistical significance

Of the sixteen independent variables introduced in Chapter 3, three are now redundant, namely: CASE (because all remaining PPOPs are, or are assumed to be, dative), and PP COORDINATION and PP EMBEDDING (because all remaining PPOPs belong to an uncoordinated, unembedded PP). Each of the thirteen remaining variables was found to have a statistically significant correlation with PPOP placement in Goldvarb's stepping-up and stepping-down procedures. The results are presented in full at Appendix G.

In terms of model fit, Goldvarb's comparison of the likelihood of this thirteen-variable model to the likelihood of a model that fits the data perfectly indicates that the test model is an extremely poor fit overall. Goldvarb reports the result of this comparison as a 'Fit: χ-square' statistic. As Hoffmann (2005b: 296) explains, the Fit: χ-square statistic is satisfactory if the probability (p) is ≥ 0.05. Where $p < 0.05$, the test model cannot be said to approximate the 'perfect' model. The probability for the Fit: χ-square statistic for the model used in this thesis (summarised at Appendix G) is < 0.0001. This result indicates that a significant proportion of variation in PPOP placement is unrelated to the independent effects of <u>any</u> of the thirteen variables included. There are three main reasons for a poor Fit: χ-square statistic. Firstly, some proportion of the variation may be completely unstructured, i.e. free variation may be involved to a lesser or greater extent such

that even the best possible model could never approximate the data. Secondly, the inventory of independent variables may be incomplete, i.e. there may be at least one variable — and possibly many more than one — that play an important role in conditioning the variation but which the model does not take account of. Thirdly, the model may be too complex for chi-square tests to be considered reliable (Hoffmann 2005b: 295). The problem of thinly spread data was identified earlier (see section 5.2), and one of its consequences can be high error values where the expected frequency for a cell falls below 5. As Sigley (1997: 246) points out, models of linguistic data typically have a large proportion of empty and/or poorly represented cells, as is the case here, and so expected frequencies very often fall below the minimum reliable threshold of 5. Crucially, however, a high proportion of expected frequencies < 5 does not invalidate Goldvarb's measures of independent effects: the problem lies in the potential loss of statistical power which could lead to a failure to detect significant correlations and an underestimation of the degree of model fit. As all thirteen of the variables in the present model reach statistical significance, we need not be concerned that any of these thirteen have slipped under the radar. The probable underestimation of the degree of model fit is also less of a concern in exploratory studies where the goal is to identify which of the independent variables in the inventory contribute to the variable phenomenon and which do not. Any combination of these three reasons could be responsible for the poor Fit: χ-square statistic for the model I have used.

Although each of the thirteen variables reaches statistical significance, it is a fundamental mistake to assume that statistical significance is synonymous with substantive significance, or 'strength of association' (e.g. Thompson 1999, Ziliak & McCloskey 2004, Mauder 2008: 78, Babbie 2010: 486–8). In order for a given independent variable to reach statistical significance, it is necessary only for *some non-zero* correlation to exist between that variable and the dependent variable in the particular sample at the chosen level of confidence (usually 0.05, as is the case in Goldvarb). Where the strength of association is strong, we would expect the correlation to be statistically significant. The reverse, however, does not necessarily apply: a correlation may reach statistical significance and yet be of no substantive significance.

As there are no objective tests of *substantive* significance, effect size is widely used throughout all fields of science as a good indicator of what Ziliak & McCloskey (*ibid*: 527) call the 'oomph' of a result: the larger the effect size, the larger the 'oomph', i.e. the more substantive the effect indicated. Effect sizes for each of the thirteen variables included in the model of variation in PPOP placement are quantified in the following section. In deciding whether an effect size is big enough to be taken seriously, the researcher must be guided by what is already known about or believed about the effects of the independent variable in question along with a good dose of simple common sense (e.g. Gold 1970: 177–8, Porte 2010: 199–200). In exploratory studies in particular, common sense may very well be the only available guide.

It is also important to keep in mind that the existence of a statistically significant, substantive correlation between the dependent variable — in this case, PPOP placement — and an independent variable does not mean there is necessarily a cause-and-effect relationship. This point will be illustrated towards the end of Chapter 6, where a fourteenth variable, introduced purely to tease out the combined effects of two of the thirteen variables, gives new and unexpected insight into the independent effects of another existing variable.

5.3.3 Substantive significance

Goldvarb enables estimations of effect size by reference to factor weight *ranges* (Tagliamonte 2006: 242), which is simply the difference between the highest and lowest factor weight values associated with the variable. A factor weight range indicates the strength of the variable's association with the dependent variable when all other variables in the model are controlled for: the larger the range, the stronger the correlation. Whereas factor weight *values* are interpretable by reference to the model's input value, factor weight *ranges* are not. They are, however, interpretable relative to each other, and so can be used to rank the relative importance of variables. The factor weight ranges for the thirteen variables included in the model of variation in PPOP placement in Old English are given in Table 5.3 in decreasing order of size.

Table 5.3 Ranking of variables by effect size

Ranking	Variable	Factor weight range
1	PREPOSITION	0.825
2	LATIN INTERFERENCE	0.492
3	MAIN VERB	0.489
4	CLAUSE TYPE	0.296
5	PERSON	0.257
6	LINEAR ORDER OF PP AND V	0.240
7	SUBJECT FORM	0.224
8	DIALECT	0.160
9	NARRATIVE MODE	0.156
10	ADJACENCY OF PP AND V	0.124
11	DATE	0.120
12	REFLEXIVITY	0.110
13	NUMBER	0.002

One shortcoming of inferring a variable's substantive significance from its factor weight range is that range values provide no indication of how much variation is *uniquely* associated with each variable (Hinrichs & Szmrecsanyi 2007: 463). A variable, *A*, may have a relatively large effect size, but some of the variation with which *A* is associated may also be associated with other variables. Conversely, another variable, *B*, may have a relatively small effect size, but much of the variation associated with *B* may be associated with *B* alone.

The difference between a variable's effect size and its unique explanatory power can be illustrated by reference to Figure 5.1. Let the three circles in Figure 5.1 represent the amount of variation associated with three independent variables, *A*, *B* and *C*. In Figure 5.1 the effect size of each variable (as indicated by its factor weight range) is indicated by the size of the circle — i.e. *C* > *A* > *B* — whereas the unique explanatory power of each variable is indicated by the size of the non-overlapping portion of its corresponding circle — i.e. *C* > *B* > *A*.

Figure 5.1 Illustration of effect size *vs*. unique explanatory power
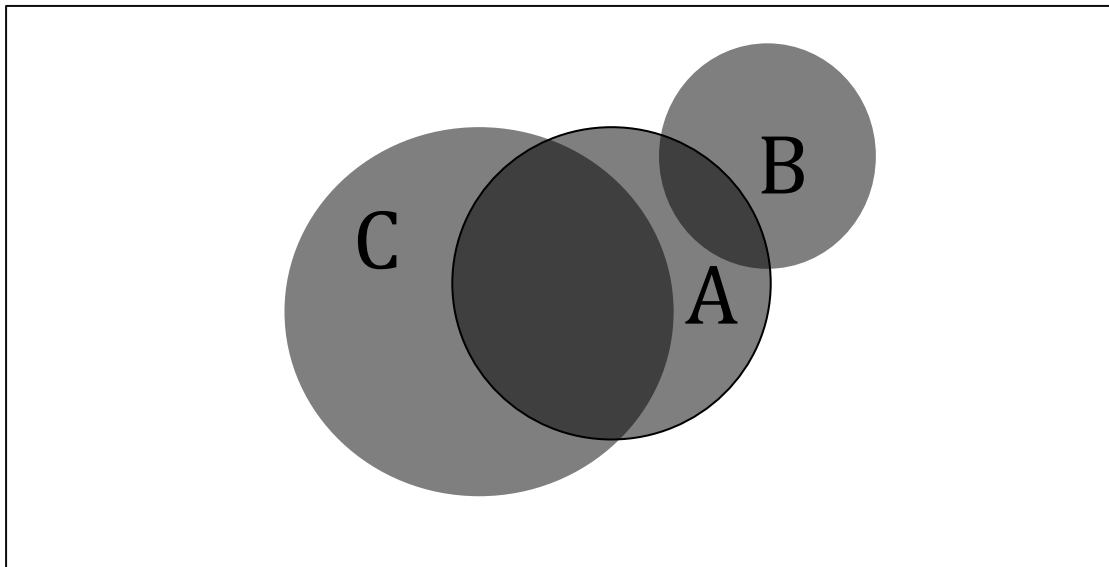
In this illustrative example, variable *C* is more important than both *A* and *B* in terms of both effect size and unique explanatory power, but the importance of variables *A* and *B* relative to each other depends on which of the measures of substantive significance is used. Both are perfectly valid ways of gauging substantive significance, but it is important to recognise that one may give a different perspective to the other.

A variable's unique explanatory power can be calculated by comparing the -2 log likelihood value (a goodness-of-fit measure in logistic regression) for the model in which all variables are present to the -2 log likelihood value for the same model from which the variable in question is excluded (Hinrichs & Szmrecsanyi 2007: 463–4). The decrease in model likelihood then provides an indication of how much of the variation accounted for by the model as a whole is uniquely attributable to the omitted variable. This information is presented in Figure 5.2. Note that the variables' values in Figure 5.2 do not have an interpretation in absolute terms: they are interpretable only in terms of their ranking and size relative to each other.

Figure 5.2 Ranking of variables by unique explanatory power
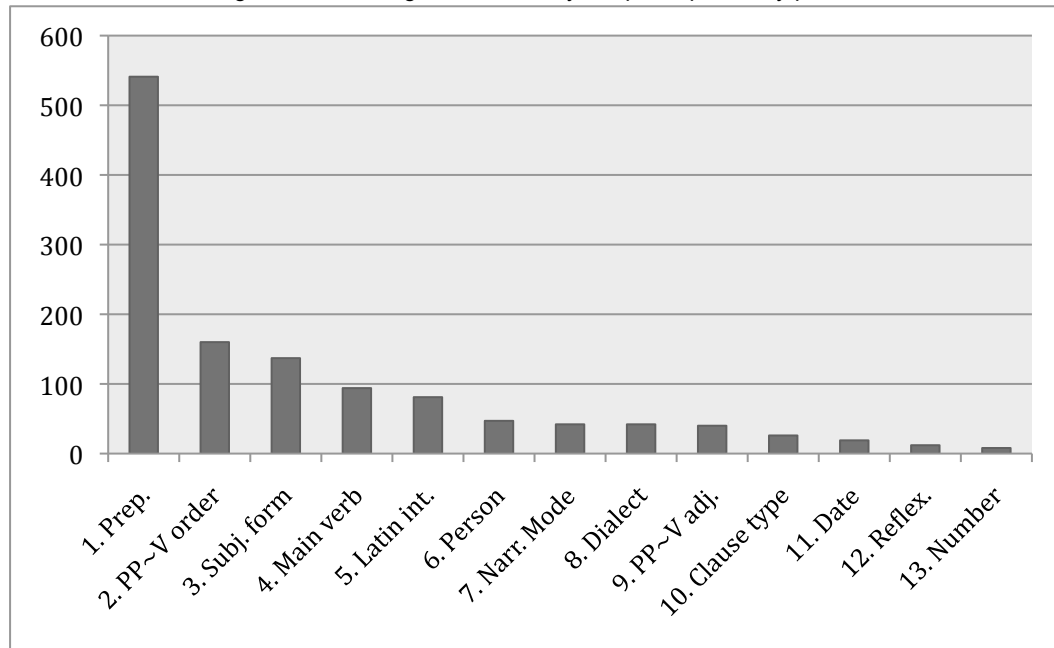
Whether the substantive and relative importance of the thirteen variables is interpreted by reference to their effect size, as in Table 5.3, or to their unique explanatory power, as in Figure 5.2, PREPOSITION ranks as the most important variable, and DATE, REFLEXIVITY and NUMBER rank as the least important. Among the remaining variables, the biggest differences between the two perspectives involves: firstly, LINEAR ORDER OF PP AND V and SUBJECT FORM, which are ranked sixth and seventh respectively in Table 5.3 but second and third respectively in Figure 5.2; and, secondly, CLAUSE TYPE which ranks fourth in Table 5.3 but tenth in Figure 5.2. This indicates that a considerable proportion of the variation with which LINEAR ORDER and SUBJECT FORM are associated is unique to those variables and, conversely, that a considerable proportion of the variation associated with CLAUSE TYPE is not unique to that variable.

5.3.4 Non-significant contrasts

As explained in section 5.3.1, the closer a factor's weight value is to 1, the greater the probability of left-of-P placement in the context defined by that factor, and as the value approaches 0, the smaller probability of left-of-P placement. The more closely two or more weight values for factors associated with the same variable approximate each other, the more probable it is that those factors do not establish

a statistically significant contrast. For example, if variable *A* is defined by three contexts, 1, 2 and 3, with weight values of 0.800, 0.300 and 0.286 respectively, the proximity of the last two values immediately suggests there may be no statistical basis for differentiating between contexts 2 and 3. This assumption can be tested by means of the likelihood ratio test (LRT). Put simply, and using the same example, the LRT uses a chi-square test to determine whether the model in which variable *A* is defined by three contexts, i.e. 1, 2 and 3 (the baseline model), is significantly better than a model in which variable *A* is defined by two contexts, i.e. 1 and not-1 (the test model) (Guy 1988: 132–3, Tagliamonte 2006: 145–51). If the test results indicate there is no significant difference between the log likelihood for each of the two models, then there is statistical justification for conflating factors 2 and 3. The identification of non-contrastive factors is, by itself, potentially just as valuable a source of information for the analysis of variation as the identification of contrastive factors. This is perhaps especially true in exploratory studies, where little, if anything, is known about what the relevant distinctions are. From a statistical point of view, the elimination of non-significant contrasts is highly desirable as it yields a more parsimonious model. Whether there is *linguistic* justification for conflating the factors is, of course, another matter. As Guy (*ibid*) is at pains to emphasise, just because two categories of a particular independent variable correlate with the dependent variable in a similar way is not sufficient reason to conflate them: unless the two categories can be shown to form a natural class, they must be kept separate, otherwise the results would be linguistically meaningless. In the remainder of this chapter and in the next, a number of non-significant contrasts are identified.

5.4 Number

I begin with the results for the least important variable in the model in terms of both effect size, as indicated by factor weight range (see Table 5.3), and unique explanatory power (see Figure 5.2), i.e. NUMBER. The results show for each category of the variable: the total number of number-unambiguous PPOPs in the analysed sample, i.e. after the exclusion of pronouns that appear in a context where right- or left-of-P placement occurs at least 95% of the time (see the summary in section 3.6); the number and proportion placed left-of-P; and a factor

weight value. The percentages in column three indicate the actual, or 'raw', frequencies of left-of-P placement, while the weight values in column four indicate the probability of left-of-P placement relative to the model's input value of 0.093 when all other variables in the model are controlled for. This format will be used throughout this chapter and the next to present the results for each variable.

Table 5.4 Factor weights for NUMBER

Factor	PPOPs	Left-of-P	Factor weight
Plural	895	172 (19.2%)	0.501
Singular	1,546	293 (19.0%)	0.499
Total / Range	2,441	465 (19.0%)	0.002

The factor weight values for NUMBER indicate that plural PPOPs very slightly favour left-of-P placement while singular PPOPs very slightly disfavour it, but the effect size is so small that I assume that there is in fact no linguistically meaningful correlation between NUMBER and PPOP placement at all.

I have one piece of independent evidence to support this conclusion. This evidence was obtained using a simple heuristic suggested by Johnson (2009: §2.2) as a way of gauging the reliability of factor weight values (or equivalent measures) when using software that does not provide estimations of their precision, as is the case with Goldvarb and with Johnson's own software, Rbrul. The viability of this technique rests on the assumption that if a statistically significant correlation in a sample as a whole reflects ordered heterogeneity — which is, after all, what we are interested in — then the correlation in question should also be evident in two random halves of the sample.

Following Johnson's suggestion, I split my sample of data into two random halves. To minimise the possibility of confounds from extra-linguistic variables, I assigned 50% of PPOPs from each text file to each group. To maximise randomness, PPOPs from each text file were alternately assigned to Group 1 or Group 2 in order of their appearance in the text, e.g. the first, third and fifth PPOPs in coaelive (*Lives of Saints*) were assigned to Group 1, and the

second, fourth and sixth to Group 2, and so on.[81] I then ran a one-step analysis on each group of data and compared the two sets of weight values in order to identify any marked differences. For all variables except NUMBER, the only differences in direction and/or ranking of factor effects involved a small number of factors, each associated with less than 100 PPOPs in both groups. As factor weight values can become increasingly unstable as the number of associated examples decreases (Johnson 2009: §2.2), these differences are not of particular concern. So, for all variables except NUMBER, the results of this exercise provide some assurance that the factor weight estimations are internally consistent. For NUMBER, however, the results conflicted: for Group 1, the factor weight for 'plural' was 0.514 (21.6% left-of-P) and for 'singular' 0.486 (18.7% left-of-P); but for Group 2 the results were in reverse, i.e. 0.486 (18.6% left-of-P) for 'plural' and 0.514 (19.1% left-of-P) for 'singular'.

Since the direction of NUMBER effects switches between two random halves of the sample, I conclude that the factor weight values for NUMBER for the sample as a whole, as shown in Table 5.4, are not internally consistent. Further, as the values consistently hover around the neutral weight value of 0.500, I take this as evidence that there is no linguistically meaningful correlation between PPOP placement and the grammatical number of the pronoun.

5.5 Reflexivity

Table 5.5 gives the results for pronoun REFLEXIVITY, the second least important variable in terms of both effect size (see Table 5.3) and unique explanatory power (see Figure 5.2).

---

[81] This was very easy to do. The token file lists the coding string for each PPOP on a separate line, arranged by text file and listed in order of appearance in each text file. The two groups were created by adding an extra column to the coding strings to which the value 1 or 2 was alternately assigned. Each group was then identified by reference to its coding value in that column.

Table 5.5 Factor weights for REFLEXIVITY

Factor	PPOPs	Left-of-P	Factor weight
Reflexive	608	191 (31.4%)	0.555
Non-reflexive	6,386	2,236 (35.0%)	0.445
Total / Range	6,994	2,427 (34.7%)	0.110

The factor weight values indicate that left-of-P placement is very slightly favoured by reflexive PPOPs and very slightly disfavoured by non-reflexive PPOPs, giving a relatively small effect size. These findings mirror those reported in Taylor (2008: 364, Table c6 for *to* PPOPs; *ibid*: 352, Table 9 for other PPOPs), although the correlation was not statistically significant in her study, probably because her sample is smaller. The effect of sample size on statistical significance can be elucidated by an analogy. If a coin were tossed five times and heads came up four times, i.e. 80% of the time, we could accept this was simply down to chance, but if heads came up in 80% of 1,000 tosses, we might start to suspect that the coin was biased. In other words, the smaller the bias in the coin, the greater the number of tosses needed to detect that bias. Likewise, the weaker the correlation between a dependent and independent variable, the larger the sample required for the correlation to reach statistical significance.

There is, however, reason to doubt the practical significance of the results in Table 5.5. In the following section, I identify one group of PPOPs that are largely responsible for the statistical significance of DATE, namely those governed by *fram* 'from'. Having excluded these 457 pronouns from the sample to measure the impact on DATE effects, I found that REFLEXIVITY no longer reached statistical significance. The outlying behaviour of *fram*-PPOPs with respect to two independent variables raises the possibility that there is something unique about this set of pronouns, although if there is, I cannot see it. *Fram*-PPOPs do not cluster in any distinctive fashion, for example they do not co-occur with any particular verb or class of verbs, they do not tend to occur in a certain type of clause, they do not tend to have third rather than non-third person reference or *vice versa*, etc. They are also spread across 52 text files of varying dates, types and dialects, so their outlying behaviour is unlikely to be attributable to any particular extra-linguistic factor. Wende (1915: 74) did find one peculiarity of

fram-PPs in his sample, namely that right-of-P placement is invariable when the PP expresses the oblique agent in a passive construction, as in (1). Could this particular use of *fram*-PPs help explain the exceptional behaviour of *fram*-PPOPs?

(1) Ond he wæs gehalgad **from him** mid micelre are

  and he was consecrated from him with great reverence

  'And he was consecrated by him with great reverence'

                (cobede,Bede_3:20.244.32.2515)

Fram is not the only preposition to head the equivalent of a *by*-phrase in passives, but there is a strong tendency for *fram* or *þurh* 'through' to introduce a personal agent and for *mid* 'with' to introduce a non-personal agent or instrument (Mitchell *ibid*: §§807, 820). As PPOPs appear rarely to refer to things rather than people in Old English (Alcorn 2009: 438–41), I assume that those belonging to a *by*-phrase are most likely to be governed by *fram* or *þurh*. And since *þurh*-PPOPs are excluded from the analysis of PPOP placement on the basis that they are invariably right-of-P, I assume that most of the PPOPs belonging to a *by*-phrase in the analysed sample will be governed by *fram*. The question, then, is whether there is something distinctive about agentive *fram*-PPOPs and, if so, whether *fram*-PPOPs cease to behave as outliers when the agentive examples are excluded from the sample.

The number of potentially agentive *fram*-PPOPs in the YCOE is not large. Just 97 *fram*-PPOPs co-occur with a form of *beon*, *wesan* or (*ge*)*weorðan* and a past participle. Not all of these PPs express an oblique agent, e.g. (2), and in some cases the PP's function cannot be determined without examining the clause in context, e.g. (3).

(2) ... þæt ic beo **fram ðe** ascired

   that I am from you separated

  '... that I am separated from you'

                (coaelive,ÆLS_[Mark]:71.3251)

(3) ge eac   Uespassianus **fram** **him** sended wæs

  and moreover Vespasian {from / by} him sent was

'and moreover Vespasian was sent {from / by} him

<div align="right">(cobede,BedeHead:1.6.8.4)</div>

Because of the number of ambiguous examples, I have not examined all 97 potentially agentive *fram*-PPs to distinguish agentive examples, as in (1), from non-agentive examples, as in (2). However, it is certainly the case that there are no agentive examples among the eight *fram*-PPs with a left-of-P PPOP: in each of these cases the PP expresses SOURCE, e.g. (4).

(4)   a. ... æfter ðan þe Loth wæs totwæmed **him fram**
          after that that Loth was separated   him from'
       '… after Lot was separated from him'

<div align="right">(cootest,Gen:13.14.519)</div>

      b. ... þær   **us** bið afyrred   æghwylc yfel **fram** &   æghwylce yrmþo
            where us is   removed each   evil fram and each   misery
       '... where each evil and each misery will be removed from us'

<div align="right">(coverhom,LS_19_[PurifMaryVerc_17]:148.2220)</div>

So it does appear to be the case that right-of-P placement of *fram*-PPOPs is invariable when expressing the oblique agent in a passive construction, although the total number of examples involved is fairly small. Excluding the 97 *fram*-PPOPs in a passive from the sample has the following effect on results: there is no material change to the factor weight values shown in Table 5.5; the effect size of REFLEXIVITY reduces slightly (to 0.086); the direction of the effects remains the same; and the variable remains statistically significant, as does (manuscript) DATE. So the outlying behaviour of *fram*-PPOPs with respect to REFLEXIVITY and DATE cannot be attributed to those that belong to a *by*-phrase.

In terms of raw frequencies of left-of-P placement, *fram*-PPOPs certainly show the greatest sensitivity to REFLEXIVITY, but as only four other prepositions govern more than 10 reflexive PPOPs in the analysed sample and there is no

consistent pattern among these prepositions' data, as is evident from Table 5.6, it cannot be said that *fram*-PPOPs are bucking an otherwise regular trend.[82]

Table 5.6 Left-of-P frequency by preposition and reflexivity

Factor	Reflexive	Non-reflexive	Total
fram 'from'	35/51 (69%)	105/406 (26%)	140/457 (31%)
to 'to'	99/167 (59%)	1,235/3,036 (41%)	1,334/3,203 (42%)
on 'in, on'	10/56 (18%)	176/401 (44%)	186/457 (41%)
betwux 'between'	15/87 (17%)	2/40 (5%)	17/127 (13%)
mid 'with'	7/170 (4%)	119/1,082 (11%)	126/1,252 (10%)
Other prepositions	25/77 (32%)	599/1,421 (42%)	624/1,498 (42%)
Total	191/608 (31%)	2,236/6,386 (35%)	2,427/6,994 (35%)

A correlation between placement of personal pronoun objects and pronoun reflexivity is identified by Hopper (1975: 37–38), who claims that 'the rule which 'clusters' pronouns at the beginning of the clause fails to apply to just those reflexive pronouns which might be confused with non-reflexives.' The potentially ambiguous pronouns in question are, of course, third person forms: first and second person pronouns are either unambiguously reflexive (when they co-refer with the subject) or unambiguously non-reflexive (when they do not). Although Hopper's claim extends only to pronouns governed a verb, the suggestion that reflexivity influences the placement of personal pronouns is certainly in line with the behaviour of PPOPs.

If reflexive third person PPOVs tend to remain close to the verbal complex, as Hopper (*ibid*: 38) suggests, then we would expect to find few, if any, among the specially placed PPOVs identified in Chapter 4. This is indeed what we find. Table 5.7 shows that third person PPOVs appear in clitic position B significantly less often when reflexive than when non-reflexive ($\chi^2 = 5.32$, $p =$

---

[82] 15 of the 35 reflexive left-of-P PPOPs governed by *fram* co-occur with *aweorpan* 'to throw, throw away, cast down', e.g. (i). Old English dictionaries lists *framaweorpan* 'to cast from, throw away' but none of the examples they cite show *framaweorpan* used with a reflexive object. Consequently, I have no independent evidence to support interpreting these 15 examples this way.

(i) and bædon þæt hi   awurpon heora wæpna **him fram**
     and bid   that they cast   their weapons them from
     'and bid that they should cast their weapons from them'

(coaelive,ÆLS_[Maurice]:59.5720)

0.02) and Table 5.8 shows the correlation is even stronger with respect to clitic position C ($\chi^2 = 28.35, p < 0.001$).

Table 5.7 Frequency of third p. PPOV placement in pos. B by pronoun reflexivity

	($\flat a/\flat onne$ –) V_{FIN} – PPOV – Subject	($\flat a/\flat onne$ –) V_{FIN} – Subject (…) PPOV	Total
Reflexive	15 (31%)	33 (69%)	48
Non-reflexive	247 (49%)	259 (51%)	506
Total	262 (47%)	292 (53%)	554

Table 5.8 Frequency of third p. PPOV placement in pos. C by pronoun reflexivity

	COMP – PPOV – Subject	COMP – Subject (…) PPOV	Total
Reflexive	5 (10%)	45 (90%)	50
Non-reflexive	210 (50%)	213 (50%)	423
Total	215 (45%)	258 (55%)	473

Hopper (*ibid*: 38) argues that this difference in the syntax of reflexive and non-reflexive third person pronouns 'is a consequence of the distinction between two 'different' morphemes which happen to coincide in phonological shape', and speculates that this syntactic difference might explain why the morphological distinction between reflexive and non-reflexive forms that was present in proto-Germanic became redundant in Old English. However, when we take into account the effects of PPOV case and subject form (i.e. *man vs.* nominal) on PPOV placement in clitic positions B and C, as shown in section 4.4.5, it transpires that the proportion of reflexive and non-reflexive PPOVs placed in these positions is very close to what we would expect. In other words, there does not appear to be an effect of pronoun reflexivity on special placement of PPOVs over and above the apparent effect of pronoun case and subject form.

Overall, I conclude that there is no convincing evidence of a systemic relationship between PPOP placement and REFLEXIVITY. The effect size for the sample as a whole is small, and the statistical significance of the variable can be localised to *fram*-PPOPs, which appear to behave somewhat exceptionally also with respect to DATE albeit for reasons I am unable to fathom. Data for PPOVs

appear to lend some weight to Hopper's (1975: 38) claim that there is a tendency to avoid placing reflexive third person PPOVs early in the clause, but this can be wholly attributed to factors identified in the previous chapter as correlating with placement of third person PPOVs in clitic positions B and C, namely pronoun case and subject form.

5.6 Date

The results for the variable DATE are given in Table 5.9. Recall that the weight values for DATE have been estimated without reference to data in biblical translations as all such data in the YCOE come from late manuscripts, so the results are for PPOPs in non-biblical translations and non-translated texts only. Data in the 'unclassified' category were also excluded from the estimate of DATE effects on the basis that at least some examples properly belong to the early and/or late category.

Table 5.9 Factor weights for DATE

Factor	PPOPs	Left-of-P	Factor weight
Late (post-AD 975)	4,526	1,868 (41.3%)	0.560
Early (pre-AD 925)	805	291 (36.1%)	0.440
Total / Range	5,331	2,159 (40.5%)	0.120

The factor weights indicate that left-of-P placement is slightly favoured in late Old English manuscripts and slightly disfavoured in early ones. This is surprising given that left-of-P placement ceased to be an option at some point during the Middle English period (van Kemenade 1987: 190). Taylor (2008), who measured the diachronic trend by reference to text composition date (personal correspondence), found broadly the same result and localised the unexpected trend to PPOPs governed by *to* 'to' (*ibid*: 346–7, 364 Table c9). PPOPs governed by other prepositions in her sample showed no significant diachronic change in left-of-P frequency. Although unable to explain this aspect of *to*-PPOPs, Taylor notes that their increasing frequency in left-of-P placement over time is manifest in both translated and non-translated texts, which led her to conclude that the trend probably cannot be attributed to an oddity of any particular scribe or author.

In my sample, two prepositions show a statistically significant increase in frequency of left-of-P placement over time according to raw frequencies, as shown in Table 5.10. Taylor (*ibid*: 357, fn. 5) found a similar trend for *fram*-PPOPs but the increase was not statistically significant in her sample.

Table 5.10 Left-of-P placement of PPOPs governed by *to* and *fram* by DATE

Factor	*to*	*fram*
Late (post-AD 975)	1,103/2,102 (52%)	112/300 (37%)
Early (pre-AD 925)	111/270 (41%)	8/71 (11%)
Total	1,214/2,372 (51%)	120/371 (32%)

As there is a reasonably large number of *to*-PPOPs, it is possible to model variation in their placement alone using the same model as that used for all PPOPs. When the placement of these 2,372 PPOPs is analysed separately, the factor weight range for DATE turns out to be smaller (Late 0.526, Early 0.474, range 0.052) in comparison to that for the sample as a whole, and DATE is not identified as statistically significant in either the stepping-up or stepping-down procedure. This indicates there is no significant difference in likelihood of left-of-P placement according to manuscript date for *to*-PPOPs when the other variables in the model are controlled for. The 11% increase in left-of-P placement of *to*-PPOPs over time shown in Table 5.10 must therefore be due to factors other than manuscript date. There are too few *fram* PPs to model their placement separately, but when they alone are <u>omitted</u> from the full sample, the factor weight range for DATE reduces by half (Late 0.529, Early 0.471, range 0.058), although DATE remains statistically significant ($p = 0.001$).[83] This indicates that half of the effect size of DATE in my sample as a whole can be attributed to variation exhibited by *fram*-PPOPs in particular. As there are so few early *fram*-PPOPs, I suggest nothing meaningful should be read into the fact that PPOPs governed by *fram* behave differently in late manuscripts than in early manuscripts. And although there remains a statistically significant increase in left-of-P placement over time for PPOPs other than those governed by *fram* when

---

[83] All variables remained statistically significant when *fram*-PPOPs are excluded from the sample except for REFLEXIVITY, as discussed in the previous section.

other variables in the model are controlled for, I suggest that effect size — 0.058 — is sufficiently small to be treated as negligible. In sum, I find there is no convincing evidence of a statistically significant increase or decrease in left-of-P placement according to manuscript date.

In order to investigate Taylor's findings that there is a significant increase in left-of-P placement for PPOPs governed by *to* according to text composition date, I replaced manuscript date with text composition date as the means of measuring diachronic change, leaving all other aspects of my model exactly the same. The results are in line with those of Taylor (2008), i.e. *to* is the only preposition whose objects show a statistically significant difference in placement according to text composition date, with left-of-P placement favoured in texts with a late composition date (factor weight 0.565) and disfavoured in those with an early date of composition (factor weight 0.435). The effect size for my sample (0.130) is, however, considerably smaller than that for Taylor's sample (0.342) (Taylor 2008: 364, Table c9), for which there are several possible explanations. Firstly, there are a larger number of independent variables in my model in comparison to Taylor's, so some of the variation associated with text composition date in Taylor (2008) may have been absorbed by one or more of my additional variables. Secondly, there is a difference in our respective samples and, thirdly, there is a difference in the proportions of data controlled for Latin interference effects. In addition, whereas I have excluded certain groups of data from my estimations of DATE and LATIN INTERFERENCE (see section 5.2), Taylor did not.[84] Any or all of these factors may account for the difference in the effect size of text composition date for placement of PPOPs governed by *to*. As my data sample and statistical model are more comprehensive than in Taylor (2008), I would suggest my findings are the more reliable.

Is there a difference in PPOP placement between early and late Old English? The only evidence that there might be concerns pronouns governed by two prepositions — *fram* and *to* — for which there are indications of an

---

[84] Not only do all biblical translations in the YCOE come from late manuscripts, they are all associated with a late composition date. Accordingly, the problem discussed in section 5.2 and illustrated in Table 5.2 is present whether diachronic trends are measured by reference to text composition date or manuscript date. For this reason, data from biblical translations were excluded from the analysis of the effects of composition date (as they were when DATE was measured by reference to manuscript date) to guarantee the independence of the results for DATE and LATIN INTERFERENCE from one another.

increasing preference for left-of-P placement from the early to late period. For *fram*-PPOPs, the difference is apparent when time is measured by reference to manuscript date, but there are rather too few *fram*-PPOPs in early manuscripts for their behaviour to be judged as representative. When time is measured by reference to text composition date, no diachronic change can be detected. For *to*-PPOPs, the difference is apparent when time is measured by reference to composition date, although the effect size is hardly impressive (0.130). When time is measured by reference to manuscript date, no diachronic change can be detected. On the basis that evidence of a difference in PPOP placement between early and late Old English is limited to two prepositions for which counter-evidence is available, I conclude that there is no compelling evidence of diachronic change.

5.7 Main verb

Table 5.11 shows the frequency and probability of left-of-P placement according to the particular main verb with which the PP co-occurs.

Table 5.11 Factor weights for MAIN VERB

Factor	PPOPs	Left-of-P	Factor weight
faran 'to go'	107	32 (29.9%)	0.628
sendan 'to send'	187	74 (39.6%)	0.612
niman 'to take'	136	45 (33.1%)	0.608
cuman 'to come'	722	333 (46.1%)	0.600
cweðan 'to say'	1,241	543 (43.8%)	0.573
sprecan 'to speak'	156	51 (32.7%)	0.567
clipian 'to speak, call'	114	53 (46.5%)	0.565
gan 'to go'	157	58 (36.9%)	0.527
bringan 'to bring'	102	40 (39.2%)	0.509
miscellaneous	3,106	1,023 (33.9%)	0.484
BE 'to be'	622	125 (20.1%)	0.443
lædan 'to lead'	158	31 (19.6%)	0.340
habban 'to have'	276	19 (6.9%)	0.139
Total / Range	6,994	2,427 (34.7%)	0.489

In spite of this extensive array of factor weight values, I find there is little compelling evidence of a statistically significant correlation between PPOP placement and the particular main verb involved — with the possible exception of *habban*.

The first indication that placement of PPOPs differs minimally according to the accompanying main verb comes from the fact that most of the factor weight values are close to the neutral value of 0.5. This indicates that when all of the other variables in the model are controlled for, the probability of left-of-P placement is not hugely sensitive to the particular main verb involved unless the verb is *lædan* or *habban*. A second indication is the proximity of each factor weight value to those that are adjacent, again with the exception of that for *habban* and, to a lesser extent, *lædan*. This indicates that PPOP placement does not differ significantly between most of the verbs identified. Indeed, a series of supplementary statistical tests show that when the PPOP co-occurs with any of the first nine verbs listed in Table 5.11, i.e. *faran* to *bringan* inclusive, the probability of left-of-P placement is no more or less likely with one rather than another. These nine verbs, then, can be seen as belonging to an internally consistent set in terms of their correlation with PPOP placement. In addition, since the weight values for some of these nine verbs are very close to the neutral value of 0.5, I interpret this as evidence that the presence of any one of these nine verbs has no significant impact on PPOP placement.

Working further down the verbs listed in Table 5.11, the miscellaneous category represents all those that co-occur with fewer than 100 PPOPs in the analysed sample. Little can be inferred about the behaviour of PPOPs in the presence of any of these verbs individually, but the proximity of this category's factor weight value to the neutral value of 0.5 suggests there is no statistical correlation with PPOP placement. For PPOPs co-occurring with BE, statistical tests show left-of-P placement is significantly less likely in comparison to the first four verbs listed in Table 5.11, but no more or less likely in comparison to the next five. The factor weight value for BE is also very close to 0.5. Consequently, it seems to me that it is not unreasonable to interpret all of the results discussed so far, i.e. those for *sendan* to BE inclusive, in terms of a short and relatively uninteresting continuum of probabilities centred around the neutral value of 0.5.

The weight value for *lædan* clearly falls outside this continuum, but this is partly due to the 58 *lædan mid* examples, all of which have a right-of-P PPOP, e.g. (5).

(5)   Se ealdorman ða   þa apostolas **mid him** to ðam cyninge Xerxes gelædde

           the chief   then the apostles with him to the   king   Xerxes led

           'Then chief then led the apostles with him to Xerxes the king'

                                           (cocathom2,ÆCHom_II,_38:282.70.6347)

As the majority (80%) of all possible combinations of preposition and verb occur less than twenty times with a PPOP in the analysed sample, it is not possible to obtain reliable estimations of the combined effects of preposition and verb on PPOP placement for any particular combination. However, indications of a strong dispreference for left-of-P placement of pronouns governed by *mid* has already been identified (see Table 3.10), and *lædan* is the only verb itemised in Table 5.11 that co-occurs with a right-of-P *mid*-PPOP at least three times but never with a left-of-P *mid*-PPOP.[85] These 58 examples amount to 37% of all the data for *lædan* in Table 5.11, so are bound to have a noticeable impact on its weight value. Indeed, when these examples are excluded from the sample, the weight value for *lædan* increases from 0.340 to 0.426, i.e. to a value close to that for BE.

I therefore conclude that, with exception to *habban*, to which I return in a moment, there is no compelling evidence that PPOP placement is sensitive to the particular main verb it occurs with. There are some differences to be sure, but as each probability value in Table 5.11 does not differ significantly from (a value which is very close to) the neutral value of 0.5 — *habban* aside — these differences may be reasonably ascribed to random fluctuations in the sample.

Habban then remains as the only verb with a factor weight value that is significantly different from each of the others. Unlike the data for *lædan*, there is no sizeable cluster of examples involving *habban* where the PPOP is always right-of-P. Around half of the PPOPs that co-occur with *habban* (130/276, 47%)

---

[85] The only other P+V combination that occurs more than a handful of times but never with a left-of-P PPOP is *beon æfter* (N=23). The only combinations that occur more than a handful of times but never with a right-of-P PPOP involve *betweonum* (discussed in Chapter 4) and *ongean* and *togeanes* (discussed in Chapter 6).

are governed by *mid*, and this certainly contributes to its low factor weight: just 7 of these 130 PPOPs (5%) are left-of-P. The next most frequently occurring preposition among the *habban* data is *on* (64/276), and these PPOPs also show a very low frequency of left-of-P placement: just 4/64 (6%) are left-of-P. This is a much lower proportion than might be expected according to data in Table 3.10, which shows the overall frequency of left-of-P placement of *on*-PPOPs is about 25%. *Habban* occurs with other individual PPOP-governing prepositions only infrequently, but there is a consistent pattern of a lower-than-expected frequency of left-of-P placement. It is not impossible that this is due to the small number of examples involving prepositions other than *mid* and *on*, but other verbs suffer from similar problems without it leading to a clearly deviant factor weight value. It is not the case, however, that *habban* is associated with a low frequency of left-of-P placement of PPOPs regardless of the preposition involved: PPOPs governed by *betweonum* are not included in the analysed sample but each of the 16 examples that co-occur with *habban* is left-of-P, although this seems to be the only context in which the correlation with *habban* clearly gives way to another factor that is strongly associated with PPOP placement.

Neither the left- nor right-of-P PPOPs accompanied by *habban* cluster in any particular set of text files defined by dialect, date or genre (i.e. ± translation), nor do they congregate in any particular context that strongly favours right-of-P placement, other than those governed by *mid*. *Mid* examples apart, there is no obviously common feature to potentially explain the generally strong preference for right-of-P placement in the presence of *habban*. In a small group of examples, the verb co-occurs with a deverbal nominal to form a composite predicate (Akimoto & Brinton 1999), e.g. (6a) (cf. *andian* 'to envy') and (6b) (cf. *wunian* 'to dwell').

(6) a. &   hæfdon andan **to him**

         and had   envy   to him

         'and envied him'

                                                    (cogenesiC,Gen_[Ker]:37.8.18)

    b. ... þonne næfþ   Godes gast nane wununge   **on him**

            then   not-has God's spirit no   dwelling-place in them

'… then God's spirit will have no dwelling-place in them'

<div align="right">(coaelhom,ÆHom_4:252.656)</div>

Some other examples are given at (7) to illustrate the miscellany of PPs that co-occur with *habban*.

(7) a. Seo sunne ðe ofer us scinð   is lichamlic gesceaft. &   hæfð swa ðeah

        the sun   that over us shines is bodily   creature and has   nevertheless

        ðreo agennyssa **on hire**

        three properties in it

        'The sun that shines over us is a physical creature and has, nevertheless,

        three properties in itself'

<div align="right">(cocathom1,ÆCHom_I,_20:338.100.3934–5)</div>

   b. ... ðeah   hie geseon ðæt ða yfelan hie   hæbben **ongemong him**

        though they see   that the evil   them have   among   them

        '... though they see that the evil ones have them among them'

<div align="right">(cocura,CP:37.263.9.1712)</div>

   c. ... efne swa swa seo ea   **in hire** nænigne wætres stream hæfde

        even so   as   the river in it   not-any water's stream had

        '... even though the river had no stream of water in it'

<div align="right">(cogregdC,GD_1_[C]:2.15.30.144)</div>

In conclusion, MAIN VERB ranks as the third most important variable in the multivariate model in terms of effect size, i.e. factor weight range (see Table 5.3), and the fourth most important in terms of unique contribution to model likelihood (see Figure 5.2). The factor weight range is considerably extended by *habban*, which seems to have a generally inhibiting effect on left-of-P placement unless the pronoun is governed by *betweonum*. When *habban* data are excluded from the sample along with the 58 *lædan+mid* PPOPs, all of which are right-of-P, the factor weight range for MAIN VERB reduces by more than 60% (from 0.473 to 0.182) and all individual weight values then fall squarely within 0.1 of the neutral value of 0.5. This, I suggest, should be interpreted as evidence that PPOP placement does not vary according to the particular verb involved in any

significantly meaningful way, unless — for reasons I cannot discern — the verb is *habban*.

5.8 Subject form

The results for the variable SUBJECT FORM are given below. As explained in Chapter 3, the 'other' category represents PPOPs in clauses with neither a nominal subject nor a nominative personal pronoun subject.

Table 5.12 Factor weights for SUBJECT FORM

Factor	PPOPs	Left-of-P	Factor weight
Nominal	2,179	969 (44.5%)	0.583
Other	2,613	1,006 (38.5%)	0.561
Personal pronoun	2,202	452 (20.5%)	0.359
Total / Range	6,994	2,427 (34.7%)	0.224

The inclusion of this variable in the model of variation in PPOP placement was motivated solely on the basis of an impression that left-of-P placement occurs less frequently in the presence of another personal pronoun in the same clause. The reason for focusing on the form of the subject in particular was because PPOPs co-occur considerably more frequently with a personal pronoun subject than with another personal pronoun object. The reason for formulating SUBJECT FORM as a three-way rather than two-way variable was to help tease out whether left-of-P placement occurs less frequently in the presence of another personal pronoun in comparison to all other types of subject or only in comparison to some. The factor weight values for the nominal and 'other' categories are almost identical, and the likelihood ratio test confirms there is no significant difference in probability of left-of-P placement between these two groups of PPOPs. In most (90%) of 'other' cases, there is no overt subject. In the 10% of cases with an overt 'other' subject, the subject is either the indefinite pronoun *man* (more on which in a moment), an overt relative pronoun, a quantifier or a non-nominative subject in a non-finite clause. The probability of left-of-P placement does not differ significantly between these two sub-sets of 'other' data either.

These results indicate that there is a significant difference in likelihood of left-of-P placement between PPOPs that co-occur with a nominative personal pronoun subject, which strongly disfavour left-of-P placement, and those that occur with anything <u>but</u> a nominative personal pronoun subject, which slightly favour left-of-P placement. Having obtained this quantitative evidence for my impression about the effect of another personal pronoun on PPOP placement, at least as far as the form of the subject is concerned, I am not entirely sure how to interpret it.

On two previous occasions, I have turned to two groups of personal pronoun objects of verbs ('PPOVs') for independent evidence of an effect of a particular variable on the special placement of simple personal pronoun objects in general. Unfortunately, the same tactic cannot be used to explore whether a co-occurring personal pronoun subject correlates with a lower frequency of PPOV placement in clitic positions B and C. One of the criteria for recognising clitic object pronouns in these positions in particular is their placement immediately before the subject. While there are lots of qualifying examples in which the subject is a nominal argument or the indefinite pronoun *man*, there are <u>no</u> examples in which the subject is a simple personal pronoun.[86] While this might <u>appear</u> to be evidence of an effect of SUBJECT FORM on the special placement of PPOVs, an independent explanation for this asymmetry has already been discussed in Chapter 2, namely that simple personal pronoun objects do not precede a simple personal pronoun subject unless the object pronoun is the topic. We have, however, seen data, presented and discussed in section 4.4.5, that show placement of PPOVs in clitic positions B and C is significantly more frequent when the subject is *man* than when it is nominal. This is also true for PPOPs: 68%

---

[86] Actually the YCOE provides one example in which a PPOV separates a fronted verb from an inverted personal pronoun subject. This example, given at (ii), is also the only one found by Allen (1995: 109, fn. 14) in which a pronominal dative experiencer argument precedes a pronominal nominative theme argument (nom. > dat. is the regular order for this type of 'impersonal' construction when both arguments are pronominal). Allen (*ibid*) explains, however, that <hit> was supplied by the text's editor (Sedgefield 1899), which indicates that <hit> does not appear in the 10th- or 17th-century text witnesses. The 12th-century witness has the expected order <hit me>.

(ii)   Ne þuhte **me**   hit   no rihtlic, ne eac nauht   gerisenlic, gif him sceolden
        Not seemed me-DAT it-NOM not right   nor also not-at-all fitting   if him must
        þiowe men þenian
        servile men serve
        'It did not seem right to me, nor also at all fitting, if servile men had to serve him'
                                                                        (coboeth,Bo:41.142.4.2834)

of PPOPs that co-occur with *man* in the analysed sample are left-of-P, which is about 50% more frequent than for PPOPs that co-occur with a nominal subject, as shown in Table 5.12. However, as there are only 104 examples involving *man* in total in Table 5.12, this has little effect on the overall proportions. Once again, I see no obvious explanation for why left-of-P placement should be most frequent when the subject is the impersonal pronoun *man* and least frequent when the subject is a personal pronoun.

When the trend evident in Table 5.12 is considered on a text-by-text basis a rather confused picture emerges. In Appendix H, I show the frequency of left-of-P placement according to SUBJECT FORM for each of the eleven texts that supply at least 50 examples to each of the three SUBJECT FORM categories. Although these data are not controlled for the possible effects of other variables, the data in Table 5.12 do suggest that raw frequencies of left-of-P placement according to SUBJECT FORM are a reasonable guide to probability when other variables are controlled for. In six of these texts, i.e. *Ælfric's Catholic Homilies I* and *II, Bede, Blickling Homilies, Heptateuch* and *West Saxon Gospels*, PPOP placement conforms to the pattern suggested in Table 5.12. In *Vercelli Homilies* and *Orosius*, on the other hand, there is no significant difference in placement between PPOPs that co-occur with a personal pronoun subject and those that co-occur with a nominal subject, while in *Supplementary Homilies* and *Gregory's Dialogues* ms. C there is no significant difference between PPOPs that co-occur with a personal pronoun subject and those that co-occur with a subject belonging to the 'other' category. So in four of the eleven texts, one of the expected contrasts does not obtain. Further, in *Lives of Saints* and *Supplementary Homilies* left-of-P placement is significantly more frequent when the subject is a full NP than when it belongs to the 'other' category, while in *Vercelli Homilies* the opposite is true. So in three of the eleven texts, a contrast is manifest that is not expected. The results of this analysis for individual texts contrast starkly with that undertaken for NARRATIVE MODE, which are remarkably consistent, as we will see in the following chapter. Nevertheless, in all but *Vercelli Homilies*, left-of-P placement occurs least frequently in absolute terms when the subject is a personal pronoun.

The analysis has shown that left-of-P placement is significantly less likely in the presence of a nominative personal pronoun subject, but I can offer no linguistically plausible interpretation for this finding. Lacking any other insight, it seems best to regard the overall correlation between SUBJECT FORM and PPOP placement as a spurious or 'nonsense' correlation, a term used to describe a situation where two variables are correlated without being causally related to one another and which is often explained by reference to a third variable to which they are both related (Upton & Cook 2008). For example, statistical textbooks often cite the strong and spurious correlation between ice-cream sales and the number of deaths by drowning. This correlation, of course, makes sense by assuming a third variable, temperature: in hot weather, people consume more ice-cream and more often go swimming. From this perspective, SUBJECT FORM does not in fact belong in a model of variation in PPOP placement in Old English but could nevertheless provide a clue to some as yet unidentified and non-spurious variable that does belong in such a model. The relevant question, then, is what correlates with [± personal pronoun subject] that might also correlate with PPOP placement? I do not have an answer, but hypotheses along this line could very well yield some valuable insights.

5.9 Clause type

The factor weights for CLAUSE TYPE are given in Table 5.13. As explained in section 5.2, the 'other' category represents PPOPs occurring in participle phrases (N=55) and small clauses (N=17).

Table 5.13 Factor weights for CLAUSE TYPE

Factor	PPOPs	Left-of-P	Factor weight
Main	1,987	642 (32.3%)	0.610
Subordinate	2,315	777 (33.6%)	0.575
Main conjunct	2,139	812 (38.0%)	0.558
Infinitival	209	78 (37.3%)	0.483
Subordinate conjunct	272	97 (35.7%)	0.467
Other	72	21 (29.2%)	0.314
Total / Range	6,994	2,427 (34.7%)	0.296

The values of these factor weights relative to each other paint a rather confusing picture. A series of likelihood ratio tests (LRTs) show that the probability of left-of-P placement for each clause type category does not differ significantly from that of its immediately adjacent neighbour(s). In other words, when all other variables in the model are controlled for, left-of-P placement is no more or less probable: in non-conjunct subordinate clauses than in main clauses (conjunct or non-conjunct); in infinitival clauses than in conjunct clauses (main or subordinate); or in 'other' clauses than in subordinate conjunct clauses. This suggests there is no significant effect of any of the factors that most obviously distinguish these clause types from one another, i.e. clause embedding, clause conjunction and clause finiteness. If there were a significant systemic correlation between PPOP placement and clause embedding, we would expect to see a significant difference between the factor weight values for each of the two main clause categories on the one hand and the factor weight value for each of the other clause types on the other, but this is clearly not the case. Likewise, if there were a significant systemic correlation between PPOP placement and clause conjunction, we would expect to see a significant difference between the factor weight values for main clauses and for main conjunct clauses as well as between the factor weight values for subordinate clauses and for subordinate conjunct clauses. Again, this is not the case. Lastly, if there were a significant systemic correlation between PPOP placement and clause finiteness, we would expect to see a significant difference in factor weight values between each type of finite clause on the one hand and each type of non-finite clause, i.e. infinitival and 'other', on the other. Once again, this is not the case.

In section 3.5.3.1, I noted that Taylor (2008) found the effect of clause type to vary according to whether or not the PPOP is governed by *to* 'to'. The reason Taylor split her data in this way was due to a diachronic difference between *to* data and other data in her particular model of variation in PPOP placement in her particular sample. As discussed earlier in this chapter, I found no clear evidence of such a difference between *to*-PPOPs and other PPOPs, at least not when time is modelled according to manuscript date. Having split her sample in this way, Taylor found that *to*-PPOPs are slightly more likely to be specially placed in non-conjunct main clauses than in non-conjunct subordinate clauses

(*ibid*: 364), while other PPOPs are significantly *less* likely to be specially placed in non-conjunct main clauses than in non-conjunct subordinate clauses (*ibid*: 351–2).[87] When my sample of *to*-PPOPs are analysed separately from my sample of other PPOPs, I find that special placement is no more or less likely in non-conjunct main clauses than in non-conjunct subordinate clauses for either set of pronouns. In other words, Taylor's finding of a difference in the behaviour of *to*-PPOPs in comparison to other PPOPs according to CLAUSE TYPE disappears when the sample is extended and a larger number of variables are controlled for.

CLAUSE TYPE is ranked fourth in terms of effect size in Table 5.3, but the data in Table 5.13 show that its factor weight range is considerably extended by the 'other' category. As this category represents a very small number of PPOPs in two different constructions, it would not be unreasonable to exclude these data altogether. In addition, the much lower ranking of CLAUSE TYPE in terms of its unique explanatory power (see Figure 5.2) indicates that only a small proportion of the variation with which CLAUSE TYPE is associated is uniquely attributable to this variable and, further, that it uniquely accounts for a very small proportion of variation in PPOP placement overall. As the factor weights for CLAUSE TYPE depict a cline of probabilities that cannot be grouped into distinctive statistical sets according to the features that most obviously characterise different clause types, I conclude that the small amount of variation in PPOP placement associated with CLAUSE TYPE has no obvious linguistically meaningful interpretation. We will, however, encounter an alternative perspective on these results in the final section of Chapter 6. There I show that the correlation between CLAUSE TYPE and PPOP placement can be perfectly captured by the introduction of a new variable that distinguishes between finite and non-finite main verbs.

5.10 Summary

This chapter has highlighted some of the problems that can emerge when undertaking a multivariate analysis of a non-designed sample of data, and has suggested some simple ways in which the consequences of these problems can be

---

[87] Note that Taylor's (2008) factor weight values represent the predicted probability of *right*-of-P placement. A predicted probability of left-of-P placement (pL) can be converted into a predicted probability of right-of-P placement (pR) by means of the formula: $1 - p$L. Likewise, pR can be converted into pL by the formula $1 - p$R.

minimised. The results of the regression for six of the independent variables introduced in Chapter 3 were presented and discussed. I argued that the results indicate PPOP placement in Old English is insensitive to pronoun number and reflexivity, and that there is no real change in the frequency of left-of-P placement over the course of the Old English period. I argued, further, that the evidence suggests PPOP placement is also insensitive to the type of clause in which it appears and, with one exception, to the particular main verb it co-occurs with. *Habban* 'to have' appears to inhibit special placement of PPOPs quite significantly, although the reason for this is presently unclear. Evidence of a statistically significant correlation between PPOP placement and subject form was considered from a number of angles but this too escapes any obvious linguistic interpretation.

Chapter 6 Substantive variables

6.1 Introduction

The variables dealt with in this chapter are those that appear to correlate in a linguistically meaningful, but non-categorical, way with PPOP placement once other variables in the model are controlled for, namely: the extra-linguistic variables LATIN INTERFERENCE (section 6.2) and DIALECT (section 6.3); the pronoun-related variable PERSON (section 6.4); the clause-related variable NARRATIVE MODE (section 6.5); the lexical variable PREPOSITION (section 6.6); and the variables related to PP position, i.e. LINEAR ORDER OF PP AND V and ADJACENCY OF PP AND V (section 6.7). The background to the generation and interpretation of the results presented in this chapter is explained in Chapter 5.

6.2 Latin interference

The factor weight values for LATIN INTERFERENCE are given in Table 6.1. Recall that these values have been estimated by reference to late West Saxon data only as this is the only sample of data that allow for LATIN INTERFERENCE effects to be estimated independently of DATE and DIALECT effects (see section 5.2). The 24 PPOPs occurring in late West Saxon manuscripts which are classified neither as translations nor non-translations were excluded from the analysis on the basis that a factor weight value for this small set of pronouns would be meaningless. Horizontal lines are added to Table 6.1 to indicate the statistically significant divisions of the variable's categories.

Table 6.1 Factor weights for LATIN INTERFERENCE

Factor	PPOPs	Left-of-P	Factor weight
Non-translations	1,956	915 (46.8%)	0.755
Non-biblical, unmatched	220	112 (50.9%)	0.745
Non-biblical, unsampled	137	48 (35.0%)	0.602
Biblical, unmatched	482	93 (19.3%)	0.388
Non-biblical, matched	86	19 (22.1%)	0.385
Biblical, matched	554	58 (10.5%)	0.341
Biblical, unsampled	428	38 (8.9%)	0.263
Total / Range	3,863	1,283 (33.2%)	0.492

These factor weight values are exactly in line with Taylor's (2008: 354–5) description of direct and indirect interference effects of Latin PP word order on PPOP placement in translated texts.

The direct interference effect is evidenced by a clear dispreference for left-of-P placement when the Old English PP is a direct translation of a Latin PP (0.385 for matched non-biblical data, 0.341 for matched biblical data) compared to a clear preference for left-of-P placement in non-translations (0.755). As the likelihood ratio test ('LRT') shows there is no statistically significant difference between the weight values for directly translated PPs in biblical and in non-biblical translations, the results support Taylor's conclusion that the direct interference effect is equally strong in both types of translation.

For unmatched PPs in non-biblical translations, i.e. those that do not correspond directly to a PP in the Latin, there is no evidence of any influence of Latin PP word order. This is apparent from the fact that the LRT shows that the factor weight value for unmatched non-biblical data (0.745) does not differ statistically from the factor weight value for data in non-translations (0.755). On the reasonable assumption that unsampled non-biblical data includes both matching and non-matching PPs, we would expect the factor weight value for unsampled non-biblical data to fall somewhere between that for unmatched data (0.745) and matched data (0.385), which is indeed what we find (0.602).

Data in biblical translations, by contrast, exhibit an indirect as well as a direct interference effect. The indirect effect is evidenced by a clear dispreference

for left-of-P placement when the Old English PP is <u>not</u> a direct translation of a Latin PP (0.388) compared to a clear preference for left-of-P placement in non-translations (0.755). Further, as the LRT shows there is no statistically significant difference between the weight values for matched and unmatched biblical data, we may also conclude that the indirect effect is just as strong as the direct effect. In Taylor's (2008) sample of biblical data, the two types of interference effect were equally strong for *to*-PPOPs (ibid: 365, table c10), but the direct effect was found to be stronger than the indirect effect in her sample of PPOPs governed by other prepositions (*ibid*: 354, table 14). Exactly the same contrast is manifest between my sample of *to*-PPOPs and my sample of other PPOPs. No doubt other differences would emerge if results for PPOPs governed by each preposition were separately calculated, but there are insufficient data for most prepositions to undertake a comprehensive analysis of the interaction between LATIN INTERFERENCE and PREPOSITION. We may conclude, then, that the direct effect and the indirect effect are of the same magnitude overall, but that the indirect effect seems to be weaker, although still statistically significant, for PPOPs governed by the most frequently occurring preposition in the corpus, i.e. *to* 'to'. As the direct and indirect effect are of the same magnitude overall, we would expect the factor weight value for unsampled biblical data not to differ significantly from that of matched data (0.341 for data in biblical translations, 0.385 for data in non-biblical translations) and that of unmatched biblical data (0.388). Although the factor weight value for unsampled biblical data — 0.263 — is lower than each of these values, it is not significantly lower: the LRT shows no significant change in model fit when data belonging to the last four factors listed in Table 6.1 are treated as belonging to a single category.

In sum, the direct and indirect effects of Latin interference identified by Taylor (2008: 354–5) could hardly be any clearer. As Taylor's sample was not limited to PPOPs in late West Saxon manuscripts, the data in Table 6.1 provide additional reassurance that her findings are not confounded by either DATE or DIALECT effects. Nothing need be added to her conclusions — or those of Minkoff (1976, 1977) — about why indirect interference is found only in biblical texts (*ibid*: 355–6). As Taylor notes, extant accounts of translation strategies employed during the Old English period describe a deliberate effort to represent

the language of the exemplar more faithfully when translating biblical texts. Consequently, it is not surprising that the output of biblical translators exhibits a greater sensitivity than that of non-biblical translators to the general regularity of head-initial PPs in their Latin exemplars.[88]

The only context in which translators appear to have been completely uninfluenced by PP word order in their Latin exemplar is when translating head-final *cum*-PPs. As was shown in Table 3.24, 147 PPs in translated texts are matched with a Latin *cum*-PP and only five occur with a left-of-P PPOP in total. 128 (87%) of the PPs matched with a *cum*-PP are headed by *mid* 'with' and example (1), from the Old Testament book of *Exodus*, illustrates the typical correspondence between a Latin *cum*-PP and an Old English *mid*-PP in the analysed sample where the latter is a direct translation of the former.

(1) a.   TULIT quoque Moyses ossa   Ioseph   **secum**

carried also   Moses   bones of-Joseph him-with

'Moses also carried Joseph's bones with him'

(Crawford 1922: 248)

   b.   WITODLICE Moyses nam Iosepes   ban   **mid him**

truly   Moses   took Joseph's bones with him

'Truly Moses took the bones of Joseph with him'

(cootest,Exod:13.19.2907)

Data presented in Chapter 3 show that *mid* 'with' is the second most frequently occurring PPOP governor in the corpus and is also among those that occur most frequently with the pronoun to the right rather than the left apart from *be*, *for* and *þurh*, which never occur with a left-of-P PPOP (see Table 3.10). Later in this chapter, we will see that *mid* is in fact the preposition most likely to govern a right- rather than left-of-P PPOP (apart from *be*, *for* and *þurh*) when the effects of other variables are controlled for. Clearly, when it came to *mid*-PPs, the translators' native syntax was strongly resistant to the influence of Latin PP word order.

---

[88] For an historical study of hermeneutics, including an introduction to medieval translation theory, see Copeland (1991).

6.3 Dialect

In section 5.2 I elected to conflate some of the DIALECT categories originally identified in Chapter 3 due to their small numbers of PPOPs. Table 6.2 gives the results of the multivariate analysis for the revised categories. I also elected to exclude data from biblical translations from estimations of DIALECT effects to ensure the results are not confounded by DATE and/or LATIN INTERFERENCE effects. The results in Table 6.2 are therefore for PPOPs in non-translations and non-biblical translations only.

Table 6.2 Factor weights for DIALECT

Factor	PPOPs	Left-of-P	Factor weight
Other	851	404 (47.5%)	0.586
West Saxon	3,189	1,387 (43.5%)	0.548
West Saxon+Anglian	972	282 (29.0%)	0.440
West Saxon+Anglian Mercian	518	165 (31.9%)	0.426
Total / Range	5,530	2,238 (40.5%)	0.160

The results show that there is a slightly greater preference for left-of-P placement over right-of-P placement among data in the West Saxon category and a slightly greater preference for right-of-P placement over left-of-P placement among data in the Anglian influenced categories. It is interesting to see that the factor weight values for the two Anglian influenced categories are almost the same. This reinforces the contrast, albeit a small one, between the behaviour of PPOPs in Anglian influenced texts and the behavour of those in West Saxon texts. Table 6.2 also indicates that the probability of left-of-P placement is greatest in texts whose language has not been classified as belonging to either of these categories. It is important to remember, however, that texts in the 'other' category may very well belong to one or more of the categories in Table 6.2: for the most part, inclusion in the 'other' category simply indicates that the dialectal profile of the text is not identified in the documentation for the Helsinki Corpus of English Texts.

The LRT confirms, firstly, that there is no statistical reason for maintaining the distinction between the West Saxon+Anglian and West

Saxon+Anglian Mercian categories. In other words, the probability of left-of-P placement does not differ significantly between these two categories. Secondly, the LRT confirms that the probability of left-of-P placement in Anglian-influenced materials is significantly lower in comparison to the probability in West Saxon materials. This suggests a genuine dialectal difference, but the difference is much too slight for the frequency of left-of-P placement to be regarded as a useful diagnostic tool for unprovenanced materials. Thirdly, the LRT shows there is no statistical reason for maintaining the distinction between the West Saxon category and the 'other' dialect category. This could indicate that a significant amount of data in the 'other' category derives from West Saxon materials or from materials whose language is indistinguishable from West Saxon in terms of PPOP placement. On the other hand, if a significant amount of data in the 'other' category properly could be shown to derive from Anglian-influenced materials, this would indicate that the relationship between dialect and PPOP placement is not accurately reflected in Table 6.2.

Overall, the results suggest that left-of-P placement is slightly more likely to occur in West Saxon materials than in Anglian-influenced materials. Although this difference is large enough to reach statistical significance, it is not sufficiently large for practical application. In addition, without knowing where data in the 'other' category belong, we cannot be confident that the difference in PPOP placement between West Saxon and Anglian-influenced materials is not in fact smaller, or indeed greater.

6.4 Person

The results for PERSON, given below, show that left-of-P placement is favoured when the PPOP has third person reference and disfavoured when it has first or second person reference.[89] This confirms that the correlation between PPOP placement and the grammatical person of the pronoun evident from the bare

---

[89] In Chapter 5 I elected to exclude all PPOPs that are neither clearly nor probably dative from the variation analysis. This removes all non-third person PPOPs governed by *on* 'on, in', *ongean* 'towards, against' and *wið* 'against' from the sample. As explained in section 5.2, the sample was re-balanced by excluding *all* PPOPs governed by these three prepositions from the estimation of PERSON effects. This explains why the number of PPOPs analyzed in Table 6.3 (6,341) is less than the entire modelled sample (N=6,994); the difference of 653 represents the number of third person PPOPs governed by *on* (N=457), *ongean* (N=105) and *wið* (N=91).

frequencies in Chapter 3 remains statistically significant when other variables in the model are controlled for.

Table 6.3 Factor weights for PERSON

Factor	PPOPs	Left-of-P	Factor weight
Third person	4,597	1,879 (40.9%)	0.666
First person	956	114 (11.9%)	0.421
Second person	788	80 (10.2%)	0.409
Total / Range	6,341	2,073 (32.7%)	0.257

The close proximity of the factor weight values for first and second person indicates that left-of-P placement of first person PPOPs is no more or less probable than left-of-P placement of second person PPOPs, and this is confirmed by the likelihood ratio test. This provides strong evidence that the correlation between grammatical person and PPOP placement distinguishes third person pronouns on the one hand from 'non-third' person pronouns on the other.

In Alcorn (2009: 436–7), I reported that third person PPOPs more frequently occur in a left-of-P position in comparison to non-third person PPOPs in all but two of the 33 text files in the YCOE that contain at least ten third person and at least ten non-third person PPOPs. The correlation is thus remarkably consistent. The two noted exceptions are *History of the Holy Rood Tree* (corood), in which left-of-P placement of non-third person PPOPs (N=20, 8 [40%] left-of-P) occurs with approximately the same frequency as third person PPOPs (N=82, 31 [38%] left-of-P), and *Alexander's letter to Aristotle* (coalex), which provides an unusually high number of non-third person PPOPs (N=43, 20 [47%] left-of-P) in comparison to third person PPOPs (N=15, 4 [27%] left-of-P). All but four of the 43 non-third person PPOPs in coalex are first person pronouns. As the text is a personal travelogue, its high proportion of PPOPs with first person reference is therefore unremarkable. Once pronouns associated with (near) knockout factors are excluded, the picture changes only slightly. Corood ceases to be an exception (first/second person: N=14, 3 [21%] left-of-P; third person: N=70, 26 [37%] left-of-P) and left-of-P proportions for coalex remain about the same (first/second person: N=27, 12 [44%] left-of-P; third person: N=8, 2 [25%] left-of-P). Although

three new exceptions emerge, two arise on the basis of a single example — cogenesiC (*Genesis*, as found in Cambridge, Corpus Christi College, 201), in which there is one non-third person left-of-P PPOP out of 30 (3%) but all 34 third person PPOPs are right-of-P; and cowsgosp,Jn (*West Saxon Gospels*, *John*), in which there is one non-third person left-of-P PPOPs out of 60 (2%) but all 225 third person PPOPs are right-of-P — and in the third text, comary (*Mary of Egypt*), left-of-P placement of non-third person PPOPs (N=20, 11 [55%] left-of-P) occurs with the same frequency as third person PPOPs (N=33, 18 [55%] left-of-P). Overall, the only text file in which left-of-P placement occurs significantly more frequently with non-third person PPOPs than with third person PPOPs is coalex, although non-third person PPOPs in comary — and in coalex — do show a remarkably high frequency of left-of-P placement.[90]

There are also indications that the correlation between PPOP placement and grammatical person is not confined to the prose. Within the York Poetry Corpus (Pintzuk & Plug 2001), 62 of the 97 simple third person PPOPs are left-of-P (64%) compared to 21 of the 83 simple non-third person PPOPs (25%). This difference is significant ($\chi^2 = 27.46$, $p < 0.0001$) although it does not take account of the effects of any other variables.

There is also evidence that the correlation between grammatical person and PPOP placement has a dialectal distribution. When PERSON and DIALECT are replaced with a cross-product variable that allows their combined effects to be estimated, there is a significant improvement in model likelihood — a clear sign that the effects of PERSON and the effects of DIALECT are not entirely independent of one another.[91] The locus of this interaction can be identified by comparing the factor weight values for each of the contexts defined by this cross-product variable. These values are provided in Table 6.4 in which the 'other' dialect

---

[90] The governors of the non-third person left-of-P PPOPs are: in comary, *fram* x5, *to* x3 and *beforan*, *mid* and *ongean* x1 each; and in coalex, *to* x5, *mid* x4, *togeanes* x2 and *æfter* x1.

[91] Goldvarb is not configured to test for the independence of variables, although it does supply information that allows this to be done by hand (Guy 1988: 134–6, Sigley 2003: 234–8, Tagliamonte 2006: 151). The test statistic, G^2, is twice the difference in log likelihood between the two models and can be compared to a chi-square distribution at the difference in degrees of freedom between the two models. A chi-square value ≤ 0.05 indicates a statistically significant improvement in model fit. As said, there is a highly significant increase in model likelihood on replacing PERSON and DIALECT with a PERSON*DIALECT cross-product variable (G^2 (df2) = 24.198, $p < 0.0001$).

category represents data in texts whose language is unclassified for dialect by the editors of the Helsinki Corpus or else is known to include elements of at least one variety that is neither West Saxon nor Anglian. The figures in parentheses in each cell indicate the number of left-of-P PPOPs as a proportion of all PPOPs in that particular context.

Table 6.4 PERSON*DIALECT (input value = 0.106)

	West Saxon	Anglian-influenced	Other
Third	0.644	0.461	0.641
	(1,100/2,075 [53%])	(332/1,020 [33%])	(294/571 [51%])
Non-third	0.323	0.468	0.463
	(99/781 [13%])	(53/342 [15%])	(29/169 [17%])

Notice, firstly, that the figures in parentheses show that third person PPOPs are placed left-of-P more frequently than non-third person PPOPs in each of the three dialect categories. When other variables in the model are controlled for, however, the results indicate a strong effect of PERSON in the West Saxon materials (factor weight range = 0.321) and, to a lesser extent, in materials belonging to the 'other' dialect category (factor weight range = 0.178), but no statistically significant correlation is evident for data in the Anglian-influenced texts (factor weight range = 0.007). A large proportion of the West Saxon data in Table 6.4 (53% of third person and 58% of non-third person) comes from the YCOE's four large Ælfrician texts, i.e. *Catholic Homilies I* and *II*, *Supplementary Homilies* and *Lives of Saints*. When these data are excluded from the analysis, the effect of PERSON remains statistically significant for West Saxon. So the correlation between PPOP placement and PERSON in the West Saxon materials is not confined to these four texts.

It is widely acknowledged that language exhibits a fundamental opposition between the first and second person on the one hand, and the third person on the other. Manifestations of this opposition play out in different ways in different languages (Bhat 2004: 91–118, Siewierska 2004: 5–8). For example, in languages with a grammatical category of gender, gender is distinguished in the third person but only rarely in the second and hardly ever in the first (Bhat *ibid*: 109–12, Siewierska *ibid*: 104–7). Old English is a classic example of this: the personal

pronoun paradigm (see Table 3.3 for object forms) shows gender distinctions in the third person alone. An opposition between first and second person on the one hand and third person on the other is evident in two other areas of Old English morphosyntax: accusative and dative are distinct for third person personal pronouns but not for first and second person personal pronouns; and whereas third person genitive forms are indeclinable (Mitchell 1985: §289), first and second person genitive forms can be declined like strong adjectives for case, gender and number, e.g. (2).[92]

(2)   hi   beoþ mine
       they are   mine-NOM-MASC-PL
       'they will be mine'

(cootest,Gen:48.5.2104)

So there is already evidence that Old English grammar treats first and second person pronouns differently from third person pronouns, although admittedly each involves a categorical distinction rather than the probabilistic distinction that is clearly evident in terms of PPOP placement. A fourth distinction, and one which is probabilistic in nature, involves subject pro-drop, i.e. elision of a topical subject. Both Berndt (1956) and van Gelderen (2000: 132–4) provide evidence that shows first and second person *pro*-drop is considerably less frequent than third person *pro*-drop in Old English.

Cross-linguistic manifestations of the third person *vs.* non-third person opposition are commonly ascribed to the fact that the first and second person alone denote participants of the speech act (e.g. Bhat 2004: 91, Siewierska 2004: 7). This view is supported by the existence of what Bhat (*ibid*: 134) refers to as 'two-person' languages, i.e. languages with a personal pronoun system that expresses first and second person alone, and which typically use forms identical to, or derived from, demonstratives to refer to parties other than speaker or addressee, e.g. Basque, Imbabura Quechua and Lak (Bhat *ibid*: 132–47, Siewierska *ibid*: 5–6). There is no reason to think that manifestations of the

---

[92] Mitchell (1985: §§289–2) notes that *sin* 'his, her, their' is declinable. However, this form is almost completely absent from the prose and occurs only spasmodically in the poetry so was probably archaic even in Old English.

person opposition in Old English have a different basis. An alternative characterisation, in terms of animacy of the pronoun's referent, is considered and rejected in Alcorn (2009: 438–42), as discussed in Chapter 3.

If Lyons (1977: 638) is correct in his statement that the third *vs.* non-third person opposition in natural language is 'fundamental and ineradicable', we would not be surprised to find evidence of it in varieties closely related to Old English and perhaps in the behaviour of simple personal pronoun objects of verbs. The fact that first and second person genitive forms are declinable but third person genitive forms are not in Old Frisian (Bremmer 2009: 57) as well as in Old English provides one indication that the person effect in Old English has a long history, as do similarities between Old English and other older Germanic varieties with respect to *pro*-drop patterns (van Gelderen 2000: 136). The placement of PPOPs in Old Saxon poetry in particular provides a direct indication that the correlation between PPOP placement and PERSON in Old English has a history older than the language itself. Wende's (1915: 235) analysis of PPOP placement in the *Heliand*, the largest known Old Saxon text, shows that 57 (65%) of the 88 simple third person PPOPs occur in a left-of-P position compared to 9 (18%) of the 50 simple non-third person PPOPs ($\chi^2 = 27.95$, $p < 0.0001$). Although the numbers are small, this is a strikingly similar pattern to that found for the Old English poetry, for which the comparative proportions, given earlier, are 64% and 25% respectively. Wende's data are certainly consistent with the hypothesis that the correlation between person and placement of PPOPs in Old English could go back at least to the West Germanic origins of the language.

Data on Old English person personal pronoun objects of verbs ('PPOVs'), however, suggest that the robust correlation between PPOP placement and grammatical person does not go beyond the PP domain. Using the same sets of data for third person PPOVs that were introduced in Chapter 4, Table 6.5 compares their frequency of occurrence in clitic position B with equivalent data for non-third person PPOVs, and Table 6.6 does likewise with respect to clitic position C. Percentages in parentheses indicate the proportion of the row total.

Table 6.5 Frequency of PPOV placement in pos. B by grammatical person

	($\flat a$/$\flat onne$ –) V_{FIN} – PPOV – Subject	($\flat a$/$\flat onne$ –) V_{FIN} – Subject (…) PPOV	Total
Third person	262 (47%)	292 (53%)	554
Non-third person	92 (85%)	16 (15%)	108
Total	354 (53%)	308 (47%)	662

Table 6.6 Frequency of PPOV placement in pos. C by grammatical person

	COMP–PPOV–Subject	COMP–Subject–(…)PPOV	Total
Third person	215 (45%)	258 (55%)	473
Non-third person	39 (30%)	89 (70%)	128
Total	254 (42%)	347 (58%)	601

The data in the second of these tables are consistent with the trend evident among PPOPs in that PPOV placement in position C occurs significantly less frequently with first and second person PPOVs than with third person PPOVs (χ^2 = 9.27, p = 0.002), the first table shows the opposite trend — and significantly so (χ^2 = 52.16, p < 0.0001).

In summary, the multivariate analysis confirms an independent correlation between PPOP placement and the grammatical person of the pronoun. These results are in line with previous univariate analyses, which show that: third person PPOPs are significantly more likely than first and second person PPOPs to appear in a special position; and that there is no significant difference in the frequency of special placement of first and second person PPOPs. The analysis of PPOP placement in a corpus of Old English poetry provides evidence that the effect is not confined to the prose. The present analysis has also found evidence that suggests that the correlation is not present in all dialects of Old English — there are no signs of it in data that derive from Anglian-influenced texts — and that it is confined to pronominal objects of prepositions in particular. Third *vs.* non-third person asymmetries are manifest in many different ways in many different languages, and Old English is already known to provide three such examples. Such asymmetries are generally associated with the fact that the first and second person denote speech act participants, and I have not found a more insightful way

to describe the Old English data. Finally, Wende's data from the Old Saxon *Heliand* suggests that the correlation between PPOP placement and PERSON in Old English may have been inherited from England's West Germanic settlers. If so: (a) we might expect to find further evidence in other, especially historical, West Germanic varieties, a matter I leave for future research; and (b) the presence of the feature in West Saxon but not in Anglian-influenced materials might then be seen as a conservatism rather than an innovation of West Saxon.

6.5   Narrative mode

The results for NARRATIVE MODE, given in Table 6.7, show that left-of-P placement is significantly less likely in sequences of direct speech, i.e. in mimetic contexts, than elsewhere, i.e. in diegetic contexts, when all other variables in the model are controlled for.

Table 6.7 Factor weights for NARRATIVE MODE

Factor	PPOPs	Left-of-P	Factor weight
Diegetic	4,804	2,093 (43.6%)	0.578
Mimetic	2,190	334 (15.3%)	0.422
Total / Range	6,994	2,427 (34.7%)	0.156

This correlation appears to be quite robust within the corpus, as can be seen from the data in Appendix I, which compares bare frequencies of left-of-P placement in mimetic contexts with frequencies in diegetic contexts in each of the twenty-eight texts that supply at least fifty PPOPs to Table 6.7. The data in Appendix I show that left-of-P placement is consistently more frequent in diegetic contexts in all but two texts: *Boethius* (coboeth), in which only 5 out of 138 PPOPs occur in a diegetic context; and *Genesis* (cogenesiC), in which only 1 of all 64 PPOPs is specially placed. In all of the other texts for which a chi-square test is possible except *Bede*, left-of-P placement is significantly more frequent in diegetic contexts.

On the basis of the data in Appendix I, it might rightly be asked whether it is appropriate to measure the extent of the correlation between PPOP placement

and NARRATIVE MODE by reference to data from all of the YCOE's text files. As said, almost all PPOPs in *Boethius* occur in mimetic contexts, whereas almost all PPOPs in the four versions of the *Anglo-Saxon Chronicle*, for example, occur in diegetic contexts (450/453). However, when the multivariate model is adjusted so that the correlation between PPOP placement and NARRATIVE MODE is estimated only from data in text files that contain at least ten PPOPs in mimetic contexts and at least ten in diegetic contexts, there is no material change to the factor weight values given in Table 6.7.

Another set of data that might justifiably be excluded from the analysis of NARRATIVE MODE are those PPOPs that occur as the object of a *cweðan to* 'to say to' construction. This is by far the most frequently occurring verb + preposition combination in the YCOE, accounting for 17% of the data in Table 6.7. Given that this construction typically introduces a sequence of direct speech, it is not surprising to find that 90% of *cweðan to* PPOPs occur in diegetic contexts. Again, however, when the multivariate model is yet further adjusted to exclude *cweðan to* PPOPs also, there is still no material change to factor weight values in Table 6.7.

Too few PPOPs occur in mimetic contexts in the poetry to permit a meaningful analysis of PPOP placement by narrative mode in that genre. Special placement of simple personal pronoun objects of verbs ('PPOVs') in the prose, however, shows no statistically significant correlation with narrative mode. Using the same sets of data for third person PPOVs that were introduced in Chapter 4, Table 6.8 compares their frequency of occurrence in clitic position B according to narrative mode, and Table 6.9 does likewise with respect to clitic position C. Percentages in parentheses indicate the proportion of the row total.

Table 6.8 Frequency of third p. PPOV placement in pos. B by narrative mode

	(*þa*/*þonne* –) V_{FIN} – PPOV – Subject	(*þa*/*þonne* –) V_{FIN} – Subject (…) PPOV	Total
Diegetic	232 (46%)	272 (54%)	504
Mimetic	30 (60%)	20 (40%)	50
Total	262 (47%)	292 (53%)	554

Table 6.9 Frequency of third p. PPOV placement in pos. C by narrative mode

	COMP – PPOV – Subject	COMP – Subject – (…) PPOV	Total
Diegetic	186 (47%)	210 (53%)	396
Mimetic	29 (38%)	48 (62%)	77
Total	215 (45%)	258 (55%)	473

The data in the second of these tables are consistent with the trend evident among PPOPs in that PPOV placement in position C occurs less frequently in mimetic contexts than in diegetic contexts, although the difference is not significant ($\chi^2 = 2.25$, $p = 0.13$). The trend in Table 6.8 is in the opposite direction, but again the difference is not significant ($\chi^2 = 3.56$, $p = 0.06$). As with the correlation with PERSON, it appears that the correlation with NARRATIVE MODE does not go beyond the PP domain.

What, then, can be made of the correlation between PPOP placement and NARRATIVE MODE? Waterhouse (1976: 83) argues that a writer's choice between different methods of presenting or suggesting speech in Old English 'is one of his most potent means of creating characterisation and of controlling the effect of his story.' By contrasting Ælfric's use of direct and indirect speech as didactic devices in some of his saints' lives, Waterhouse finds evidence that Ælfric tended to reserve direct speech for his 'good' characters — especially the central saint — and indirect speech for his 'bad' characters (*ibid*: 103). So there is some evidence that direct speech is used purposefully in Old English to convey a greater meaning than the sum of the words it consists of, at least in hagiographic texts. Bjork (1985) goes further and examines the syntax, rhetoric and diction of the dialogues of characters in the Old English verse saints' lives and concludes that, by these means, each of these poems

> 'tends to concentrate a great deal of meaning in direct discourse, consistently juxtaposing a confusion or lack of focus in the dryht of hell [i.e. when the character is under Satan's power (Russell 1984: 135) — RA] or in the mutable world to a perfect symmetry and harmony in the dryht of heaven [i.e. when under Christ's power (Russell *ibid*) — RA] or in the real world'
>
> Bjork (1985: 127)

This suggests that, at least in the texts examined by Bjork, syntax is used as a stylistic device to reinforce the central themes of the text, although it is difficult to imagine how PPOP placement in direct speech in particular could possibly contribute to the text's meaning.

Mitchell (1985: §1635) admits it is appealing to suppose that written representations of direct speech might mirror features of the spoken language, but is not convinced that investigations along this line would 'produce anything very solid'. His reservation is not without support:

> In Old and Middle English it was rare to present speech as a different form of language from narrative [...] In Middle English there were some developments in the representation of speech, but these were not generally towards a realistic representation. Most authors continued to use the same vocabulary and syntax in speech as elsewhere.
>
> Blake (1979: 157)

There are, however, two points at issue here. The first is that Mitchell and Blake appear to be commenting on qualitative rather than quantitative differences. It is very probably true that there is no syntactic phenomenon found in mimetic contexts that it is not also found in diegetic contexts or *vice versa* in Old English, but an asymmetric distribution of *variants* of a syntactic variable is a possibility neither Mitchell nor Blake explicitly rule out. The second point concerns how differences between the language of direct speech and the language of the narrative should be interpreted. Mitchell and Blake identify — and reject — one possibility, i.e. that the language of direct speech might more closely approximate the spoken language, but there is at least one other possibility. In a study of the syntax of direct speech in Vedic Sanskrit, Jamison (1991) found a number of differences in comparison to the syntax of the narrative. For example, she found an almost complete absence of non-finite verb forms in direct speech, which contributes to 'the peculiar flavour of direct speech in Vedic prose, [i.e.] its stripped down, abrupt quality' (*ibid*: 99). Jamison (*ibid*: 96) argues that the differences she identifies need not necessarily reflect differences between the spoken and written language, but rather that 'there may well be a stereotyped style for speech as artificial as that of the other rhetorical patterns.' To put it another way, the essence of Jamison's argument is a suggestion that the language of direct

speech might have a distinctive register — i.e. a variety bound to a particular discursive situation (Halliday, McIntosh & Strevens 1964: 87) — which would provide a plausible way to interpret quantitative (and qualitative) linguistic differences between the two modes of narration for any language. If PPOP placement is indeed sensitive to register in this way, we might reasonably expect to find other quantitative differences between the language of direct speech and the language of the surrounding narrative, but this is virgin territory as far as the language of Old English prose is concerned.

An alternative way to look at the correlation between PPOP placement and NARRATIVE MODE is to assume the existence of some other variable that correlates with both PPOP placement and narrative mode in such a way that the results in Table 6.7 and Appendix I may be seen as epiphenomenal. One possible candidate variable is PERSON. In section 6.4, it was noted that first and second person PPOPs are significantly less likely than third person PPOPs to be specially placed. But it is also the case that the majority of third person PPOPs (86%) appears in diegetic contexts, while the majority of the (predominantly deictic) non-third person PPOPs (85%) appears in mimetic contexts.[93] Could it be, then, that the probability of left-of-P placement according to NARRATIVE MODE simply reflects this asymmetric distribution of PPOPs according to person? The answer is no. When PERSON and NARRATIVE MODE are replaced with a cross-product variable that models their combined effects, there is no significant change in model likelihood. This indicates that the correlation between PPOP placement and NARRATIVE MODE on the one hand, and the correlation between PPOP placement and PERSON on the other, are entirely independent of one another.

The data considered in this section suggest there is a significant difference in the frequency of special placement of PPOPs between the language of direct speech and the language of the narrative proper in Old English. There is little support for assuming that the language of direct speech more closely approximates spoken Old English, so an explanation for the correlation between pronoun placement and NARRATIVE MODE is more likely to lie elsewhere. One possibility, which has indirect support from the comparison of the syntax of direct

---

[93] Unsurprisingly, most first and second person PPOPs in diegetic contexts occur in texts that some might regard as composed largely, if not entirely, of speech, e.g. homilies and letters. As Taylor (2003: Direct speech) points out, the line between narrative and personal comment is often difficult to establish for these types of text in particular.

speech and the syntax of the narrative proper in at least one other text language, is that the correlation reflects a difference in register.

6.6 Preposition

The placement of PPOPs governed by the (near) knockout prepositions has already been discussed in Chapter 4. I concluded that the invariable right-of-P placement of PPOPs governed by *þurh* 'through' is probably linked to this preposition's strong association with accusative case, and that the invariable right-of-P placement of PPOPs governed by *be* and *for* is due to the prosodic weakness of these forms, which requires them to be phonologically dependent on their object. I also suggested that the near invariable left-of-P placement of PPOPs governed by *betweonum* 'between' is the result of a reanalysis of the variant form types of BETWEEN at a very early stage of Old English, resulting in *betweonum* forms being interpreted as the type for special clitic objects in particular.

Factor weight values for the other prepositions are given below. The low factor weight value for non-*betweonum* forms of BETWEEN is expected under the analysis of the variant types of Old English BETWEEN proposed in Chapter 4.

Table 6.10 Factor weights for PREPOSITION

Factor	PPOPs	Left-of-P	Factor weight
ongean 'towards, against'	105	101 (96.2%)	0.968
togeanes 'against, towards'	122	107 (87.7%)	0.927
miscellaneous	484	242 (50.0%)	0.523
to 'to'	3,203	1,334 (41.6%)	0.514
fram 'from'	457	140 (30.6%)	0.496
on 'on, in'	457	186 (40.7%)	0.475
beforan 'before'	213	41 (19.2%)	0.423
of 'of'	175	58 (33.1%)	0.389
æt 'at'	164	25 (15.2%)	0.259
'between' (except *betweonum*)	139	20 (14.4%)	0.244
æfter 'after'	223	47 (21.1%)	0.225
mid 'with'	1,252	126 (10.1%)	0.143
Total / Range	6,994	2,427 (34.7%)	0.825

As noted in section 5.2, the exclusion of first and second person PPOPs not governed by a dative-favouring preposition leaves *ongean* and *on* represented by third person pronouns only. So whereas the factor weight values for each of the other prepositions in Table 6.10 are controlled for PERSON effects, the weight values for *ongean* and *on* are not. The same is true for *wið* 'against', which is included in the 'miscellaneous' group. Consequently, factor weight values for these three prepositions may be somewhat inflated since third person PPOPs are significantly more likely than non-third person PPOPs to be specially placed.

PREPOSITION ranks as the most important variable in the multivariate model in terms of both effect size, i.e. factor weight range (see Table 5.3), and unique contribution to model likelihood (see Figure 5.2). The factor weight values for individual prepositions show that almost half of this variable's effect size results from the exceptionally high factor weight values for *ongean* and *togeanes*. The 227 PPOPs governed by these two prepositions also make a considerable contribution to its explanatory power. As more than 95% of *ongean*-PPOPs favour the left-of-P variant, there would be justification for excluding these PPOPs from the analysis of variation. I have, however, retained them to draw attention to the similarity between placement of these PPOPs and those governed by *togeanes*. When these PPOPs are excluded from the variation analysis, the value for PREPOSITION in Figure 5.2 reduces by approximately one-third although the ranking of variables remains the same.

Besides their similar semantics and their strong preference for left- rather than right-of-P PPOPs, *ongean* and *togeanes* have other features in common. Etymologically, both are complex forms that derive from the combination of a preposition plus adverb. The adverbial element of both forms goes back to the same word, whose original primary meaning seems to have been 'direct, straight' (*OED* 1989: *again*, *to-gains*). In *togeanes*, the adverbial element shows a case ending, in this case gen. sg. *-es*, which is found with a number of Old English adverbs, e.g. *dæges* 'daily' and *hamweardes* 'homewards', and with some other complex prepositions, e.g. *tomiddes* 'amidst' (Mitchell 1985: §1389). The strong preference for left- rather than right-of-P PPOPs with *ongean* and *togeanes* is not confined to the prose. Although there are just sixteen PPOPs in total in the York Poetry Corpus (Pintzuk & Plug 2001) governed by *ongean* and *togeanes*, in each

case the pronoun occurs in a left-of-P position. Like *ongean* and *togeanes*, *beforan* is also a compound preposition formed from a combination of preposition + adverb. However, the weight value for *beforan* in Table 6.10 indicates that this preposition does not favour special placement of PPOPs, instead seeming to pattern along with simple, i.e. non-compound, prepositions.

Data from the Old Saxon *Heliand* suggest that the preference for left- rather than right-of-P PPOPs with *ongean* and *togeanes* is an inherited feature of Old English. Wende's (1915: 232–3) analysis shows that Old Saxon *angegin* and *togegnes* occur 31 and 25 times, respectively, with a simple personal pronoun object in the *Heliand* and that each of these 56 pronouns is left-of-P. These are the only prepositions occurring with a simple personal pronoun object more than two or three times in the *Heliand* which show a strong preference for the PPOP to be left-of-P, so evidently there was something special about PPOP placement when governed by these two prepositions in Old Saxon too.[94] It seems, then, that the exceptionally high factor weight values for *ongean* and *togeanes* is an echo of their earlier Germanic grammar, although I do not see what might have led to this.

Of the nineteen examples of *ongean* and *togeanes* with a <u>right</u>-of-P PPOP, just two date to early Old English. The example at (3) is from the *Parker Chronicle* and comes from an entry in a hand dated to *c*. 915–930 (Bately 1986: xxi–xliii). The example at (4) is from the early tenth-century *Orosius* (Ker 1957). The concentration of examples with a right-of-P PPOP in the late Old English materials cannot, however, be interpreted as evidence that placement of PPOPs to the right of *ongean* and *togeanes* is a late Old English innovation: the majority of data in the YCOE comes from late manuscripts (see Table 3.20), so we would expect to find far fewer examples in the early material in any event.

(3)   &   þa scipu foran be suðan east andlang sæ **togenes him**

        and the ships went by south east along   sea towards him

        'and the ships sailed by the south-east along the sea to meet him'

                                    (cochronA-2b,ChronA_[Plummer]:911.3.1207)

---

[94] Wende (1915: 233) identifies seven Old Saxon prepositions, including *thuru* 'through' and *be* 'by', which invariably occur with a right-of-P object in the *Heliand*, but none governs more than two PPOPs in total.

(4)   Þa   gegaderade Pholomeus micle fird   **ongean him**

        then gathered   Ptolemy   great army against him

        'Then Ptolemy gathered a great army against him'

<div align="right">(coorosiu,Or_3:11.79.3.1555)</div>

Once weight values for *ongean* and *togeanes* are set aside, those that remain can be seen to form a cline of values rather than distinct sets defined, for example, by prepositional meaning or phonological weight. The likelihood ratio test ('LRT') shows that the value for the miscellaneous category does not differ significantly from the value for *to*, and that the value for *to* does not differ significantly from the value for *fram*. In fact, comparing each weight value below that for *togeanes* with the next in Table 6.10 by means of the LRT shows that there is no significant contrast between any adjacent weight values apart from those for *of* and *æt* and those for *æfter* and *mid*.

At the upper end of this cline lies *to*, the most frequently occurring preposition and also the preposition most likely to govern a left- rather than right-of-P PPOP after *ongean* and *togeanes*. At the other end is *mid*, the second most frequently occurring PPOP governor in the corpus and the preposition *least* likely to govern a left- rather than right-of-P PPOP. The position of the two most frequently occurring prepositions at either end of this cline of weight values suggests that the probability of left-of-P placement of PPOPs is unlikely to be related to preposition frequency. As in the prose, *to*-PPOPs are specially placed more frequently than *mid*-PPOPs in the York Poetry Corpus, although the difference is not statistically significant: 12/31 [39%] *to*-PPOPs are left-of-P compared to 7/29 [24%] *mid*-PPOPs ($\chi^2 = 1.47$, $p = 0.225$). Once again, these patterns mirror patterns in the *Heliand*, in which 28/63 [44%] *to*-PPOPs are left-of-P compared to 1/20 [5%] *mid*-PPOPs ($\chi^2 = 10.39$, $p = 0.001$) (Wende 1915: 233).

As most of the prepositions individuated in Table 6.10 are represented by relatively few pronouns, it is impossible to obtain reliable estimations for interactions involving PREPOSITION. Nevertheless, by comparing the placement of *to*- and *mid*-PPOPs in each of the contexts defined by each of the other variables in the model, I found no indications of any potential interactions that might

explain why *mid*-PPOPs are significantly less likely than *to*-PPOPs to be specially placed. Evidence of a potential interaction would be apparent if, for example, third person PPOPs were to occur significantly more frequently than first or second person PPOPs in a left-of-P position when governed by *to* but not when governed by *mid*, although such evidence would not, in itself, be definitive. However, interaction effects are not always apparent from bare frequencies in cross-tabulations, so there may be interactions involving PREPOSITION which are simply undetectable in the available sample.

The tendency for PPOPs governed by *mid* to occur to the preposition's right has been noted in several places in this study. In Chapter 4, it was noted that *mid* is the preposition that most frequently governs a PPOP when the PP is the constituent of an NP. As right-of-P placement is invariable in this context, these examples are not included in the data in Table 6.10. Earlier in the present chapter, it was also noted that the tendency to place PPOPs to the right rather than the left of *mid* triumphs over the tendency of translators to conform to Latin word order when translating *cum*-PPs: of the 128 Old English *mid*-PPs matched with a *cum*-PP in the sample, just three (2%) govern a left-of-P PPOP. An example with a right-of-P PPOP was given earlier at (1). An example with a left-of-P PPOP is given at (5).

(5) a. … &   þæra manna dæl   þe **me mid** comon

       and of-the men's portion that me with came

    '… and the portion of the men who came with me'

<div align="right">(cootest,Gen:14.22.546)</div>

   b. … et   partibus uirorum,   qui uenerunt **mecum**

       and portion of-the-men who came   me-with

    '… and the portion of the men who came with me'

<div align="right">(Crawford 1922: 120)</div>

It does not appear to be the case that *mid* is a prosodically deficient preposition, i.e. phonologically dependent on its object as I have shown to be the case with *be* and *for*. In Chapter 1, it was noted that *mid* is by far the most common preposition to occur with a null object, and there are no regular spelling distinctions between

these instances of *mid* and the ones discussed in Chapter 4. It may also be noted that *mid* is stranded just as often (N=124) as it is pied piped (N=125) in relative clauses and that there are around fifty examples of *þærmid* 'therewith, with it' in the YCOE.

Whether the data in Table 6.10 should be interpreted as reflecting lexical effects of the governing preposition on PPOP placement is open to question. On the one hand, the history of English includes many examples of grammatical variation with a lexical dimension. For example, Ellegård (1953: 201) suggests that each verb has its own history in terms of *do*-support, as do Rydén & Brorström (1987) with respect to the transition from BE to HAVE as the perfective auxiliary and as does Tottie (1991: 463) with respect to the change from *no*-negation (as in *he saw no books*) to *not*-negation (as in *he did not see any books*). During the periods that each of these changes were underway, we would expect to find some verbs favouring the older variant and others favouring the innovative variant. Although the YCOE's data show no clear evidence that the frequency of left-of-P placement was changing during the course of Old English (see section 5.6), the different factor weight values in Table 6.10 might nevertheless be indicative of the lexical diffusion of an increasing preference for right-of-P placement of PPOPs over a much greater time-depth.

On the other hand, what appear to be idiosyncratic lexical effects could have a systemic basis. Hoffmann (2005a: 269–70), for example, derives obligatory pied piping in Present Day English when the preposition's object co-refers with a certain group of nouns, e.g. *the way in which*, *the extent to which*, *the time at which*, not from lexical idiosyncrasies of the nouns, but from the syntactic functions of these PPs: they are all sentence adjuncts, either of manner, degree or time. It is possible that what appears to be variation in PPOP placement according to governing preposition in Old English might be also be restricted by PP function, although I suspect that supporting evidence will be extremely difficult to find. Adjuncts are typically distinguished from complements by reference to semantic and syntactic criteria (e.g. Dowty 2003), which usually require native speaker intuitions about, for example, optionality and substitutability. In a corpus such as the YCOE, evidence of relevant contrasts is more likely to be the exception than the rule, and without a full appreciation of the rich semantics of

individual lexical items, the task of investigating the syntactic function of Old English PPs would be an especially daunting one.

For the moment, then, it is not possible to go much beyond the observation that there are some significant differences in the probability of left-of-P placement according to the particular preposition involved. The only aspect of this observation that seems fairly clear is that variation in PPOP placement does not appear to be related to the frequency of individual prepositions.

6.7 PP position

I now turn to the results for the two variables that encode the position of the PP relative to the main verb ('V'). The factor weights for each of these variables are given below.

Table 6.11 Factor weights for LINEAR ORDER OF PP AND V

Factor	PPOPs	Left-of-P	Factor weight
[PP(…)V]	3,012	1,380 (45.8%)	0.620
[V(…)PP]	3,982	1,047 (26.3%)	0.380
Total / Range	6,994	2,427 (34.7%)	0.240

Table 6.12 Factor weights for ADJACENCY OF PP AND V

Factor	PPOPs	Left-of-P	Factor weight
Adjacent	4,843	1,901 (39.3%)	0.562
Non-adjacent	2,151	526 (24.5%)	0.438
Total / Range	6,994	2,427 (34.7%)	0.124

As can be seen from the factor weight ranges, the effect size of LINEAR ORDER OF PP AND V (henceforth 'LINEAR ORDER') is almost exactly twice that of ADJACENCY OF PP AND V (henceforth 'ADJACENCY'). The ranking of variables according to explanatory power in Figure 5.2 confirms that LINEAR ORDER also uniquely accounts for a greater amount of variation than ADJACENCY and, moreover, comes second only to PREPOSITION.

There is, in addition, evidence of a statistically significant interaction between LINEAR ORDER and ADJACENCY. When these two variables are replaced

with a cross-product variable that allows their combined effects to be estimated, the model likelihood shows a highly significant improvement.[95] As all other variables were held constant, this improvement can be directly attributed to an interaction effect between these two variables. The locus of this interaction can be seen from the factor weight values for the categories of the cross-product variable, as shown in Table 6.13. The figures in parentheses in each cell indicate the number of left-of-P PPOPs as a proportion of all PPOPs in that particular context.

Table 6.13 Factor weights for LINEAR ORDER*ADJACENCY (input value = 0.068)

	[PP(…)V]	[V(…)PP]
Adjacent	0.761	0.473
	(1,347/2,695 [50%])	(554/2,148 [26%])
Non-adjacent	0.287	0.464
	(33/317 [10%])	(493/1,834 [27%])

These weight values show that ADJACENCY has a considerable effect when the PP precedes the main verb (0.761 *vs*. 0.287) but a negligible effect when it follows the main verb (0.473 *vs*. 0.464). Indeed, the LRT confirms that left-of-P placement is no more or less likely in [V PP] contexts than in [V...PP] contexts. Overall, these results show that the probability of left-of-P placement increases considerably when the PP is left-adjacent to the main verb and decreases considerably when the PP precedes but is separated from the main verb. When the PP follows the main verb, the probability of left-of-P placement is close to the overall norm, i.e. to the expected probability when all variables in the model are controlled for.

This strong correlation between left-of-P placement and left-adjacency of the PP to the main verb immediately brings to mind a generalisation about preposition stranding in Modern Dutch. As in Old English, PPs do not occupy a fixed position in Dutch: they can optionally scramble from their base, i.e. VP-internal, position into the IP domain as in (6), in which *niet* marks the left edge of the VP; or they can optionally appear in extraposition, as in (7).

---

[95] G^2 (1df) = 108.978, $p < 0.0001$ (see fn. 91).

(6)   ... dat Jan {op deze beslissing} niet {op deze beslissing} had gerekend

        that Jan {on this decision}   not {on this decision}   had counted

        '... that Jan had not counted on this decision'

<div align="right">(Ruys 2008: 547, ex. 1)</div>

(7)   ... dat Jan {op de beslissing} wacht {op de beslissing}

        that Jan {for the decision}   waits   {for the decision}

        '... that Jan waits for the decision'

<div align="right">(Ruys 2008: 547, ex. 7a)</div>

Dutch prepositions can be stranded by scrambling of an R-pronoun object, as in (8), or by *wh*-movement of an R-pronoun, as in (9) (van Riemsdijk 1982: 134–6).

(8)   ... dat Jan {daar / er} vaak op rekent

        that Jan {that / it} often on counts

        '... that Jan often counts on {that / it}'

<div align="right">(Ruys 2008: 549, ex. 5c)</div>

(9) a. Waar rekent Jan op?

        What counts Jan on

        'What does Jan count on?'

<div align="right">(Ruys 2008: 549, ex. 5b)</div>

     b. ... de beslissing waar   Jan op wacht

        the decision   which Jan for waits

        '... the decision which Jan waits for'

<div align="right">(Ruys 2008: 550, ex. 7b)</div>

There are, however, certain constraints against P-stranding in Dutch. Firstly, it is generally impossible with locative and temporal adjunct PPs (Ruys 2008: 549). Secondly, a few prepositions resist stranding altogether (Zwarts 1997), e.g. *ondanks* 'despite'. Thirdly, it has been noted that prepositions cannot be stranded when they occur in a scrambled position, as in (10), or in extraposition, (11).

(10)   ... de beslissing waar   Jan {*op} niet {op} had gerekend

         the decision   which Jan {*on} not {on} had counted

         '... the decision which Jan had not counted on'

<div align="right">(Ruys 2008: 547, ex. 2)</div>

(11)   ... de beslissing waar   Jan {op} wacht {*op}

         the decision   which Jan {for} waits {*for}

         '... the decision which Jan waits for'

<div align="right">(Ruys 2008: 550, ex. 7b)</div>

This third constraint has been described in a number of ways. For Bennis & Hoekstra (1984) and van Riemsdijk & Williams (1986), for example, the relevant observation is that stranding requires the preposition to be left-adjacent to the base position of the main verb. Various types of counter-example, however, show that stranding is still possible when P is left-adjacent to the base position of the verbal *complex*. This is illustrated in (12), where the stranded P is separated from the base position of V by a particle.

(12)   ... de trein waar   ik mee terug ga

         the train which I with back go

         '... the train which I go back with'

<div align="right">(Ruys 2008: 551, ex. 10c)</div>

So, stranding in Dutch seems to be best in general when the PP is left-adjacent to the base position of V, and impossible when the PP follows the base position of V, i.e. when the PP is in extraposition. When the PP occurs to the left of, but is not adjacent to, the base position of V, the possibility of stranding depends on whether the intervening material is VP-internal.

There is a small body of evidence that suggests preposition stranding in Old English — whether by left-of-P placement of a PPOP or by *wh*-movement of the object in a relative clause — is restricted to the same contexts in which preposition stranding is possible in Dutch. However, in each case the evidence is in some way problematic. Firstly, while the data in Table 6.13 show the probability of left-of-P placement increases significantly when the PP is

left-adjacent to the main verb, these data reflect the PPs' position relative to the verb's *surface*, not its *base*, position. Secondly, while Harris (2006: 37–8) found evidence to suggest that right-of-P placement is invariable when the PP is in extraposition, the number of examples is small (N=28) and he shows that the generalisation holds only for non-reflexive pronouns. Thirdly, while van Kemenade (1987: 155) claims that stranded prepositions have 'a regular position: preceding the verb in non-V2 clauses and therefore preceding the verbal trace in V2 clauses' — thereby making an explicit link between stranding and left-adjacency to the base position of V for Old English — her claim is unsupported by quantitative evidence, as is Traugott's (1992: 231) statement that a stranded preposition 'usually precedes the verb'. Lastly, while Pintzuk & Haeberli (2008) show that stranded prepositions are 'overwhelmingly' left-adjacent to the base position of the main verb (*ibid*: 378), they are concerned with identifying elements with a fixed position (which they then use to diagnose underlying clause structure), so the placement of non-stranded Ps is not examined.

In order to find out whether left-of-P placement of PPOPs is significantly more likely when the PP is left-adjacent to the base position of the main verb we need to introduce some sort of control for the main verb's position. Just because a PP immediately precedes the main verb does not necessarily mean it immediately precedes the base position of the main verb; conversely, just because a PP follows the main verb does not necessarily mean it is in extraposition. There are, however, no straightforward means by which main verb position can be fully controlled for. While it is safe to assume that non-finite main verbs occur in clause-final position, data in Pintzuk (1995: 247, Table 3) show very clearly that finite main verbs vary between verb-second and verb-final word orders in main clauses and in subordinate clauses. Consequently, there are only a few contexts in which the finite verb's position can be reliably determined by reference to a single criterion, e.g. verb- and operator-initial clauses, in which the finite verb is definitely not in its base position. So let us instead make the extremely simple but reasonable assumption that finite main verbs are much less likely than non-finite verbs to occur in final position. On this assumption, it follows that PPs that are left-adjacent to the main verb are much less likely to be left-adjacent to the *base position* of the main verb when the main verb is finite than when it is non-finite.

Consequently, if left-of-P placement of PPOPs were significantly more likely when the PP is left-adjacent to the base position of the main verb, we would expect left-of-P placement to be significantly more frequent/likely when the PP is left-adjacent to a non-finite main verb than when left-adjacent to a finite main verb.

Consider, then, the data in Table 6.14. These data expand on the results given for the LINEAR ORDER*ADJACENCY cross product variable in Table 6.13 by providing separate factor weight values for data in [PP V] and [V(...)PP] contexts according to whether the main verb is finite (V_{FIN}) or not (V_{NF}).

Table 6.14 LINEAR ORDER*ADJACENCY*MV FINITENESS (input value = 0.086)

Factor	PPOPs	Left-of-P	Factor weight
[PP V_{FIN}]	1,996	1,095 (54.9%)	0.798
[PP V_{NF}]	699	252 (36.1%)	0.602
[V_{FIN}(...)PP]	3,572	977 (27.4%)	0.489
[V_{NF}(...)PP]	410	70 (17.1%)	0.309
[PP...V]	317	33 (10.4%)	0.281
Total / Range	6,994	2,427 (34.7%)	0.517

These data show, firstly, that PPOPs are significantly more likely to be realised as a special clitic in [PP V_{FIN}] contexts than in [PP V_{NF}] contexts. This is exactly the opposite of what is predicted by the idea that left-of-P placement is most likely to occur when the PP is left-adjacent to the base position of the main verb. Secondly, the behaviour of PPOPs in [V_{NF}(...)PP] contexts suggests that left-of-P placement is not entirely precluded when the PP is in extraposition, at least on the assumption of underlying OV for the majority of the 410 examples.[96] Thirdly — and quite unexpectedly — the weight values show that when the PP precedes the main verb, there is a significant difference in probability of left-of-P placement verb according to whether the verb is finite or not, with finiteness again

---

[96] Pintzuk (1991) found evidence of underlying VO in clauses where VO and OV can be distinguished at a frequency of up to 10% in main clauses (*ibid*: 178, Table 3.6) and up to 3.2% in subordinate clauses (*ibid*: 179, Table 3.7). Data in Koopman (2005: 59, Table 4) suggest the frequency may be considerably higher in later Old English texts in particular but that OV is still more common overall.

correlating more strongly than non-finiteness with PPOP special placement. This points clearly to a surprising correlation between PPOP placement and main verb (non-)finiteness. Further, factor weight ranges indicate that the size of this 'effect' is about the same for PPs that are left-adjacent to the main verb (0.798 – 0.602 = 0.196) as it is for PPs that follow the main verb (0.489 – 0.309 = 0.180).

To test the independence of this apparent correlation, I re-ran the multivariate analysis with MV FINITENESS added as a new and separate variable to the model specified in Appendix G. The addition of this variable produced a significant improvement in model likelihood, indicating that the correlation between main verb (non-)finiteness and PPOP placement is both statistically significant and independent of the effects of each of the other thirteen variables in the model.[97] The factor weight values for this variable also indicate that the presence of a finite MV correlates with a significantly higher probability of left-of-P placement (0.615) while the presence of a non-finite MV correlates with a significantly lower probability (0.385). Moreover, in terms of unique contribution to model likelihood, as represented in Figure 5.2, MV FINITENESS ranks fourth, i.e. between SUBJECT FORM and MAIN VERB, although its inclusion makes no material difference to the importance of each of the other thirteen variables relative to each other, as indicated by relative bar heights in Figure 5.2. The fact that these bar heights change minimally when MV FINITENESS is added indicates there is minimal overlap between the variation in PPOP placement that is uniquely associated with MV FINITENESS on the one hand, and the variation that is uniquely associated with each of the original thirteen variables on the other. In terms of effect size, as represented in Table 5.3, the addition of MV FINITENESS also makes little difference, with one exception: CLAUSE TYPE is demoted from

---

[97] Adding MV FINITENESS as an additional, separate variable to the model shown in Appendix G yields a significant improvement in model likelihood (G^2 (1df) = 96.37, $p < 0.0001$) but no significant improvement in model fit. The input value for this model is 0.083. The other thirteen variables remain statistically significant. Only one change was made to the model shown in Appendix G to accommodate the inclusion of MV FINITENESS. The 281 PPOPs in non-finite constructions were excluded from the analysis of CLAUSE TYPE effects to avoid the risk of 'indeterminacy' (Sigley 2003: 229, fn. 3; also discussed in terms of 'structural zeros' in Paolillo 2002: 69–70, 'distributional interactions' in Sigley 2003: 228–9 and 'non-orthogonality' in Tagliamonte 2006: 182), which is present whenever a particular context is described by more than one variable and which can skew factor weight estimations. The risk of indeterminacy here results from the fact that PPOPs in non-finite constructions naturally co-occur with a non-finite main verb, i.e. all data that belong to either the infinitival or 'other' categories of CLAUSE TYPE also belong to the non-finite category of MV FINITENESS.

fourth (of thirteen) to tenth (of fourteen), i.e. to between NARRATIVE MODE and ADJACENCY OF PP AND V.[98] MV FINITENESS itself ranks sixth with a weight range of 0.230. The impact of MV FINITENESS on the effect size of CLAUSE TYPE indicates that a significant proportion of the variation associated with CLAUSE TYPE in the stepping-up phase of the thirteen-variable model is absorbed by MV FINITENESS in the stepping-up phase of the fourteen-variable model.

The identification of a relationship between PPOP placement and MV FINITENESS is quite unexpected, and is all the more surprising since the correlation is so strong. This correlation has not been identified in any previous study of PPOP placement in Old English (or of P-stranding in West Germanic languages more generally as far as I am aware). It seems highly improbable that main verb finiteness would *directly* influence PPOP placement (why should it?) so it seems much more likely that this apparent effect is a reflex of some as yet unidentified variable (or variables), although I have no intuition about what this (or these) might be. For the majority of clauses containing a PPOP, the only regular difference between those with a finite main verb and those with a non-finite main verb is the absence or presence, respectively, of a finite auxiliary, and by no stretch of the imagination does that make the answer to the problem any clearer.

I can be no more enlightening about why left-of-P placement is more likely when PP is left-adjacent to the main verb regardless of where the verb is situated. The observed correlation fits neatly with Quirk & Wrenn's (1957: §141) claim that left-of-P placement 'is most frequent [...] when it enables the preposition to stand before a verb form', at least descriptively if not also in terms of the motivation they infer for it. This correlation and the one involving MV FINITENESS must regretably be left here as extremely interesting puzzles that would appear to demand some imaginative hypotheses to guide further research.

My closing comments concern an observation in Ogura (1991: 276) and Alcorn (2009: 445) that left-of-P placement is especially frequent when the PP is left-adjacent to the (surface) position of the main verb in *cweðan to* 'to say to' constructions. A comparison of data for this construction compared to all other

---

[98] The factor weight range value for CLAUSE TYPE almost halves to 0.160. The range values for each of the other twelve original variables change minimally, i.e. between +0.01 and -0.03.

data is given at Table 6.15. These results are line with my earlier findings (*ibid*), with both sets of results showing that *cweðan to* PPOPs show highly marked behaviour in comparison to other PPOPs in [PP V] contexts in particular.

Table 6.15 Frequency of left-of-P placement by PP position: (*cweðan to vs.* other)

	[PP...V]	[PP V]	[V(...)PP]	Total
cweðan to	0/4	245/269	292/947	537/1,220
	(0%)	(**91%**)	(31%)	(44%)
other P+V combos.	33/313	1,102/2,426	755/3,035	1,890/5,774
	(11%)	(**45%**)	(25%)	(33%)
Total	33/317	1,347/2,695	1,047/3,982	2,427/6,994
	(10%)	(50%)	(26%)	(35%)

It is certainly the case that in [PP V] contexts, a significantly greater proportion of *cweðan to* PPOPs are associated with a number of factors that strongly favour left-of-P placement, e.g.: 96% of the 269 *cweðan to* PPOPs have third person reference compared to 77% of the 2,426 PPOPs in other constructions; 91% of the *cweðan to* PPOPs occur in a diegetic context compared to 73% of the other PPOPs; and 93% of the *cweðan to* PPOPs co-occur with a finite MV compared to 72% of the other PPOPs. However, similar differences hold between *cweðan to* PPOPs and other PPOPs in [V(...)PP] contexts too, yet the difference in PPOP placement is not nearly so marked.

Without some understanding of the left-adjacency effect on PPOP placement that appears to hold quite generally (see Table 6.14), it is difficult to imagine why the effect is especially strong with *cweðan to* constructions. One possibility is that the high concentration of left-of-P favouring factors associated with this particular construction coupled with the generally promoting effect on left-of-P placement of left-adjacency of PP to MV produce a greater frequency of left-of-P placement than would be expected by reference to the probabilities associated with each factor individually.[99] In other words, perhaps we are witnessing a somewhat formulaic word order pattern.

---

[99] There is a woefully inadequate number of PPOPs to test this hypothesis statistically. In order to estimate the combined effect of PREPOSITION, VERB, PERSON and NARRATIVE MODE, for example, we would need a cross-product variable consisting of all possible combinations of the factors that

6.8 Summary

In this chapter, the results of the regression for seven of the independent variables introduced in Chapter 3 were presented and discussed. For each of these variables, a statistically and linguistically significant correlation with PPOP placement was shown to obtain, although the results for DIALECT indicate that the difference in PPOP placement between the two major dialect categories used in this study is too weak to have any practical application. The results for LATIN INTERFERENCE confirm the findings of Taylor (2008) and extend her findings by showing that the indirect effect is of approximately the same magnitude as the direct effect. The correlation between PPOP placement and PERSON, first noted by Wende (1915: 76–81) was shown to be a particularly strong one, although appears be confined to West Saxon materials, and may possibly go back to the Germanic varieties from which English descends. The effect of PERSON does not straightforwardly extend to object personal pronouns when governed by a verb. Evidence was presented to show that the correlation between PPOP placement and NARRATIVE MODE is too robust to be dismissed as a 'nonsense' variable, and I have suggested it may reflect a difference in register. The results for PREPOSITION indicate that the probability of left-of-P placement is greater with some prepositions than with others, although there is no obvious pattern to the data, either in terms of the prepositions' relative frequency or their semantics. While this correlation appears to be purely lexical, I suggested that future research might be able to detect some or other pattern relating to PP function and/or to the demise of the left-of-P option in Middle English. Lastly, it was shown that left-of-P placement is strongly favoured when the PP occurs immediately before the main verb. Contrary to predictions of accounts of preposition stranding more generally, evidence was produced which shows quite clearly that this left-adjacency effect holds regardless of whether the verb is in its base position or not. Evidence of a further — and quite unexpected — correlation was also encountered, namely a strong effect of verb finiteness. It was shown that this variable captures the same variation in PPOP placement that is associated with CLAUSE TYPE, and some more besides.

---

define these variables. This variable would be defined by 936 factors — being the product of the number of factors that define PREPOSITION (13), VERB (13), PERSON (3) and NARRATIVE MODE (2) — and very few of these factors are likely to be associated with more than one or two PPOPs.

Chapter 7 Conclusions

7.1 The variety of special positions

The thesis has shown that theories which allow simple personal pronouns governed by a preposition to appear as special clitics to the left of P^0 but within PP and to attach to a host at the CP/IP boundary are able to account for the placement of the majority of left-of-P PPOPs. The thesis has additionally shown that at least one other position — somewhere below the host at the CP/IP boundary — must be recognised to fully account for the data on PPOPs. A third clitic position — at the left edge of VP — is provided for in van Kemenade (1987) and in Pintzuk (1991), but this position is later abandoned in Hulk & van Kemenade (1997), who appear to abandon the idea of special clitics altogether in favour of a weak pronoun analysis, and in Pintzuk (1996: 392, fn. 14), who leaves a not-inconsiderable number of examples rather poorly explained. I have argued against a weak pronoun analysis of specially placed PPOPs, primarily on the basis that such an analysis fails to account satisfactorily for the placement of two-thirds of all examples to the immediate left of the preposition regardless of where PP is situated. Van Bergen (2003: 126, fn. 6) suggests the Spec-VP option may yet be rescued if the possibility of scrambling of nominal objects of verbs is allowed for, a possibility for which there is some independent evidence (Koopman 1991: 114–7, Haeberli 1999, Trips 2002: 188–97). I concluded that further work is needed to determine whether recognising an additional host at the IP/VP boundary, e.g. Spec-VP, would be sufficient to fully account for the variety of positions in which clitic PPOPs appear, noting that the range of positions identified for clitic PPOPs must additionally be reconciled with the placement of nominal and pronominal objects of verbs.

Thirdly, the thesis has shown how main clause-initial PPOPs pose a considerable challenge to theories of Old English syntax in general. These data have tended to be overlooked in the theoretical literature, but the number of examples is large enough to raise important questions, such as whether topicalisation is obligatory in main clauses and how the verb second constraint

can be satisfied. I have argued that the empirical evidence strongly suggests that main clause-initial PPOPs are not topics and that they behave no differently than other left-of-P PPOPs. A clitic analysis of these pronouns would explain why there are so few similar examples involving nominal objects of prepositions, but I have left open the question about whether these examples involve: narrative inversion with the PPOP proclitic on the finite verb, as suggested by van Kemenade (1987) although not exactly in these terms; an initial expletive subject topic, as suggested by Axel (2009: 30–3) for verb-initial clauses in Old High German, with the PPOP clitic situated on a host at the CP/IP boundary, i.e. between topic and finite verb — a possibility only in clauses with an unaccusative predicate; or something else entirely.

7.2 Structured variation

Overall, the thesis has shown clear evidence that there is indeed some degree of structure to the variable positioning of PPOPs in Old English prose, although for most of the correlating factors identified the thesis has concluded only that *causa latet, vis est notissima* (the cause is hidden, but the effect is very clear).

I have shown that the variable associated with the largest share of predictably placed PPOPs is pronoun case, with left-of-P placement found to be extremely rare unless the pronoun is dative. Data on pronouns governed by case-alternating prepositions indicate this cannot be interpreted as a lexical effect of the preposition, and a number of minimal pairs suggest it cannot be interpreted as an effect of PP semantics either. The behaviour of PPOPs in the poetry and of PPOVs in the prose add some weight to accepting the correlation between pronoun placement and pronoun case at face value, although genitive personal pronouns are relatively infrequent in general while accusative PPOPs in the poetry and accusative PPOVs in the prose more frequently occur in a special position in comparison to accusative PPOPs in the prose. Quite why dative pronouns should be more frequently realised as special clitics than non-dative object pronouns in general, however, is not at all clear. The thesis has further argued that the correlation between PPOP placement and pronoun case could explain why pronouns governed by *þurh* 'through' are invariably situated to the preposition's right since *þurh* rarely governs dative, although this does not explain why *þurh*,

unlike most other prepositions, strongly prefers pied piping to stranding in relative clauses nor why *þærþurh* does not appear until early Middle English. The thesis has also suggested that the correlation between PPOP placement and pronoun case may have been the starting point for a process by which forms of Old English BETWEEN came to exhibit a remarkable distribution according to object type. An explanation along any other line would be hard pressed to explain why *betweonum* forms are predominantly associated with special clitic objects and why other BETWEEN forms are predominantly associated with simple clitic and non-clitic objects.

Two categorical lexical correlations have been shown to have a phonological basis. I have presented evidence, firstly, that *be* is the unaccented variant of *bi* and that *for* is the unaccented variant of *fore* and, secondly, that *be* and *for* are simple clitics which are phonologically dependent on their object. As clitic hosts cannot themselves be phonologically dependent, it follows that objects of *be* and *for* cannot be clitics, and so the invariable right-of-P placement of objects of these forms in the poetry as well as the prose is correctly predicted. The proposed analysis also explains why *þærbi* and *þærfore* are attested when *þærbe* and *þærfor* are not, and why *bi* and *fore* but not *be* and *for* can be stranded in a relative clause. These findings are particularly relevant to Old English dictionaries, as some identify these phonologically strong~weak variants as separate prepositions.

There is no direct evidence for pronoun prosody in Old English prose, but the thesis has identified two contexts in which the pronoun is very likely to carry contrastive or emphatic stress: where it belongs to a PP that is coordinated with another PP headed by the same preposition and where the preposition expresses. For these data, I have suggested that right-of-P placement follows directly from stress, from which the pronoun's phonological independence may be assumed.

The thesis has shown that the Subjacency condition is able to provide a principled explanation for the invariable right-of-P placement of PPOPs when the PP modifies a noun, but only at the expense of claiming that left-of-P placement always involves movement of the PPOP out of PP which would have serious consequences for otherwise elegant accounts of the placement of two-thirds of special clitic PPOPs. This is a pity since Subjacency would provide a way to

account for the only clearly syntactic factor found to correlate categorically with right-of-P placement of PPOPs in this study.

Other evidence of structured variation presented in the thesis was of a probabilistic rather than categorical nature. Among the extra-linguistic variables examined, there is clear evidence of a direct interference effect of Latin word order on PPOP placement in the general class of Latin translations and of an indirect interference effect in biblical translations in particular. These findings expand slightly on those of Taylor (2008) by showing that the size of the indirect effect is statistically equivalent to the size of the direct effect. There is evidence that left-of-P placement is slightly more common in West Saxon materials than in Anglian-influenced materials, although the difference is too slight for this finding to have any practical application. However, as the highest probability of left-of-P placement is associated with texts not fully classified for dialect (if at all), it is possible that the findings for dialect may not be wholly reliable.

Only one pronoun-related variable other than case shows signs of having a statistical effect on PPOP placement. The thesis has shown that third person PPOPs are significantly more likely than first person and second person PPOPs to be specially placed. Previous univariate analyses show this correlation is remarkably consistent across individual prose texts (Alcorn 2009: 436–7), and the thesis has found evidence that the same effect is manifest in the poetry. Evidence of the same effect on PPOPs in Old Saxon poetry indicates that the person effect is a feature inherited from the West Germanic varieties from which English descends, in which case its presence in West Saxon prose materials but not in Anglian-influenced prose materials may be seen as a conservatism of the former. It is tempting to suppose a connection between the person effect and the case effect: after all, PPOPs are rarely realised as special clitics unless dative, and dative is unambiguous only among third person forms. There is a large body of literature on syntactic consequences of morphological change (e.g. Lightfoot 2002), and van Kemenade (1987: 188–207) in particular has argued that the loss of clitic object pronouns in the history of English was a direct consequence of the loss of morphological distinctions. However, the idea that non-distinct pronoun morphology might lie behind the person effect is difficult to reconcile with the

data in Table 6.5 and Table 6.6, which show no consistent evidence of the person effect on PPOV placement.

Evidence of a statistical effect of the governing preposition for prepositions other than BETWEEN, *þurh*, *be~bi* and *for~fore* is difficult to interpret, but the thesis has shown, firstly, that correlations are unrelated to the prepositions' frequencies and, secondly, that the range of probability values reveal no obvious semantically-defined groupings. Two prepositions, *ongean* and *togeanes*, both 'towards, against', very strongly favour left- over right-of-P placement in the poetry as well as the prose, and a similar trend is evident with their cognates in the Old Saxon *Heliand*. The originally adverbial component of these prepositions sets them apart from many Old English prepositions, but not from *beforan* 'before', which does not show a strong preference for left-of-P PPOPs. One preposition, *mid* 'with', shows a strong preference for right- over left-of-P placement, and a similar trend is once again evident in the Old English and Old Saxon poetry. Given clear evidence of the influence of Latin PP word order, the strong tendency for pronouns to follow governing *mid* is especially remarkable in the Latin translations, in which just 2% of Old English *mid*-PPs matched with a head-final Latin *cum*-PP govern a left-of-P PPOP. I have suggested that the wide range of probability values for individual prepositions might possibly reflect the lexical diffusion of an increasing preference for right-of-P placement over a much greater time-depth than is represented by the YCOE's data. If so, we would expect to find examples of *him mid* to disappear before examples of *him ongean*, for example, in the post-Old English materials.

Of the four remaining statistically significant variables identified, just one readily lends itself to the possibility of a linguistically meaningful interpretation. The thesis has found evidence that left-of-P placement is significantly less likely to occur in direct speech than elsewhere, and univariate analyses show this correlation is sufficiently consistent across individual prose texts to indicate a genuine effect. I have suggested this is more likely to reflect a difference in register than a difference between the spoken and written language, but there are no relevant studies to draw on for evidence to support or falsify this possibility.

Each of the remaining three statistically significant variables identified is nothing short of a puzzle. I can do no more than record my findings in the hope

that they might inspire some imaginative hypotheses to guide further research. Two of these variables make reference to the main verb. The thesis has found evidence that the probability of left-of-P placement increases significantly when the PP is left-adjacent to the main verb and, regardless of where the PP is situated, when the main verb is finite. These findings indicate that, unlike preposition stranding in Dutch — which is best in general when the preposition is left-adjacent to the base position of the main verb — placement of the PP immediately before the main verb significantly increases the probability of PPOP special placement regardless of whether the main verb is in its base position or not. The evidence further suggests that PPOP special placement is possible when the PP is in extraposition. This too is unlike the situation in Dutch, which disallows stranding when the PP is in extraposition.

The third puzzling variable makes reference to the form of the subject. The results for this variable indicate that the probability of left-of-P placement decreases significantly in the presence of a personal pronoun subject. I concluded that the statistical relationship between subject form and PPOP placement is very probably spurious, but that it suggests the presence of a third variable — one that correlates with subject form and with PPOP placement.

Finally, the thesis has shown there is no substantive correlation between PPOP placement on the one hand and pronoun number, pronoun reflexivity or clause type on the other. PPOP placement also appears to vary minimally according to co-occurring main verb apart from with *habban* 'to have', which appears to significantly constrain the probability of left-of-P placement. There is also no evidence that the likelihood of special placement of PPOPs changed significantly over the course of the Old English period apart from those governed by *fram* 'fram' and *to* 'to'. These two sets of data show an increasing preference for left-of-P placement according to one method of measuring time but not according to another, so the evidence for diachronic variation for each set is weak at best. The fact that there is no evidence of a statistically significant *decrease* in left-of-P placement over time strongly suggests that the decline did not begin until after the Old English period. The diachronic evidence in general does not sit comfortably with van Kemenade's suggestion (1987: 193–6) that left-of-P placement is rare after the early 12th century. Either the loss of this option was

remarkably rapid or van Kemenade's small sample (*ibid*: 4) is unrepresentative of PPOP placement in Middle English. The history of special placement of PPOPs beyond the Old English period is a much under-studied topic, with little known about the progress of the regularisation of right-of-P placement. For example, why did left-of-P placement cease to be an option? Was it connected to other contemporaneous word order changes or did it happen independently? How gradually did left-of-P placement disappear? Was it lost sooner in contexts in which left-of-P placement was disfavoured in Old English or did it progress at a steady rate across the board? These and other questions are ripe for novel research.

While the statistical analyses have separated the wheat from the chaff in terms of previously identified variables, the results of the regression show that a significant proportion of the observed variation is unaccounted for by *any* of the variables investigated. It is impossible to say whether this is due to the presence of a degree of free variation or to the presence of other relevant variables that have yet to be identified.

References

Electronic corpora

Pintzuk, Susan & Leendert Plug (eds.). 2001. *The York-Helsinki Parsed Corpus of Old English Poetry*. Department of Language and Linguistic Science, University of York.

Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths (eds.). 2003. *The York–Toronto–Helsinki parsed corpus of Old English prose*. Department of Language and Linguistic Science, University of York. [Corpus documentation is available online at http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm]

Editions consulted (Old English)

coalex: Orchard, Andy (ed.). 1995. The letter of Alexander to Aristotle: Old English text. In Andy Orchard (ed.), *Pride and prodigies: Studies in the monsters of the Beowulf manuscript*, 224–52. Cambridge: D. S. Brewer.

cobede: Miller, Thomas (ed.). 1890–98. *The Old English version of Bede's Ecclesiastical History of the English People*, 4 vols. London: Trübner & Co. [EETS OS 95, 96, 110, 111.]

coblick: Morris, Richard (ed.). 1874–80. *The Blickling Homilies of the tenth century*, 3 vols. [Repr. in 1 vol. 1967.] London: N. Trübner & Co. [EETS OS 58, 63, 73.] Also: Kelly, Richard J. (ed.). 2003. *The Blickling Homilies*. London: Continuum. Also: Willard, Rudolph (1960). The Blickling Homilies. Copenhagen: Rosenkilde & Bagger. [Early English Manuscripts in Facsimile X.]

coboeth: Sedgefield, Walter J. (ed.). 1899. *King Alfred's Old English version of Boethius' De consolatione philosophiae*. Oxford: Clarendon Press.

cocathom2: Godden, Malcolm (ed.). 1979. *Ælfric's Catholic Homilies: The Second Series*. London: Oxford University Press. [EETS SS 5.]

cochronC: Rositzke, Harry A. (ed.). 1940. *The C-text of the Old English Chronicles*. Bochum-Langendreer: Verlag Heinrich Pöppinghaus. [Beiträge zur englischen Philologie 34.]

cocura: Sweet, Henry (ed.). 1871. *King Alfred's West-Saxon version of Gregory's Pastoral Care*. London: Oxford University Press. [EETS OS 45, 50.]

cogregdC: Hecht, Hans (ed.). 1900–7. *Bischof Wærferth von Worcester Übersetzung der Dialoge Gregors des Grossen*. [Repr. Darmstadt 1965.] Leipzig: G. H. Wigand. [Bibliothek der angelsächsischen Prosa 5.]

cogregdH: see cogregdC.

colacnu: Grattan, John H. G. & Charles Singer (eds.). 1952. *Anglo-Saxon magic and medicine*. London. Oxford University Press. [Publications of the Wellcome Historical Medical Museum, n.s. 3.] Also: Cockayne, Oswald (ed.). 1866. *Leechdoms, wortcunning and starcraft of early England*, vol. 3. London: Longmans, Green, Reader and Dyer.

colawaf: Liebermann, Felix (ed.). 1903. *Die Gesetze der Angelsachsen*, vol. 1, *Text und Übersetzung*. Halle: Max Niemeyer.

colsigewZ: see cootest.

comart3: Kotzor, Günter (ed.). 1981. *Das altenglische Martyrologium*, vol. 2, *Edition, Anmerkungen und Indices*. Munich: Verlag der Bayerischen Akademie der Wissenschaften.

Conf 4: Fowler, Roger (ed.). 1965. A late Old English handbook for the use of a confessor. *Anglia* 83, 1–34. [Text not included in YCOE.]

conicodA: Cross, James E. (ed.). 1996. *Two Old English apocrypha and their manuscript source*. Cambridge: Cambridge University Press. [Cambridge studies in Anglo-Saxon England 19.]

coorosiu: Bately, Janet (ed.). 1980. *The Old English Orosius*. London: Oxford University Press. [EETS SS 6.]

cootest: Crawford, Samuel J. (ed.). 1922. *The Old English version of the Heptateuch, Ælfric's treatise on the Old and New Testament and his preface to Genesis*. London: Oxford University Press. [EETS OS 160.]

coverhom: Scragg, Donald G. (ed.) 1992. *The Vercelli Homilies and related texts*. Oxford: Oxford University Press. [EETS OS 300.]

cowulf: Bethurum, Dorothy (ed.). 1957. *The homilies of Wulfstan*. Oxford: Clarendon Press. [Reprint 1998, Sandpiper Books.]

Thorpe, Benjamin (ed.). 1832. *Cædmon's metrical paraphrase of parts of the Scripture in Anglo-Saxon*. London: The Society of Antiquaries of London. [Text not included in YCOE.]

Editions consulted (Old Frisian)

Buma, Wybren Jan. 1949. *Die Brokmer Rechthandschriften*. The Hague: Nijhoff. [Oudfriese Taal- en Rechtsbronnen 5.]

Hoekstra, Jelle (ed.). 1950. *De eerste en de tweede Hunsinger Codex*. The Hague: Nijhoff. [Oudfriese Taal- en Rechtsbronnen 6.]

Sjölin, Bo (ed.). 1970. *Die 'Fivelgoer' handschrift, I. Einleitung und Text*. The Hague: Nijhoff. [Oudfriese Taal- en Rechtsbronnen 12.]

Editions consulted (Latin)

Clayton, Mary & Hugh Magennis (eds.). 1994. *The Old English lives of St. Margaret*. Cambridge: Cambridge University Press.

Colunga, Alberto & Laurentio Turrado (eds.). 1985. *Biblia Sacra: iuxta Vulgatam Clementinam*, nova editio. Matriti: Biblioteca de Autores Cristianos.

Colgrave, Bertram & R. A. B. Mynors (eds.). 1969. *Bede's ecclesiastical history of the English people*. Oxford: Clarendon Press.

Cross, James E. (ed. & transl.). 1996. *Two Old English Apocrypha and their manuscript source*. Cambridge: Cambridge University Press. [Cambridge studies in Anglo-Saxon England 19.]

Secondary sources (except those already listed as primary sources)

Adams, Michael. 1997. Elliptical *with*. *American Speech* 72(2), 220–4.

Akimoto, Minoji & Laurel J. Brinton. 1999. The origin of the composite predicate in Old English. In Laurel J. Brinton & Minoji Akimoto (eds.), *Collocational and idiomatic aspects of composite predicates in the history of English*, 21–58. Amsterdam: John Benjamins.

Alcorn, Rhona. 2007. *The distribution of pronominal objects of prepositions in Old English: An empirical study of the significance of grammatical person*. Diss. M.Sc., University of Edinburgh.

——— 2009. Grammatical person and the variable syntax of Old English personal pronouns. *English Language and Linguistics* 13(3), 433–51.

Algeo, John. 2010. *The origins and development of the English language*, 6th edn. Boston, MA: Wadsworth.

Allen, Cynthia L. 1980. Movement and deletion in Old English. *Linguistic Inquiry* 11(2), 261–323.

——— 1990. Review of van Kemenade, Syntactic case and morphological case in the history of English. *Language* 66(1), 146–52.

——— 1995. *Case marking and reanalysis: Grammatical relations from Old to early Modern English*. Oxford: Oxford University Press.

Andersen, Henning. 1973. Abductive and deductive change. *Language* 49(4), 765–93.

Axel, Katrin. 2009. The verb-second property in Old High German: Different ways of filling the prefield. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change: New approaches to word order variation in Germanic*, 17–44. Berlin: Walter de Gruyter. [Trends in linguistics: Studies and monographs 203.]

Babbie, Earl R. 2010. *The practice of social research*, 12th edn. Belmont, CA: Wadsworth.

Bately, Janet (ed.). 1986. *The Anglo-Saxon Chronicle: A collaborative edition,* vol. 3, MS. A. Cambridge: D. S. Brewer.

Bech, Kristin. 2001. *Word order patterns in Old and Middle English: A syntactic and pragmatic study*. Diss. Ph.D., University of Bergen.

Belden, Henry M. 1897. *The prepositions* in, on, to, for, fore *and* æt *in Anglo-Saxon prose: A study of case-values in Old English*. Diss. Ph.D., John Hopkins University, Baltimore, MD.

Bennis, Hans & Teun Hoekstra. 1984. Gaps and parasitic gaps. *The Linguistic Review* 4, 29–87.

Bergen, Linda van. 2003. *Pronouns and word order in Old English, with particular reference to the indefinite pronoun* man. New York & London: Routledge. [Outstanding dissertations in linguistics.]

——— 2008. Negative contraction and Old English dialects: Evidence from glosses and prose, Part II. *Neuphilologische Mitteilungen*, 109(4), 391–435.

Berndt, Rolf. 1956. *Form und Funktion des Verbum im nördlichen Spätaltenglischen*. Halle: Max Niemeyer.

Bessinger, Jess B. (ed.). 1978. *A concordance to the Anglo-Saxon Poetic Records*. Ithaca, NY: Cornell University Press.

Bhat, Darbhe N. S. 2004. *Pronouns*. Oxford: Oxford University Press. [Oxford studies in typology and linguistic theory.]

Bjork, Robert E. 1985. *The Old English verse saints' lives: A study in direct discourse and the iconography of style*. Toronto: University of Toronto Press.

Blake, Norman F. 1979. *The English language in medieval literature*. London: Methuen & Co. Ltd.

Blockley, Mary. 2001. *Aspects of Old English poetic syntax: Where clauses begin*. Urbana, IL: University of Illinois Press.

Boeckx, Cedric & Kleanthes K. Grohmann. 2004. Left dislocation in Germanic. In Werner Abraham (ed.), *Focus on Germanic typology*, 139–52. Berlin: Akademie Verlag. [Studia typologica 6.]

Bosworth, Joseph & T. Northcote Toller. 1898. *An Anglo-Saxon dictionary*. Oxford: Clarendon Press.

Bremmer, Rolf H. 2009. *An introduction to Old Frisian: History, grammar, reader, glossary*. Amsterdam: John Benjamins.

Campbell, Alistair. 1959. *Old English grammar*. Oxford: Clarendon Press.

Cameron, Angus, Ashley C. Amos, Antonette diPaolo Healey *et al.* (eds.). 2007. *Dictionary of Old English: A to G online* [*DOE*]. Toronto: Dictionary of Old English Project.

Cardinaletti, Anna. 1994. On the internal structure of pronominal DPs. *The Linguistic Review* 11(3), 195–220.

——— 1999. Pronouns in Germanic and Romance languages: An overview. In van Riemsdijk (ed.), 33–82.

Cardinaletti, Anna & Michal Starke. 1996. Deficient pronouns: A view from Germanic. In Höskuldur Thráinsson, Samuel D. Epstein & Steve Peter (eds.), *Studies in comparative Germanic syntax*, vol. II, 21–65. Dordrecht: Kluwer Academic Publishers. [Studies in natural language and linguistic theory 38.]

——— 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In van Riemsdijk (ed.), 145–233.

Carlton, Charles. 1963. Word order of noun modifiers in Old English prose. *The Journal of English and Germanic Philology* 62(4), 778–83.

Carrera-Sabaté, Josefina. 2002. Statistics in the analysis of phonetic variation: Application of the Goldvarb programme. *Noves SL, Journal on Sociolinguistics* Spring. [Accessed 2008, January 19] www6.gencat.net/llengcat/noves/hm02primavera/metodologia/a_carrera.pdf

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Cinque, Guglielmo & Luigi Rizzi (eds.). 2010. *Mapping spatial PPs*. Oxford: Oxford University Press. [Oxford studies in comparative syntax: The cartography of syntactic structures, vol. 6.]

Clark Hall, John R. 1960. *A concise Anglo-Saxon dictionary*, 4th edn. [With a supplement by Herbert D. Meritt.] Toronto, Buffalo, London: University of Toronto Press in association with the Medieval Academy of America.

Colman, Fran. 1991. What positions fit in? In Kastovsky (ed.), 51–102.

Colman, Fran. (ed.). 1992. *Evidence for Old English: Material and theoretical bases for reconstruction*. Edinburgh: John Donald. [Edinburgh studies in the English Language 2.]

Copeland, Rita. 1991. *Rhetoric, hermeneutics and translation in the Middle Ages: Academic traditions and vernacular texts*. Cambridge: Cambridge University Press. [Cambridge studies in medieval literature 11.]

Cross, James E. & Thomas D. Hill (eds.). 1982. *The prose* Solomon and Saturn *and* Adrian and Ritheus. Toronto: University of Toronto Press.

Denison, David. 1981. *Aspects of the history of English group-verbs, with particular attention to the syntax of the Ormulum*. Diss. Ph.D., University of Oxford.

——— 1993. *English historical syntax: Verbal constructions*. London: Longman.

Dikken, Marcel den. 2010. On the functional structure of locative and directional PPs. In Cinque & Rizzi (eds.), 74–126.

diPaolo Healey, Antonette (ed.). 2009. *Dictionary of Old English web corpus*. Available from: http://www.doe.utoronto.ca/

diPaolo Healey, Antonette & Richard L. Venezky (comps.). 1980. *A microfiche concordance to Old English*. Toronto: The Pontifical Institute of Mediaeval Studies. [Publications of the Dictionary of Old English 1.]

Dowty, David. 2003. The dual analysis of adjuncts/complements in Categorial Grammar. In Ewald Lang, Claudia Maienborn & Cathrine Fabricius-Hansen (eds.), *Modifying adjuncts*, 33–66. Berlin: Mouton de Gruyter.

Elenbaas, Marion. 2006. *The synchronic and diachronic syntax of the English verb-particle combination*. Utrecht: LOT. [LOT dissertation series 149.]

Ellegård, Alvar. 1953. *The auxiliary* do: *The establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell. [Gothenburg studies in English 2.]

Eynde, Frank van. 1999. Major and minor pronouns in Dutch. In Gosse Bouma, Erhard W. Hinrichs, Geert-Jan M. Kruij & Richard T. Oehrle (eds.), *Constraints and resources in natural language syntax and semantics*, 135–49. Stanford, CA: CSLI Publications.

Fakundiny, Lydia. 1970. The art of Old English verse composition. *The Review of English Studies*, New Series, 21(82), 129–42.

Fanselow, Gisbert. 2009. Bootstrapping verb movement and the clausal architecture of German (and other languages). In Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nuger & Florian Schäfer (eds.), *Advances in comparative Germanic syntax*, 85–118. Amsterdam: John Benjamins. [Linguistik Aktuell 141.]

Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2000. *The syntax of Early English*. Cambridge: Cambridge University Press.

Gehrke, Berit. 2008. *Ps in motion: On the syntax and semantics of P elements in motion events*. Utrecht: LOT. [LOT dissertation series 184.]

Gelderen, Elly van. 2000. *A history of English reflexive pronouns: Person,* self, *and interpretability*. Amsterdam: John Benjamins. [Linguistik Aktuell 39.]

Gneuss, Helmut. 1991. The Old English language. In Malcolm Godden & Michael Lapidge (eds.), *The Cambridge companion to Old English literature*, 23–54. Cambridge: Cambridge University Press.

Godden, Malcolm R. 2003. The Alfredian Boethius project. *Old English Newsletter* 37(1), 26–34.

Gold, David. 1970. Statistical tests and substantive significance. In Denton E. Morrison & Ramon E. Henkel (eds.), *The significance test controversy: A reader*, 172–81. London: Butterworths.

Gorman, Kyle. (to appear). Hierarchical regression and the stratification of (neg). *University of Pennsylvania Working Papers in Linguistics* 16(2): *Selected papers from NWAV 38.*

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins. [Linguistik Aktuell 66.]

Gupta, Sunil. 2008. *Understanding the essentials of multivariable testing*: *Counting effects*. San Jose, CA: Interwoven Inc. http://publications.autonomy.com/pdfs/Promote/White%20Papers/ wp_optimost_counting_effects.pdf (14 Nov. 2009).

Guy, Gregory R. 1988. Advanced Varbrul analysis. In Kathleen Ferrara, Becky Brown, Keith Walters & John Baugh (eds.), *Linguistic change & contact: Proceedings of the sixteenth annual conference on New Ways of Analysing Variation*, 124–36. Austin, TX: University of Texas.

Haeberli, Eric. 1999. *Features, categories and the syntax of A-positions: Synchronic and diachronic variation in the Germanic languages*. Diss. Ph.D., University of Geneva.

——— 2002. Observations on the loss of verb second in the history of English. In C. Jan-Wouter Zwart & Werner Abraham (eds.), *Studies in comparative Germanic syntax: Proceedings from the 15th workshop on comparative Germanic syntax*, 245–72. Amsterdam: John Benjamins. [Linguistik Aktuell 53.]

Haegeman, Liliane. 1999. Cardinaletti and Starke *The typology of structural deficiency*: Empirical support from West Flemish. In van Riemsdijk (ed.), 257–62.

Haider, Hubert. 1990. Topicalisation and other puzzles of German syntax. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers*, 93–112. Amsterdam: John Benjamins. [Linguistik Aktuell 5.]

Halliday, Michael A. K., Angus McIntosh & Peter Strevens. 1964. *The linguistic sciences and language teaching*. London: Longmans.

Harris, Christopher D. 2006. *On the distribution of personal pronouns in the syntax of Old English*. Diss. M.A., University of York.

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Heltveit, Trygve. 1958. Attribute and anaphora in the gender-system of English. *Norsk Tidsskrift for Sprogvidenskap* 18, 357–69.

Hinrichs, Lars & Benedikt Szmrecsanyi. 2007. Recent changes in function and frequency of Standard English genitives. *English Language and Linguistics* 11(3), 437–74.

Hoffmann, Thomas. 2005a. Variable *vs.* categorical effects: Preposition pied piping and stranding in British English relative clauses. *Journal of English Linguistics* 33(3), 257–97.

——— 2005b. Review of John C. Paolillo, Analysing linguistic variation: Statistical models and methods, 2002. *Journal of Sociolinguistics* 9(2), 293–8.

Hogg, Richard M. (ed.). 1992. *The Cambridge history of the English language*, vol. I. *The beginnings to 1066*. Cambridge: Cambridge University Press.

Hogg, Richard M. 2002. *An introduction to Old English*. Edinburgh: Edinburgh University Press.

Hopper, Paul. 1975. *The syntax of the simple sentence in Proto-Germanic*. Den Haag: Mouton de Gruyter.

Hough, Carole A. 1991. *Freo man*: A note on word division. *Neophilologus* 75(4), 641–3.

Huddleston, Rodney & Geoffrey K. Pullum (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Huddleston, Rodney, Geoffrey K. Pullum & Peter Peterson. 2002. Relative constructions and unbounded dependencies. In Huddleston & Pullum (eds.), 1031–96.

Hulk, Aafke & Ans van Kemenade. 1997. Negation as a reflex of clause structure. In Danielle Forget, Paul Hirschbühler, France Martineau & María-Luisa Rivero (eds.), *Negation and polarity: Syntax and Semantics*, 183–207. Amsterdam: John Benjamins. [Current issues in linguistic theory 155.]

Jamison, Stephanie W. 1991. The syntax of direct speech in Vedic. In Hans H. Hock (ed.), *Studies in Sanskrit syntax: A volume in honor of the Centennial of Speijer's Sanskrit syntax*, 95–112. Delhi: Motilal Banarsidass Publishers PVT Ltd.

Johansson, Christine & Christer Geisler. 1998. Pied piping in spoken English. In Antoinette Renouf (ed.), *Explorations in corpus linguistics*, 67–82. Amsterdam: Rodopi.

Johnson, Daniel E. 2009. *Rbrul manual*. http://www.ling.upenn.edu/~johnson4/ Rbrul_manual.html (19 Jun. 2008).

Kahlas-Tarkka, Leena, Matti Kilpiö & Aune Österman. 1993. Old English. In Matti Rissanen, Merja Kytö & Minna Palander-Collin (eds.), *Early English in the computer age: Explorations through the Helsinki Corpus*, 21–32. Berlin: Mouton de Gruyter.

Kastovsky, Dieter (ed.). 1991. *Historical English syntax*. Berlin: Mouton de Gruyter.

Kemenade, Ans van. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.

———— 1997a. Negative-initial sentences in Old and Middle English. *Studia Anglica Posnaniensia* 31, 91–104.

———— 1997b. V2 and embedded topicalisation in Old and Middle English. In van Kemenade & Vincent (eds.), 326–52.

———— 2002. Word order in Old English prose and poetry: The position of finite verbs and adverbs. In Donka Minkova & Robert P. Stockwell (eds.), *Studies in the history of the English language: A millennial perspective*, 355–72. Berlin: Mouton de Gruyter. [Topics in English linguistics 39.]

Kemenade, Ans van & Nigel Vincent (eds.). 1997. *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press.

Ker, Neil R. 1957. *Catalogue of manuscripts containing Anglo-Saxon*. Oxford: Clarendon Press.

Kitson, Peter. 1993. Geographic variation in Old English prepositions and the location of Ælfric's and other literary dialects. *English Studies* 74(1), 1–50.

———— 1996. The dialect position of the Old English *Orosius*. *Studia Anglica Posnaniensia* 30, 3–35.

Klavans, Judith L. 1985. The independence of syntax and phonology in cliticisation. *Language* 61(1), 95–120.

Kohonen, Viljo. 1978. *On the development of English word order in religious prose around 1000 and 1200 AD: A quantitative study of word order in context*. Åbo: Publications of the research institure of the Åbo Akademi Foundation.

Koopman, Hilda. 2010. Prepositions, postpositions, circumpositions and particles. In Cinque & Rizzi (eds.), 26–73.

Koopman, Willem F. 1991. The order of dative and accusative objects in Old English and scrambling. *Studia Anglica Posnaniensia* 25, 109–21.

———— 1992. Old English clitic pronouns: Some remarks. In Colman (ed.), 44–87.

———— 1995. Verb-final main clauses in Old English prose. *Studia Neophilologica* 67, 129–44.

———— 1997a. Another look at clitics in Old English. *Transactions of the Philological Society* 95(1), 73–93.

———— 1997b. Topicalisation in Old English and its effects: Some remarks. In Stanislaw Puppel & Raymond Hickey (eds.), *Language history and language modeling: A Festschrift for Jacek Fisiak on his 60th birthday*, vol. I: *Language history*, 307–21. Berlin: Mouton de Gruyter.

———— 2005. Transitional syntax: Postverbal pronouns and particles in Old English. *English Language and Linguistics* 9(1), 47–62.

Kroch, Anthony & Ann Taylor. 1997. The syntax of verb movement in Middle English: Dialect variation and language contact. In van Kemenade & Vincent (eds.), 297–325.

Kroch, Anthony, Ann Taylor & Donald Ringe. 2000. The Middle English verb-second constraint: A case study in language contact and language change. In Susan C. Herring, Pieter van Reenen & Lene Schøsler (eds.), *Textual parameters in older languages,* 353–91. Amsterdam: John Benjamins.

Kuhn, Hans. 1933. Zur Wortstellung und -betonung im Altgermanischen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 57, 1–109.

Kytö, Merja & Matti Rissanen. 1992. A language in transition: The Helsinki Corpus of English Texts. *Computers in English Linguistics* 16, 7–27.

Langacker, Ronald W. 1977. Syntactic reanalysis. In Li (ed.), 57–139.

Lapidge, Michael. 2006. An aspect of Old English poetic diction: The postpositioning of prepositions. In Walmsley (ed.), 153–80.

Lehmann, Winfred P. 1992. *Historical linguistics: An introduction*, 3rd edn. London: Routledge.

Li, Charles N. (ed.). 1977. *Mechanisms of syntactic change*. Austin, TX: University of Texas Press.

Lightfoot, David W. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David W. (ed.). 2002. *Syntactic effects of morphological change*. Oxford: Oxford University Press.

Liuzza, Roy M. (ed.). 2000. *The Old English version of the Gospels*, vol. 2, *Notes and glossary*. Oxford: Oxford University Press. [EETS OS 314.]

Los, Bettelou. 2000. *Onginnan/beginnan* with bare and *to*-infinitive in Ælfric. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 251–74. Amsterdam: John Benjamins.

Lowry, Richard. 2010. *VassarStats*. http:///faculty.vassar.edu/lowry//odds2x2.html

Lyons, John. 1977. *Semantics*, vol. 2. Cambridge: Cambridge University Press.

Massam, Diane & Yves Roberge. 1989. Recipe context null objects in English. *Linguistic Inquiry* 20(1), 134–9.

Mauder, Elisabeth. 2008. *Pronoun variation in Latin American Spanish: A data engineer's view on* le *and* lo. Utrecht: LOT. [LOT dissertation series 191.]

McMahon, April M. S. 1994. *Understanding language change*. Cambridge: Cambridge University Press.

Minkoff, Harvey. 1976. Some stylistic consequences of Ælfric's theory of translation. *Studies in Philology* 73(1), 29–41.

——— 1977. An example of Latin influence on Ælfric's translation style. *Neophilologus* 61(1), 127–42.

Miranda-García, Antonio & Javier Calle-Martín. 2010. Post-adpositions in Old English. *English Studies* 91(1), 89–111.

Mitchell, Bruce. 1968. Two syntactical notes on *Beowulf*. *Neophilologus* 52(1), 292–9.

——— 1978. Prepositions, adverbs, prepositional adverbs, postpositions, separable prefixes, or inseparable prefixes, in Old English? *Neuphilologische Mitteilungen* 79, 240–57.

————— 1985. *Old English syntax*, 2 vols. Oxford: Clarendon Press.

Mitchell, Bruce & Fred C. Robinson. 1992. *A guide to Old English*, 5th edn. Oxford: Blackwell.

Morgan, Davita J. 2004. *Objects in Old English: Scrambled, shifted or* in situ? Diss. Ph.D., University of York.

Napier, Arthur S. (ed.). 1894. *History of the Holy Rood Tree*. London: Kegan Paul, Trench, Trübner & Co. [EETS OS 103.]

Ogura, Michiko. 1991. *Cweðan + to + dative of person. Neophilologus* 75, 270–8.

————— 1992. Why is the element order *to cwæð him* 'said to him' impossible? In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen & Irma Taavitsainen (eds.), *History of Englishes: New methods and interpretations in historical linguistics*, 373–8. Berlin & New York: Mouton de Gruyter. [Topics in English linguistics 10.]

Oxford English Dictionary [*OED*], 2nd edn. 1989. Oxford: Oxford University Press.

Paolillo, John C. 2002. *Analysing linguistic variation: Statistical models and methods*. Stanford, CA: CSLI Publications.

Patrick, Millar. 1949. *Four centuries of Scottish psalmody*. London: Oxford University Press.

Pintzuk, Susan. 1991. *Phrase structures in competition: Variation and change in Old English word order*. Diss. Ph.D., University of Pennsylvania.

————— 1993. Verb seconding in Old English: Verb movement to Infl. *The Linguistic Review* 10, 5–35.

————— 1995. Variation and change in Old English clause structure. *Language Variation and Change* 7, 229–60.

————— 1996. Cliticisation in Old English. In Aaron L. Halpern & Arnold M. Zwicky (eds.), *Approaching second: Second position clitics and related phenomena*, 375–409. Stanford, CA: CSLI Publications.

————— 2002a. Word order in Old English poetry as a window on the prose. Paper presented at the First York-Holland Symposium on the History of English Syntax, York, 20–21 April 2002.

————— 2002b. Verb-object order in Old English: Variation as grammatical competition. In David W. Lightfoot (ed.), *Syntactic effects of morphological change*, 276–99. Oxford: Oxford University Press.

————— 2005. Arguments against a universal base: Evidence from Old English. *English Language and Linguistics* 9(1), 115–38.

Pintzuk, Susan & Eric Haeberli. 2008. Structural variation in Old English root clauses. *Language Variation and Change* 20(3), 367–407.

Pintzuk, Susan & Anthony Kroch. 1985. Reconciling an exceptional feature of Old English clause structure. In Jan Terje Faarlund (ed.), *Germanic Linguistics: Papers from a symposium at the University of Chicago, April 24, 1985*, 87–111. Bloomington, IN: Indiana University Linguistics Club.

Pintzuk, Susan & Ann Taylor. 2006. The loss of OV order in the history of English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 249–78. Oxford: Blackwell Publishing.

Porte, Graeme K. 2010. *Appraising research in second language learning: A practical approach to critical analysis of quantitative research*, 2ⁿᵈ edn. Amsterdam: John Benjamins.

Quirk, Randolph & Charles L. Wrenn. 1957. *An Old English grammar*, 2ⁿᵈ edn. London: Methuen.

Raith, Josef. 1952. Ælfric's share in the Old English Pentateuch. *The Review of English Studies*, New Series, 3(12), 305–14.

Randall, Beth. 2005. *CorpusSearch 2: A tool for linguistics research*. http://corpussearch.sourceforge.net/

Riemsdijk, Henk van. 1982. *A case study in syntactic markedness: The binding nature of prepositional phrases*, 2ⁿᵈ edn. Dordrecht: Foris Publications.

Riemsdijk, Henk van (ed.). 1999. *Clitics in the languages of Europe*. Berlin & New York: Mouton de Gruyter. [Empirical approaches to language typology 20–5.]

Riemsdijk, Henk van & Edwin Williams. 1986. *Introduction to the theory of grammar*. Cambridge, MA: MIT Press.

Rissanen, Matti. 2006. Latin influence on an Old English idiom: "*To wit*". In Walmsley (ed.), 222–41.

Roberts, Ian & Anna Roussou. 2002. The Extended Projection Principle as a condition on the tense dependency. In Peter Svenonius (ed.), *Subjects, expletives and the EPP*, 125–56. Oxford: Oxford University Press. [Oxford studies in comparative syntax.]

Rosenbach, Anette. 2005. Animacy versus weight as determinants of grammatical variation in English. *Language* 81(3), 613–44.

Russell, Jeffrey B. 1984. *Lucifer: The Devil in the Middle Ages*. Ithaca, NY: Cornell University Press.

Ruys, Eddy G. 2008. Stranding, weak pronouns, and the fine structure of the Dutch Mittelfeld. *Natural Language & Linguistic Theory* 26(3), 547–87.

Rydén, Mats & Sverker Brorström. 1987. *The* be/have *variation with instransitives in English: With special reference to the late Modern period*. Stockholm: Almqvist & Wiksell International. [Stockholm studies in English 70.]

Sankoff, David, Sali A. Tagliamonte & Eric Smith. 2005. *Goldvarb X: A multivariate analysis application*. Toronto: Department of Linguistics, University of Toronto.

Sawyer, Peter H. 1968. *Anglo-Saxon charters: An annotated list and bibliography*. London: Royal Historical Society.

Sievers, Eduard. 1883. Miscellen zur angelsächsischen Grammatik. *Beiträge zur Geschichte der deutschen Sprache und Literatur* IX, 197–300.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Sigley, Robert. 1997. *Choosing your relatives: Relative clauses in New Zealand English*. Diss. Ph.D., University of Wellington.

——— 2003. The importance of interaction effects. *Language Variation and Change* 15(2), 227–53.

Sinkovich, Kathryn A. 1984. *Intermediate college Latin*. Lanham, MD: University Press of America.

Sutter, Gert de. 2009. Towards a multivariate model of grammar: The case of word order variation in Dutch clause final verb clusters. In Andreas Dufter, Jürg Fleischer & Guido Seiler (eds.), *Describing and modelling variation in grammar*, 225–54. Berlin: Mouton de Gruyter. [Trends in linguistics: Studies and monographs 204.]

Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.

Takami, Ken-ichi. 1992. *Preposition stranding: From syntactic to functional analyses*. Berlin: Walter de Gruyter. [Topics in English linguistics 7.]

Taylor, Ann. 2003. *The York-Toronto-Helsinki parsed corpus of Old English Prose: Syntactic annotation reference manual*. York: University of York. Available from: http://www-users.york.ac.uk/~lang22/YCOE/doc/annotation/YcoeRefToc.htm

——— 2008. Contact effects of translation: Distinguishing two types of influence in Old English. *Language Variation and Change* 20(2), 341–65.

Thompson, Bruce. 1999. Statistical significance tests, effect size reporting and the vain pursuit of pseudo-objectivity. *Theory & Psychology* 9(2), 191–6.

Timberlake, Alan. 1977. Reanalysis and actualisation in syntactic change. In Li (ed.), 141–77.

Toebosch, Annemarie. 2003. *Gender-animacy and the morpho-syntax of object clitics in Dutch*. Diss. Ph.D., University of Michigan-Flint.

Toller, T. Northcote. 1921. *An Anglo-Saxon dictionary: Supplement*. Oxford: Clarendon Press.

Toon, Thomas E. 1992. Old English dialects. In Hogg (ed.), 409–51.

Tottie, Gunnel. 1991. Lexical diffusion in syntactic change: Frequency as a determinant of linguistic conservatism in the development of negation in English. In Kastovsky (ed.), 439–67.

Traugott, Elizabeth C. 1992. Syntax. In Hogg (ed.), 168–289.

Trips, Carola. 2002. *From OV to VO in Early Middle English*. Amsterdam: John Benjamins. [Linguistik Aktuell 60.]

Upton, Graham & Ian Cook. 2008. *Oxford dictionary of statistics*, 2nd edn. Oxford: Oxford University Press.

Vat, Jan. 1978. On footnote 2: Evidence for the pronominal status of *þær* in Old English relatives. *Linguistic Inquiry* 9(4), 695–716.

Visser, Frederic Th. 1970. *An historical syntax of the English language*, vol. I: *Syntactic units with one verb*, 2nd edn. Leiden: E. J. Brill.

Visser, Mario van de. 2002. On *er*: Distribution of multiple R-words in Dutch sentences. In Hans Broekhuis & Paula Fikkert (eds.), *Linguistics in the Netherlands 2002*, 175–85. Amsterdam: John Benjamins.

Walmsley, John (ed.). 2006. *Inside Old English: Essays in honour of Bruce Mitchell*. Oxford: Blackwell.

Wanner, Anja. 2009. *Deconstructing the English passive*. Berlin: Walter de Gruyter. [Topics in English linguistics 41.]

Ward, Gregory, Betty Birner & Rodney Huddleston. 2002. Information packaging. In Huddleston & Pullum (eds.), 1363–447.

Waterhouse, Ruth. 1976. Ælfric's use of discourse in some saints' lives. *Anglo-Saxon England* 5, 83–103.

Wende, Fritz. 1915. *Über die nachgestellten Präpositionen im Angelsächsischen*. Berlin: Mayer & Müller. [Palaestra 70.]

Weerman, Fred. 1988. *The V2 conspiracy: A synchronic and a diachronic analysis of verbal positions in Germanic languages*. Dordrecht: Foris. [Publications in language sciences 31.]

Wülfing, J. Ernst. 1901. *Die Syntax in den Werken Alfreds des Grossen*, vol. II. Bonn: P. Hanstein.

Wurff, Wim van der. 1987. Adjective plus infinitive in Old English. In Frits Beukema & Peter Coopmans (eds.), *Linguistics in the Netherlands 1987*, 233–42. Dordrecht: Foris.

——— 1990. The *easy-to-please* construction in Old and Middle English. In Sylvia Adamson, Vivien A. Law, Nigel Vincent & Susan Wright (eds.), *Papers from the 5th international conference on English historical linguistics*, 519–36. Amsterdam: John Benjamins. [Current issues in linguistic theory 65.]

Yerkes, David. 1982. *Syntax and style in Old English: A comparison of the two versions of Wærferth's translation of Gregory's Dialogues*. Binghampton, NY: Centre for Medieval and Early Renaissance Studies. [Medieval and Early Renaissance Studies 5.]

Young, Robert & Robert Bayley. 1996. Varbrul analysis for second language acquisition research. In Robert Bayley & Dennis R. Preston (eds.), *Second language acquisition and linguistic variation*, 253–306. Amsterdam: John Benjamins.

Ziliak, Stephen T. & Deirdre N. McCloskey. 2004. Size matters: the standard error of regressions in The American Economic Review. *The Journal of Socio-Economics* 33(5), 527–46.

Zwart, Jan-Wouter. 2005. Continental West-Germanic languages. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford handbook of comparative syntax*, 903–46. Oxford: Oxford University Press.

Zwarts, Joost. 1997. Complex prepositions and P-stranding in Dutch. *Linguistics* 35(6), 1091–112.

Zwicky, Arnold M. 1977. *On clitics*. Bloomington, IN: Indiana University Linguistics Club.

Appendix A   List of tables

Appendix B   List of figures

Appendix C Search queries for the sample of position C clitic objects

PPOVpositionC1.q (source file: ycoe\psd*)

This query identifies all complete subordinate clauses in the YCOE which match the following criteria: (i) the clause is a *that*-clause; (ii) there is an overt nominative argument of any form other than a simple personal pronoun; and (iii) there is an overt simple case-unambiguous third person personal pronoun object that is not parsed as a PP-internal PPOP (PPOPs situated immediately to the left or right of their governor are parsed as PP-internal in the YCOE).

```
node: CP-THT*
query: (CP-THT* doms [1]IP-SUB*)
AND ([1]IP-SUB* idoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
```

PPOVpositionC2.q (source file: PPOVpositionC1.out)

Eliminates clauses in which the object pronoun is parsed as a PP-external PPOP.

```
node: CP-THT*
print_complement: t
query: (CP-THT* doms [1]IP-SUB*)
AND ([1]IP-SUB* idoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
AND ([1]NP-NOM hasSister [1]PP*)
AND ([1]PP* idoms [2]P)
AND ([2]P hasSister [4]NP*)
AND ([4]NP* iDomsOnly [5]\*ICH*)
AND ([5]\*ICH* sameIndex [2]NP*DAT*|NP*ACC*|NP*GEN*)
```

PPOVpositionC3.q (source file: PPOVpositionC2.cmp)

Eliminates clauses in which the object pronoun is tagged as an adjunct.

```
node: CP-THT*
print_complement: t
query: (CP-THT* doms [1]IP-SUB*)
AND ([1]IP-SUB* idoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND ([1]NP-NOM hasSister [2]NP*ADT*)
AND ([2]NP*ADT* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
```

PPOVpositionC4.q (source file: PPOVpositionC3.cmp)

Identifies clauses with [COMP PPOV Subject ...] word order.

```
node: CP-THT*
print_complement: t
query: (CP-THT* doms [1]IP-SUB*)
AND ([1]IP-SUB* idoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iPrecedes [1]NP-NOM)
AND ([1]IP-SUB* iDomsFirst [2]NP*DAT*|NP*ACC*|NP*GEN*)
```

PPOVpositionC5.q (source file: PPOVpositionC4.cmp)

Identifies clauses with [COMP Subject (…) PPOV (...)] word order.

```
node: CP-THT*
print_complement: t
query: (CP-THT* doms [1]IP-SUB*)
AND ([1]IP-SUB* iDomsFirst [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
```

Data in PPOVpositionC4.out and PPOVpositionC5.out can be sorted if required, e.g. according to features of the object pronoun (e.g. by case or reflexivity), according to subject form (e.g. *man vs*. other) or according to features of the clause (e.g. by narrative mode). Data for first and second person PPOVs were obtained in exactly the same way, replacing only: "`[3]PRO^* idoms h*|H*|$h*|$H`" with "`[3]PRO^* idoms m*|u*|e*|+t*|+d*|i*|v*|y*| M*|U*|E*|+T*|+D*|I*|V*|Y*|$m*|$u*|$e*|$+t*|$+d*|$i*|$v*|$y*| $M*|$U*|$E*|$+T*|$+D*|$I*|$V|$Y*`".

Appendix D Search queries for the sample of position B clitic objects

PPOVpositionB1.q (source file: ycoe\psd*)

This query identifies all complete main clauses in the YCOE which match the following criteria: (i) all clause material is parsed at IP-level or below; (ii) the first constituent (or the first following a sentential conjunction and/or verbal negator) is a temporal adverb or a finite verb; (iii) there is an overt nominative subject of any form other than a simple personal pronoun; (iv) the subject follows the finite verb; and (v) there is an overt simple case-unambiguous third person personal pronoun object that is not parsed as a PP-internal PPOP (PPOPs situated immediately to the right or left of their governor are parsed as PP-internal in the YCOE).

```
node: IP-MAT*
ignore_nodes: CONJ|NEG
query: ([1]IP-MAT* iDomsFirst *VBP*|*VBD*|*VBI*|*BEP*|*BED*|
  *BEI*|*HVP*|*HVD*|*HVI*|*MDP*|*MDD*|*MDI*|*AXP*|*AXD*|
  *AXI*|ADVP-TMP)
AND ([1]IP-MAT* iDoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND (*VBP*|*VBD*|*VBI*|*BEP*|*BED*|*BEI*|*HVP*|*HVD*|*HVI*|
  *MDP*|*MDD*|*MDI*|*AXP*|*AXD*|*AXI* precedes [1]NP-NOM)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
```

PPOVpositionB2.q (source file: PPOVpositionB1.out)

Eliminates clauses with an initial temporal adverb that is neither *þa* nor *þonne*.

```
node: IP-MAT*
print_complement: t
ignore_nodes: CONJ|NEG
query: ([1]IP-MAT* iDomsFirst [1]ADVP-TMP)
AND ([1]ADVP-TMP idoms [2]ADV^T)
AND ([2]ADV^T idoms !+ta|+tonne|+Ta|+Tonne|$+ta|$+tonne|
  $+Ta| $+Tonne|+da|+donne|+Da|+Donne|$+da|$+donne|$+Da|
  $+Donne)
AND ([1]IP-MAT* iDoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND (*VBP*|*VBD*|*VBI*|*BEP*|*BED*|*BEI*|*HVP*|*HVD*|*HVI*|
  *MDP*|*MDD*|*MDI*|*AXP*|*AXD*|*AXI* precedes [1]NP-NOM)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
```

PPOVpositionB3.q (source file: PPOVpositionB2.cmp)

Eliminates clauses in which the object pronoun is parsed as a PP-external PPOP.

```
node: IP-MAT*
print_complement: t
ignore_nodes: CONJ|NEG
query: ([1]IP-MAT* iDomsFirst *VBP*|*VBD*|*VBI*|*BEP*|*BED*|
   *BEI*|*HVP*|*HVD*|*HVI*|*MDP*|*MDD*|*MDI*|*AXP*|*AXD*|
   *AXI*|ADVP-TMP)
AND ([1]IP-MAT* iDoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND (*VBP*|*VBD*|*VBI*|*BEP*|*BED*|*BEI*|*HVP*|*HVD*|*HVI*|
   *MDP*|*MDD*|*MDI*|*AXP*|*AXD*|*AXI* precedes [1]NP-NOM)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
AND ([1]NP-NOM hasSister [1]PP*)
AND ([1]PP* idoms [2]P)
AND ([2]P hasSister [4]NP*)
AND ([4]NP* iDomsOnly [5]\*ICH*)
AND ([5]\*ICH* sameIndex [2]NP*DAT*|NP*ACC*|NP*GEN*)
```

PPOVpositionB4.q (source file: PPOVpositionB3.cmp)

Eliminates clauses in which the object pronoun is tagged as an adjunct.

```
node: IP-MAT*
print_complement: t
ignore_nodes: CONJ|NEG
query: ([1]IP-MAT* iDomsFirst *VBP*|*VBD*|*VBI*|*BEP*|*BED*|
   *BEI*|*HVP*|*HVD*|*HVI*|*MDP*|*MDD*|*MDI*|*AXP*|*AXD*|*AX
   I*|ADVP-TMP)
AND ([1]IP-MAT* iDoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
AND (*VBP*|*VBD*|*VBI*|*BEP*|*BED*|*BEI*|*HVP*|*HVD*|*HVI*|
   *MDP*|*MDD*|*MDI*|*AXP*|*AXD*|*AXI* precedes [1]NP-NOM)
AND ([1]NP-NOM hasSister [2]NP*ADT*)
AND ([2]NP*ADT* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
```

PPOVpositionB5.q (source file: PPOVpositionB4.cmp)

Identifies clauses with [(*þa/þonne*) V$_{FIN}$ PPOV Subject (...)] word order.

```
node: IP-MAT*
print_complement: t
ignore_nodes: CONJ|NEG
query: ([1]IP-MAT* iDomsFirst *VBP*|*VBD*|*VBI*|*BEP*|*BED*|
   *BEI*|*HVP*|*HVD*|*HVI*|*MDP*|*MDD*|*MDI*|*AXP*|*AXD*|
   *AXI*|ADVP-TMP)
AND ([1]IP-MAT* iDoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN*|PRO$*|ADJ*|Q*|F*)
```

```
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
AND (*VBP*|*VBD*|*VBI*|*BEP*|*BED*|*BEI*|*HVP*|*HVD*|*HVI*|
   *MDP*|*MDD*|*MDI*|*AXP*|*AXD*|*AXI* iPrecedes [2]NP*DAT*|
   NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iPrecedes [1]NP-NOM)
```

PPOVpositionB6.q (source file: PPOVpositionB5.cmp)

Identifies clauses with [(*þa/þonne*) V_{FIN} Subject (…) PPOV (...)] word order.

```
node: IP-MAT*
print_complement: t
ignore_nodes: CONJ|NEG
query: ([1]IP-MAT* iDomsFirst *VBP*|*VBD*|*VBI*|*BEP*|*BED*|
   *BEI*|*HVP*|*HVD*|*HVI*|*MDP*|*MDD*|*MDI*|*AXP*|*AXD*|
   *AXI*|ADVP-TMP)
AND ([1]IP-MAT* iDoms [1]NP-NOM)
AND ([1]NP-NOM idoms N*|D*|MAN|PRO$*|ADJ*|Q*|F*)
AND ([1]NP-NOM hasSister [2]NP*DAT*|NP*ACC*|NP*GEN*)
AND ([2]NP*DAT*|NP*ACC*|NP*GEN* iDomsOnly [3]PRO^*)
AND ([3]PRO^* idoms h*|H*|$h*|$H*)
AND (*VBP*|*VBD*|*VBI*|*BEP*|*BED*|*BEI*|*HVP*|*HVD*|*HVI*|
   *MDP*|*MDD*|*MDI*|*AXP*|*AXD*|*AXI* iPrecedes [1]NP-NOM)
AND ([1]NP-NOM precedes [2]NP*DAT*|NP*ACC*|NP*GEN*)
```

Data in PPOVpositionB5.out and PPOVpositionB6.out can be sorted if required, e.g. according to features of the object pronoun (e.g. by case or reflexivity), according to subject form (e.g. *man vs*. other) or according to features of the clause (e.g. by narrative mode). Data for first and second person PPOVs were obtained in exactly the same way, replacing only: "[3]PRO^* idoms h*|H*|$h*|$H" with "[3]PRO^* idoms m*|u*|e*|+t*|+d*|i*|v*|y*| M*|U*|E*|+T*|+D*|I*|V*|Y*|$m*|$u*|$e*|$+t*|$+d*|$i*|$v*|$y*| $M*|$U*|$E*|$+T*|$+D*|$I*|$V|$Y*".

Appendix E Acc. special clitic PPOPs: questionable examples

In three examples, the accusative pronoun could be understood as a verbal object and the preposition as stranded in a relative clause. One example was given at (36) of Chapter 4, the other two are given at (1) and (2) below.

(1)   &   hio þæt ylce gewin þe   hio **hine on** bespon mid manigfealdon
        and she the same war   that she him in allured with manifold
        firenlustum twa &   feowertig wintra wæs dreogende
        luxuries   two and forty   winters was conducting
        'and for forty-two winters she conducted the same war into which she had
        allured him with manifold luxuries'

                                                        (coorosiu,Or_1:2.22.9.438)

(2)   ac sio wiðerwearde gebet   &   gelæreð ælcne þara   þe   hio **hi to**
        but the enemy   reforms and converts each   of-those that she her to
        geþiet
        attaches
        'but the enemy reforms and converts each of those to whom she attaches
        herself'

                                                        (coboeth,Bo:20.47.18.856)

In nine examples, the accusative pronoun could be understood as a verbal object and the word parsed as a preposition could be interpreted as a verbal prefix. In each case, the assumed prefixed verb is listed in at least one authoritative Old English dictionary. Three examples were discussed in Chapter 4 (see example (38) and fn. 47). The others are given at (3)–(8) below.

Examples (3)–(5) are from Wærferth's translation of *Gregory's Dialogues* as found in ms. C. The other witness to this translation (ms. O) also has an accusative pronoun in the corresponding clauses. The revised version of the translation as found in ms. H has dative <hyre> corresponding to the emboldened instance of *hi* in (3) and dative <him> corresponding to *hine* in (4). There is no text corresponding to (5) in ms. H.

(3)   Heo wæs gelæded to anre ea   &   bedypped in þæt wæter, &   hi

             she was led   to a   river and immersed in the water and her

             þær þa dryas   ongunnon ferian geond þæt wæter &   mid

             there the sorcerers began   carry over the water and with

             langum onsangum **hi** golon   **on**

             long   incantations her charmed PREFIX

             'She was led to a river and immersed in the water, and there the sorcerers

             began to carry her over the water and charmed her with long incantations'

                                              (cogregdC,GD_1_[C]:10.73.23.821–3)

(4)   **& hine mid** siðode   seo mænigeo þara   þegniendra manna

             and him with-travelled the host   of-the serving   men

             'and the host of courtiers {travelled with / accompanied} him'

                                           (cogregdC,GD_2_[C]:14.131.19.1576)

(5)   &   eac he geseah yrnende þa fulstincendan ea,   seo   […] fylnesse

             and also he saw   running the foul-stinking river which […] fulfillment

             hine geond   floweþ his lichamlicra uncysta

             him throughout flows   his carnal   vices

             'and also he saw running the foul-stinking river which […] flows through

             him in fulfillment of his carnal vices'

                              (cogregdC,GDPref_and_4_[C]:38.322.21.4839)

In the base edition for (6), *on* is marked as an addition contemporary with the manuscript (Sweet 1871: 405, l.6).

(6)   ... oððæt hit bið gewemmed midðæm ðe **hit** cnyssað **on**   unryhta

               until it is defiled   when   it oppress PREFIX unrighteous

             wilnunga, &   hit toterað

             desires   and it destroy

             '... until it is defiled when unrighteous desires strike it down and destroy

             it'

                                   (cocura,CP:52.405.3.2769)

(7) Se witega þurh   Godes gast het   ða Iudeiscan Cristes slagan
the prophet through God's spirit called the Jewish   Christ's slayers
hundas, þe **hine** mid facenfullum mode **ymbe** eodon
dogs   that him with deceitful   mind around went
'The prophet, through the spirit of God, called the Jewish slayers of Christ
— who with deceitful mind surrounded him — dogs'

(cocathom2,ÆCHom_II,_8:70.94.1414)

(8) He &   his geferan   þa   begunnon to lufienne þa micclan
he and his companions then began   to love   the great
druncennisse on nihtlicum gedwylde &   hig þa **hine on**
drunkenness on nightly   wanderings and they then him PREFIX
gebrohton þæt he begann to stelenne on heora gewunan
enjoined   that he began to steal   on their custom
'He and his companions then began to enjoy great drunkenness during
nightly wanderings and they then enjoined him, that he began to steal
according to their custom'

(colsigewZ,ÆLet_4_[SigeweardZ]:1074.488–9)

I treat *on* as a prefix in one further example, (9), but interpret the pronoun
as nom. pl., which — incidentally — is how the combination of *hi*, *on* and *wunian*
is treated in the YCOE at cocura,CPHead: 19.50.52, since *on* and *wuniað* are
written as one word rather than two in the base manuscript on that occasion.

(9) Ðætte on oðre wisan sint to manianne ða   ðe woroldare   wilniað,
that   in other way are to admonish those that worldly-honour desire
&   hie ðonne orsorglice habbað; &   on oðre ða   ðe woroldare
and they then   safely   have;   and in other those that worldly-honour
wilniað, &   ðonne hie gewilnode habbað, **hie** ðonne mid micelre
desire   and when they desired   have,   they then   with great
earfoðnesse &   mid micle broce   **on** wuniað
hardship   and with great misfortune remain

'That those who desire worldly honour, and who keep it safely, are to be admonished in one way; and in another (are to be admonished) those who desire worldly honour, and then having desired (it), they then remain with great hardship and with great misfortune'

(cocura,CP:50.387.1.2617)

In seven examples, the word parsed as a preposition could be interpreted adverbially. In (10), *beforan* co-occurs with *settan*, which is generally transitive (Bosworth & Toller 1898), but understood objects are not always expressed in Old English (Mitchell 1985: §1572–9). A translation for both possible readings of *beforan* is provided. Note also that one of the other 11th-century witnesses has <beforan him> and another has <beforan hym> (Crawford 1922: 202).

(10) Æfter þisum lædde Iosep hys fæder in to þam cyninge, & sette

after   this   led   Joseph his father in to the king   and placed

hine beforan

him before

'After this Joseph brought his father in to the king and placed him (i.e. his father) before him (i.e. the king)' (= prepositional reading of *beforan*)

'After this Joseph brought his father in to the king and placed him in front' (= adverbial reading of *beforan*)

(cootest,Gen:47.7.2040–1)

An adverbial reading of *foran* is possible for the three parallel examples from different versions of the *Chronicles* represented by (11). The parallel examples occur at cochronC,ChronC_[Rositzke]:894.18.886 and cochronD,ChronD_ [Classen-Harm]:894.22.809. In the entry for *forridan* 'to ride before, intercept', Bosworth & Toller (1898) cites (11), with 'the force rode before them' the suggested translation, but this leaves the governor of *hie* ambiguous between *forrad* and *foran*. Yet another possibility is suggested in the *DOE*, in which *foran* is listed in combination with several verbs, including *foran forridan* 'to ride before/cut off/intercept (someone *acc*.)'. This suggests that *foran* could perhaps even be regarded as an adverbial particle here.

(11)   Þa   forrad sio fierd **hie**   **foran**

            then ____ the force them ____

                             (cochronA-2a,ChronA_[Plummer]:894.22.1031)

Foran turns up in another three examples. The first, (12), is included in the *DOE*'s entry for *forgan*, for which, in combination with *foran*, is given the definition 'to cut off, intercept (someone *acc*.)', so an adverbial analysis is possible here too.

(12)   Þa   sume siðe hæfde se cyning **hi**   **forne**   forgan   mid ealre fyrde

            then some time had   the king   them in-front went-by with all   army

            'Then on one occasion the king intercepted them with the entire army'

                             (cochronC,ChronC_[Rositzke]:1009.34.1458)

The other two examples with *foran* are cited in the *DOE*'s entry for *foran* adv./prep. Example (13) is given for *forstandan foran* 'to block / obstruct / stand in the way of (something *acc*.)', and two possibilities are suggested for (14): *forsceotan foran* 'to rush before, hasten to meet / intercept (something *acc*.)' and *forsceotan* 'to forestall / anticipate (something)' + *foran* adv. 'beforehand'. My gloss and translation are in accordance with the second option.

(13)   uton   forstandan **hi**   **foran**   mid gefeohte

            let-us obstruct   them in-front with battle

            'Let us obstruct them with battle'

                          (cocathom2, ÆCHom_II,_22:192.59.4242)

(14)   Walawa þæt ða ungesæligan menn ne magon gebidon hwonne he

            alas   that the unhappy   men not can   wait   when   he

            him   to cume, ac   forsceotað **hine foran**

            them to comes but anticipate him before

            'Alas unhappy men cannot wait for when he comes to them, but anticipate him beforehand'

                          (coboeth,Bo:39.124.9.2477)

The instance of *toweard* in (15a) is recognised to be potentially adjectival by Bosworth & Toller (1898). Given the context, this seems a more likely analysis than prepositional 'toward'. An adjectival reading is also adopted in Morris (1880: 71, l.29) and Kelly (2003: 48, l.86). I give Kelly's translation to illustrate his particular reading. The only feature inconsistent with an adjectival treatment here is the absence of inflection, cf. (15b).

(15) a. ac he sende hehfæderas &   witgan   þa   **hine toweard**   sædon
  but he sent   patriarchs   and prophets who him approaching spoke

  (coblick,HomS_21_[BlHom_6]:71.103.897–8)

  'as He sent patriarchs and prophets who would prophesize His advent'

  (Kelly 2003: 49, ll. 45–6)

  b. Þa halgan ær   Cristes   cyme   on hine gelyfdon, &   hine lufodan,
  the holy   before Christ's coming in him believed   and him loved

  &   hine toweardne   sægdon
  and him approaching spoke

  (coblick,HomS_21_[BlHom_6]:81.285.1022–4)

  'The holy men before Christ's coming believed in Him, loved Him, and spoke of His coming'

  (Kelly 2003: 57, ll. 3–4)

There are two pronouns which are very probably PPOPs, but as masc. sg. acc. pronouns in particular, I see no possible antecedent. One example was given in Chapter 4, at (40). The other is given at (16).

(16)   &   þær is mid Estum   an mægð   þæt hi   magon cyle gewyrcan;
  and there is with Esthonians a   tribe-FEM that they can   cold produce
  &   þy   þær licgað þa deadan men swa lange &   ne fuliað, þæt
  and by-this there lie   the dead   men so   long and not decay that
  hy   wyrcað þone cyle **hine on**
  they produce the   cold him on

'and there is among the Esthonians a tribe that can produce cold, and by this the dead men lie there so long and do not decay, because they produce the cold in it(?)/him(?)'

For the final example, (17a), I agree *to* is prepositional but suggest it governs *þær* rather than *hine*. *To* + acc. is very unusual (see Table 3.5), and the example at (17b) provides support for the proposed analysis.

17. a. And ongean þa arfæstnysse þe of   Godes agenre gyfe cymð deofol
and against the goodness   that from God's own   gift comes devil
sæwð &   sendeð arleasnesse &   gelærð swa þæt ungesælig   man ne
sows and sends   wickedness and urges so   that unfortunate man not
arige ahwar þær   **hine to** onhagige, ne eac mæðe ne geseo on his
cares at-all where him to is-possible nor also virtue not sees   in his
underþeoddum ne on his efengelican
subjects   or in his equals
'And against the goodness which comes from God's own gift, the devil sows and sends wickedness and urges in such a way that an unfortunate man is not merciful in any way that is within his power, nor even recognises virtue in his subjects or in his equals'

   b.   Se ðe þara   mihta hæbbe, arære cirican Gode to lofe, and gif hine
he that of-the ability has   raise church God to glory and if   him
þarto onhagige, sille þar   land to
thereto is-possible give there land to
'He that is able, should raise a church to the glory of God, and if it is in his power, should give land thereto'

Appendix F   Acc. special clitic PPOPs: genuine examples?

There are fifteen pronouns in the YCOE that are parsed as an accusative special clitic PPOP for which I have found only weak support at best for an alternative analysis. Four are represented by examples (41) and (42) of Chapter 4. Another, given below at (1), is from Morris's (1967 [1874–80]) edition of *The Blickling Homilies*. In Kelly's (2003) edition, the initial conjunction and the following four words are italicised (*ibid*: 134, l. 325), indicating some kind of textual problem (*ibid*: lvi). A facsimile of the base manuscript (Willard 1960) shows that the upper margin of the relevant folio (119v) has been trimmed, obliterating the top half of the characters that Kelly italicises. Although this trimming would obscure the presence of a tilde over the vowel of the second pronoun — which would have given nom. pl. <hie> followed by dat. <hĩ> — *him* occurs in the same hand six times in the preceding four folios (it is the final instance of *him* in this particular homily), in each case as <him>. Given the shape and spacing of the visible parts of the damaged characters, original <ꝼ hie hi on asette> is more than likely.

(1)   &   hie **hi**   **on** asette ærest Sancte Petres lichoman on þære stowe þe
         and they them in placed first Saint   Peter's body   in the   place that
         nemned is Uaticanus
         named   is Vatican
         'and they placed in them [i.e. the places that had been built], firstly, Saint Peter's body in the place called the Vatican'
                        (coblick,LS_32_[PeterandPaul[BlHom_15]]:193.391.2511)

Another two examples each involve a verb that is attested only once in the extant Old English materials. I have translated both examples as if the accusative pronoun were a PPOP, although an adverbial reading of the preposition may be a possibility. In neither case is the preposition likely to be a prefix, however. *Ymbutan* does not function in this way in general, while *on* is an unlikely prefix in this case in particular since separable prefixes tend to have transparent semantics (Hiltunen 1983, van Kemenade & Los 2003: 105–6, Elenbaas 2006: 134–6), and 'on, in' this does not seem plausible with 'bellow'.

(2) Wið   oman,   genim ane grene gyrde &   læt sittan þone man on
     Against eruptions take   one green rod   and let sit   the man in
     middan huses   flore, &   bestric   **hine ymbutan**
     middle house's floor and make-stroke him around

                                   (colacnu,Med_3_[Grattan-Singer]:177.1.787–9)
     'For eruptions: take a green rod. And make the man sit in the middle of the
     floor of the house, and round him strike a circle'

                                   (Grattan & Singer 1952: 193, CLXXVIIc.)

(3) Hwilum   hi   **hine** bylgedon **on** swa fearras ond ðuton   eallswa wulfas
     for-a-time they him bellowed on like bulls   and howled as-if   wolves
     'For a time they bellowed at him like bulls and howled as if wolves'

                                   (comart3,Mart_5_[Kotzor]:Ja17,A.19.131–2)

     In three examples, the pronoun in question appears to be an accusative
special clitic PPOP, although for each example another Old English witness has a
dative rather than accusative pronoun. For (4) and (5), ms. B, an early 11th-
century version of *Bede*, has dative <him>.

(4) Þa   dyde he swa: gebæd **heo   fore** &   heo gebletsade & Gode
     then did   he so   prayed them for   and them blessed   and God
     bebead
     commended
     'Then he did so: he prayed for them, and blessed them and commended
     them to God'

                                   (cobede,Bede_3:13.198.31.2020)

(5) ... gif he his honda **hiene on** sette   &   hiene blætsian wolde
       if he his hand(s) hi   on placed and him   bless   wished
     '… if he placed his hand(s) on him and wished to bless him'

                                   (cobede,Bede_5:5.396.23.3960)

For (6), ms. H has dative <hyre> corresponding to *hi*.

(6)   ... oð þæt se deofol of   hire uteode,   þe **hi** ær   **in** gefor
               until   the devil from her out-went, that her before in left
              '... until the devil, which had previously entered her, left her'

                                          (cogregdC,GD_1_[C]:10.73.23.823)

Ms. H has no text corresponding to (7), but a few clauses later the C text (from which (7) derives) has *... þam broðrum, þe **him ymb** wæron* '...the brethren who were around him'.

(7)   ... þæt he bodode   þæs   halgan lifes word &   lare   geond   þa
               that he preached of-the holy   life's words and teachings throughout the
              mynstru,   þe **hine ymb**   tymbrode wæron
              monasteries that him around built   were
              '... that he preached the words and teachings of holy life throughout the
              monasteries, which were built about him'

                                      (cogregdC,GDPref_and_4_[C]:11.275.1.4003)

Also potentially relevant to (7), is Clark Hall's listing of *infaran* 'to enter'. This would potentially correspond with the Latin behind this translation, which has a prefixed verb rather than a verb + PP (*is qui eam invaserat*). However, a little later in the same text file from which (7) derives we find *gefor* combining once more with an acc. PP headed by *in* (cogregdC,GD_1_[C]: 10.75.20.847). This example supports the YCOE parse for (7).

     The last four examples are given below. The apparent use of natural rather than grammatical gender in (8) is not too surprising since the referent is non-human (see Mitchell 1985: §69, 3b).

(8)   syllað us eac þa cartan   þæt we **hyt**   magon **on** awrytan þæt ðæt
              give   us also the paper-FEM that we it-NEUT can   on write   that that
              we gehyrdon and eac gesawon
              we heard   and also saw

'Give us also paper that we may write on it that which we heard and also saw'

<div align="right">(conicodA,Nic_[A]:17.3.1.397)</div>

Although (9) survives only in the 12th-century Bodley ms., *smeagan* commonly combines with an acc. PP headed by *ymb(e)*.

(9)   gif ge **hit** georne   **ymbe** smeagan willað & æfter spyrigan,
        if   you it carefully about think   wish   and after examine
        'if you wish to carefully consider it and later examine it'

<div align="right">(coboeth,Bo:16.36.4.651)</div>

The example at (10) could involve an instance of *onfeohtan* 'to attack, fight with', which is how the collocation of *on* and *feohtan* is parsed later in the same text (see colawaf,LawAf_1:42.6.151). On the other hand, the same text also provides an example in which *feohtan* clearly occurs with an acc. PP headed by *on* (colawaf,LawAf_1:42.4.146), which suggests that the YCOE parse of (10) is probably correct.

(10)   Gif he mægnes hæbbe, þæt he his gefan beride   &   inne   besitte,
       if   he power   has   that he his foe   surround and within surround,
       gehealde hine VII   niht   inne   & **hine on** ne feohte, gif he inne
       keep   him seven nights within and him on not fight   if he within
       geðolian wille
       remain   wishes
       'If he have power that he can surround his foe and besiege him within, he should keep him seven nights within and not attack him, if he wishes to remain within'

<div align="right">(colawaf,LawAf_1:42.1.136–7)</div>

In (11), from a Latin translation, the relative clause could correspond to the Latin, which has a prefixed verb rather than a verb + PP (*qui hanc bene regere praevalet imponatur*), but none of the Old English dictionaries lists **ofermagan*.

(11)   ... sua is cynn   ðæt sio giemen ðære halgan ciricean, ðæt is Cristes
          so is proper that the care   of-the holy   church   that is Christ's

folces   gesomnung, sie ðam beboden ðe   **hie** wel **ofer** mæge,   &
people's assembly   is him offered   that it   well over prevails and

hiere wel rædan cunne
it   well rule   can

'... so it is proper that the care of the holy Church, that is the assembly of
Christ's people, be offered to him who prevails over it well and can rule it
well'

(cocura,CP:5.43.22.248)

Appendix G   Independent effects model

Table E.1 Independent effects model for placement of PPOPs in Old English

Variable (significance relative to this model)	Factor	Left-of-P/Total	Factor weight
PREPOSITION ($p < 0.001$)	*ongean* 'towards, against'	101/105 (96%)	0.968
	togeanes 'towards, against'	107/122 (88%)	0.927
	miscellaneous	242/484 (50%)	0.523
	to 'to'	1,334/3,203 (42%)	0.514
	fram 'from'	140/457 (31%)	0.496
	on 'on, in'	186/457 (41%)	0.475
	beforan 'before'	41/213 (19%)	0.423
	of 'of'	58/175 (33%)	0.389
	æt 'at'	25/164 (15%)	0.259
	'between' (other)	20/139 (14%)	0.244
	æfter 'after'	47/223 (21%)	0.225
	mid 'with'	126/1,252 (10%)	0.143
	Range		*0.825*
LATIN INTERFERENCE ($p < 0.001$)	Non-translations	915/1,956 (47%)	0.755
	Non-biblical, unmatched	112/220 (51%)	0.745
	Non-biblical, unsampled	48/137 (35%)	0.602
	Biblical, unmatched	93/482 (19%)	0.388
	Non-biblical, matched	19/86 (22%)	0.385
	Biblical, matched	58/554 (10%)	0.341
	Biblical, unsampled	38/428 (9%)	0.263
	Range		*0.492*
MAIN VERB ($p < 0.001$)	*faran* 'to go'	32/107 (30%)	0.628
	sendan 'to send'	74/187 (40%)	0.612
	niman 'to take'	45/136 (33%)	0.608
	cuman 'to come'	333/722 (46%)	0.600
	cweðan 'to say'	543/1,241 (44%)	0.573
	sprecan 'to speak'	51/156 (33%)	0.567
	clipian 'to call'	53/114 (46%)	0.565
	gan 'to go'	58/157 (37%)	0.527
	bringan 'to bring'	40/102 (39%)	0.509
	miscellaneous	1,023/3,106 (34%)	0.484
	BE	125/622 (20%)	0.443
	lædan 'to lead'	31/158 (20%)	0.340
	habban 'to have'	19/276 (7%)	0.139
	Range		*0.489*

CLAUSE TYPE	Main	642/1,987 (32%)	0.610
(*p* < 0.001)	Subordinate	777/2,315 (34%)	0.575
	Main conjunct	812/2,139 (38%)	0.558
	Infinitival	78/209 (37%)	0.483
	Subordinate conjunct	97/272 (36%)	0.467
	Other	21/72 (29%)	0.314
	Range		*0.296*
PERSON	Third person	1,879/4,597 (41%)	0.666
(*p* < 0.001)	First person	114/956 (12%)	0.421
	Second person	80/788 (10%)	0.409
	Range		*0.257*
LINEAR ORDER OF PP & V	[PP(…)V]	1,380/3,012 (46%)	0.620
(*p* < 0.001)	[V(…)PP]	1,047/3,982 (26%)	0.380
	Range		*0.240*
SUBJECT FORM	Nominal	969/2,179 (44%)	0.583
(*p* < 0.001)	Other	1,006/2,613 (38%)	0.561
	Personal pronoun	452/2,202 (21%)	0.359
	Range		*0.224*
DIALECT	Other	404/851 (47%)	0.586
(*p* < 0.001)	West Saxon	1,387/3,189 (43%)	0.548
	West Saxon+Anglian	282/972 (29%)	0.440
	West Saxon+Ang. Mercian	165/518 (32%)	0.426
	Range		*0.160*
NARRATIVE MODE	Diegetic	2,093/4,804 (44%)	0.578
(*p* < 0.001)	Mimetic	334/2,190 (15%)	0.422
	Range		*0.156*
ADJACENCY OF PP & V	Adjacent	1,901/4,843 (39%)	0.562
(*p* < 0.001)	Non-adjacent	526/2,151 (24%)	0.438
	Range		*0.124*
DATE	Late (post-AD 975)	1,868/4,526 (41%)	0.560
(*p* < 0.001)	Early (pre-AD 925)	291/805 (36%)	0.440
	Range		*0.120*
REFLEXIVITY	Reflexive	191/608 (31%)	0.555
(*p* < 0.001)	Non-reflexive	2,236/6,386 (35%)	0.445
	Range		*0.110*
NUMBER	Plural	172/895 (19%)	0.501
(*p* = 0.009)	Singular	293/1,546 (19%)	0.499
	Range		*0.002*
	Total	2,427/6,994 (35%)	Input 0.093
		Log Likelihood -3164.834	

Appendix H   Left-of-P frequency by subject form and text

Table F.1 Left-of-P frequency by subject form and text

	Full NP	'Other'	Personal pronoun
Ælfric's Homilies I	67/135 (50%)	64/136 (47%)	17/112 (15%)
Ælfric's Homilies II	73/129 (57%)	71/126 (56%)	26/117 (22%)
Ælfric's Lives of Saints	148/204 (73%)	104/196 (53%)	48/159 (30%)
Ælfric's Suppl. Homilies	81/119 (68%)	54/120 (45%)	49/131 (37%)
Bede	20/82 (24%)	37/135 (27%)	7/94 (7%)
Blickling Homilies	20/68 (29%)	27/79 (34%)	8/65 (12%)
Gregory's Dialogues (C)	43/124 (35%)	43/159 (27%)	29/123 (24%)
	43/124 (35%)	43/159 (27%)	29/123 (24%)
Heptateuch	33/176 (19%)	41/195 (21%)	13/180 (7%)
Orosius	55/118 (47%)	40/78 (51%)	27/79 (34%)
	55/118 (47%)	40/78 (51%)	27/79 (34%)
Vercelli Homilies	14/55 (25%)	34/78 (44%)	21/76 (28%)
West Saxon Gospels	43/273 (16%)	39/330 (12%)	20/310 (6%)
Other	372/696 (53%)	452/981 (46%)	187/756 (25%)
Total	969/2,179 (44%)	1,006/2,613 (38%)	452/2,202 (21%)

Non-significant differences are indicated for each text in Table F.1 by shading, e.g. shading indicates that there is no significant difference in frequency of left-of-P placement in *Ælfric's Homilies I* between PPOPs that co-occur with a full NP and those that co-occur with an 'other' subject type. Text files in the 'Other' category each supply less than fifty examples to one or more of the three categories.

Table G.1 Left-of-P frequency by narrative mode and text

	Mimetic	Diegetic	Total	Mimetic
Boethius	48/133 (36%)	1/5 (20%)	49/138 (36%)	96%
St. Augustine's Soliloquies	11/63 (17%)	2/4 (50%)	13/67 (19%)	94%
Nicodemus (C)	6/37 (16%)	5/15 (33%)	11/52 (21%)	71%
Nicodemus (A)	17/60 (28%)	24/32 (75%)	41/92 (45%)	65%
Genesis	1/35 (3%)	0/29 (0%)	1/64 (2%)	55%
Heptateuch	15/301 (5%)	72/250 (29%)	87/551 (16%)	55%
Mary of Egypt	13/30 (43%)	19/27 (70%)	32/57 (56%)	53%
Saint Margaret (C)	3/25 (12%)	21/29 (72%)	24/54 (44%)	46%
Apollonius of Tyre	5/23 (22%)	15/30 (50%)	20/53 (38%)	43%
Saint Eustace	2/21 (10%)	6/32 (19%)	8/53 (15%)	40%
West Saxon Gospels	14/322 (4%)	88/591 (15%)	102/913 (11%)	35%
Seven Sleepers	4/20 (20%)	18/39 (46%)	22/59 (37%)	34%
Blickling Homilies	6/69 (9%)	49/143 (34%)	55/212 (26%)	33%
Ælfric's Lives of Saints	28/178 (16%)	272/381 (71%)	300/559 (54%)	32%
Ælfric's Homilies II	18/115 (16%)	152/257 (59%)	170/372 (46%)	31%
Ælfric's Homilies I	11/115 (10%)	137/268 (51%)	148/383 (39%)	30%
Vercelli Homilies	13/75 (17%)	77/186 (41%)	90/261 (34%)	29%
Gregory's Dialogues (H)	12/43 (28%)	68/114 (60%)	80/157 (51%)	27%
Cura Pastoralis	12/54 (22%)	60/154 (39%)	72/208 (35%)	26%
Other	60/283 (21%)	357/821 (43%)	417/1,104 (38%)	26%
Ælfric's Suppl. Homilies	23/91 (25%)	161/279 (58%)	184/370 (50%)	25%
Holy Rood Tree	3/17 (18%)	28/70 (40%)	31/87 (36%)	20%
Bede	8/60 (13%)	56/251 (22%)	64/311 (21%)	19%
Martyrology 3	0/8 (0%)	43/81 (53%)	43/89 (48%)	9%
Orosius	1/9 (11%)	121/266 (45%)	122/275 (44%)	3%
Chronicles (A, C, D, E)	0/3 (0%)	241/450 (54%)	241/453 (53%)	1%
Total	334/2,190 (15%)	2,093/4,804 (44%)	2,427/6,994 (35%)	31%

NB The final column expresses the proportion of all PPOPs that occur in mimetic contexts. Text files in the 'Other' category each supply less than fifty PPOPs.