

Chapter 23

V2-relatives in Old English

Bettelou Los & Stefano Coretta

The verb-placement asymmetry between main (verb-second, “V2”) and subclauses (verb-final) found in Dutch and German is less pronounced in Old English (OE) (Pintzuk1999); the question is to what extent, as the distinction between main and subclauses cannot readily be determined on the basis of punctuation or other typographical conventions. It has long been noted that Old English shows constructions similar to the “V2-relatives” in modern Dutch and German (Mitchell1985), while being much more difficult to identify. This paper proposes that the discourse functions of such clauses in Dutch and German might be used as a diagnostic for V2-relatives in OE, as these clauses tend to be used for “second mentions” of a major protagonist who has just been introduced. V2-relatives are less embedded than other relative clauses, and can be argued to represent a midway point in the grammaticalization of relatives, from a paratactic to a hypotactic structure. Key here is the difference between a demonstrative and the relative pronoun that developed from it.

1 Introduction

In the modern West Germanic languages, movement of the finite verb to second position (“verb-second” or “V2”) is a main clause phenomenon, and hence leads to a word order asymmetry between main (V2) and subclauses (verb-final). Although the consensus is that Old English is also a V2-language, albeit with some modifications (vanKemenade1987, see also Section 2.4 below), the asymmetry between main and subclause is less pronounced (Pintzuk1999). Main and subclauses, and sentences in general, are not distinguished in the manuscripts by a consistent punctuation system, which adds to the challenges of determining the extent of the asymmetry. There is one type of clause in particular that has

long been known to be analytically ambiguous (Allen1977, Mitchell1985), and may well be a contributing factor to the difficulty of establishing to what extent OE exhibits the word order asymmetry found in the modern West Germanic languages:¹

- (1) *Maurus gemette ænne man eft **se** wæs yfele getawod*
Maurus met a man afterwards SE.NOM.M was evilly stricken
‘Maurus met a man afterwards who was evilly stricken.’
[ÆLS[Maur]:283.1661]
- (2) *And Maurus ða gemette ær he to mynstre come ænne dumbne*
and Maurus then met before he to monastery came a dumb
*cnapan 7 **se** wæs creopere eac*
boy and SE.NOM.M was cripple too
‘and Maurus then met before he came to the monastery a dumb boy and
he was a cripple as well.’
[ÆLS[Maur]:19.1507–8]

For OE, syntactic investigations have been able to rely on the York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE, TaylorTaylor2003), a 1.5 million word corpus of Old English prose based on the 1981 release of the Toronto *Dictionary of Old English Corpus*, containing all the major Old English prose works. Each word of the corpus is tagged for part of speech, and syntactically annotated by labelled bracketing in the Penn Treebank format, which has since become the industry standard for syntactic annotation. Data can be retrieved by CorpusSearch 2 (Randall2005–2007).

In YCOE, (1) has been coded as a relative clause, and hence as a subclause (IP-SUB), whereas (2) has been coded as a main clause (IP-MAT), in spite of their very similar structures including their main-clause-like word order with V2. It may well be the presence of the Tironian sign ‘7’ for “and” in (2) that guided the annotator to an analysis of a sentence consisting of two main clauses rather than a main clause and an embedded clause, which is the analysis suggested by the coding of (1). In the PDE translation in Skeat1966’s (Skeat1966 [1881–1900]) edition of Ælfric’s *Saints Lives* from which (1–2) have been taken, cases like (1) are routinely translated by relative clauses, and the *se*-strategy is recognized as one of a range of relativizing strategies available in OE, so, from that perspective, the interpretation of clauses like (1) as relatives is unproblematic. What poses a problem,

¹The demonstrative form is in bold, and the finite verb is underlined; they will be marked like this throughout the paper; also note that the demonstrative, the source of the ambiguity, will be glossed as *se*, and will be variously translated as a relative pronoun, a demonstrative pronoun or a personal pronoun.

however, is that relative clauses are not necessarily subordinate clauses – they may also be independent main clauses, and hence have main clause word order. Coding these as IP-SUB will boost the number of subordinate clauses in the data that have the same word order as main clauses, and hence obscure the word order asymmetry between main and subclauses in OE. That OE is an asymmetrical V2 language is already clear from vanKemenade1987, vanKemenade1997 work; see also Taylor2014, and particularly MeklenborgSalvesen2017, who conclude from their investigation of OE *þæt*-complement clauses that OE does not allow embedded V2. This paper adds some further data to these insights as they pertain to relative clauses, and shows that the specific characteristics of V2-relatives echo those found in modern West Germanic, and lead to their specific discourse function of marking a newly introduced protagonist.

This paper will look at all relative clauses in two OE texts: Ælfric’s *Lives of Saints* (Skeat1966 [1881–1900]) and the H manuscript (H ms, or ms H) of *Gregory’s Dialogues* (Hecht1965 [1900–1907]), and compare the use of *se*-relatives with the other types of relative in terms of position of the verb, discourse functions, and antecedents. The paper is structured as follows: Section 2 will present the types of relative clauses that have been identified for OE: *þe*-relatives, which are marked only by the invariant complementizer *þe*, the “universal embedder”; *se þe*-relatives, which are marked by the universal embedder as well as a demonstrative pronoun; and *se*-relatives, which are marked only by the pronoun, and which are the focus of the investigation. Section 3 discusses what has been said about V2-relatives in the modern West Germanic languages. Section 4 and Section 5 present the data and the methodology; as expected, it is the *se*-relative that shows the highest frequencies of V2. What is new is that they serve a particular discourse function in narrative in that they instantiate the second mention of a newly-introduced protagonist. Section 6 will place these findings in the wider context of the semantics of main and subclauses, and the development of hypotaxis. Section 7 concludes this paper.

2 Typology of OE relatives

2.1 Introduction

This section will provide a very brief overview of the main types of relative clauses in OE: *þe*-relatives, *se þe*-relatives and, finally, the *se*-relatives, the type that our investigation into V2-relatives is most concerned with. For more detailed accounts, the reader is referred to the vast literature on the history of relative

clauses in English that cannot be done justice here, see Allen1977, Mitchell1985, Traugott1992, Fischer2000, Suárez-Gómez2012.

2.2 *þe*-relatives

Out of all types, the *þe*-relative is the most frequent. In YCOE, these are marked as having C occupied by *þe* or *ðe* while the Spec,CP position remains empty. *þe*-relatives are usually verb-final, and restrictive. Unlike *se*-relatives, they allow resumptive pronouns and preposition stranding – evidence of a gap, and hence of movement/copying (Allen1977; vanKemenade1987). *þe* is a worn-down form of *þæt* ‘that’ which has become an invariant *universal embedder* in OE, marking not only relative clauses but also other clauses that we assume to be embedded (cf. Gothic *ei*). The consensus in the literature is that *þe*-relative clauses are embedded, and subordinate. Example (3) contains two *þe*-relatives (*þe* and *ðe* in bold) as well as a *se*-relative (*ðone*); all three are coded as IP-SUB in YCOE. *Se*-relatives will be discussed in Section 2.4.

- (3) *Sum æþel-boren þægn wæs philippus gehaten · ðone asende se*
 Some nobly-born thane was Philippus called SE.ACC.M sent the
casere Commodus þe on ðam dagum · rixode fram rome byrig to
 emperor Commodus þE on those days ruled from Rome city to
ðære byrig ðe is gehaten alexandria
 the city þE is called Alexandria
 ‘A certain nobly-born thane was named Philip, who the emperor
 Commodus, who was in charge in those days, sent from the city of Rome
 to the city which is named Alexandria.’ [ÆLS_[Eugenia]:5.198]

The second *þe*-relative in (3) is restrictive, and postmodifies an NP (*to ðære byrig* ‘to the city’) that also contains a demonstrative form, the singular feminine *ðære*, which is grammaticalizing into a definite article in OE; *ðære* is used here to signal to the reader that they should be able to identify the referent, not because it has been mentioned in the earlier discourse (there is no earlier mention of a city which could be the antecedent) or because readers are expected to know it (unlike Rome), but because there is going to be a postmodifier (in the form of a *þe*-relative) which will clarify what city is meant. ArgüellesLos2022 identify this use of the definite article as marking an important stage in the grammaticalization process, namely, the stage at which definite articles appear with a referent that is new in the discourse; the use of the definite article is “licensed” by a postmodifier that “activates the general knowledge necessary to locate the referents

in a contextual set” (Wagener2017). Breban2012 refers to them as “definite first mentions”. Typically, such a use of the definite article signifying identifiability is no longer compatible with a demonstrative reading.

The first instance of the *þe*-relative, however, is not restrictive – it postmodifies a named referent, as does the *se*-relative with *ðone*. The reason for this will be explored in Section 3.4.

2.3 *Se þe*-relatives

Se þe-relatives consist of a form of the distal demonstrative *se* as well as the embedder *þe*. The distal demonstrative doubles as the definite article in OE. Its paradigm is given in Table 1:

Table 1: The paradigm of the distal demonstrative in OE

	singular			plural
	masculine	feminine	neuter	all genders
nom	se	seo	þæt	þa
gen	þæs	þære	þæs	þara
dat	þæm	þære	þæm	þæm
acc	þone	þa	þæt	þa
instr	þy, þon			

Se þe-relatives have an inherent degree of ambiguity, as they may represent two or perhaps three separate types: (i) a *þe*-relative that has a demonstrative antecedent as an appositive NP with a restrictive relative clause, comparable to PDE “i.e.” (*hiera mægas him mid wæron þa þe him from noldon* ‘their kinsmen were with him, i.e. those who did not want to leave him’ in [ChronA 755.1–29]); (ii) a *þe*-relative whose demonstrative antecedent has a syntactic function in the higher clause, comparable to the generic “free relative” use of PDE *he that lives in a glass house should not throw stones* (Allen1977; see also Allen2020 for more detail and analysis of this type); (iii) there is also the possibility that *se þe* is a “compound” relativizer.

In YCOE, types (i) and (ii) differ from each other in that type (ii) has an empty Spec,CP position, as the demonstrative is best analyzed as a constituent of the higher clause, outside the relative clause. We have followed YCOE coding in our searches, which means that type (ii) will have been included under the *þe*-relative.

The *se þe*-relative has been described as “mostly restrictive” (Mitchell1959). *Se þe*-relatives are not found with preposition stranding, and this is the only type that exhibits “case attraction”, a fairly rare phenomenon in the case of type (i), in which the demonstrative appears with the case it would have if it was a constituent of the higher clause (vanKemenade1987; Allen2020); this might be a kind of appositional use, see *him* in the translation:

- (4) *ge seceað þone hælynd þone þe on rode ahangen wæs*
 you seek the.ACC savior SE.ACC.M þE on cross hanged was
 ‘You seek the Savior, him, who was hanged on the cross.’
 [cowsgosp,Mt_[WSCp]:28.5.2139]; (Allen1977)

2.4 *Se*-relatives

Se-relatives are the type that is the focus of this investigation. An example is (5):

- (5) *Se feorða godspellere is Iohannes Cristes moddrian sunu, se*
 the fourth gospelwriter is John Christ.GEN aunt.GEN son SE.NOM.M
wæs Criste swa leof þæt he hlynode uppān his breoste.
 was Christ.DAT so dear that he leaned upon his breast
 ‘The fourth gospelwriter is John, Christ’s aunt’s son, who was so dear to
 Christ that he leaned upon his breast.’ [ÆLS_[Mark]:159.3304]

Se-relatives often have V2 word order (Allen1977; vanKemenade1987). They do not show case attraction, and do not occur with resumptives or preposition stranding, that is, without a clear structural gap (Wende1915, Allen1977). *Se*-relatives in OE are, as a rule, non-restrictive (LosvanKemenade2018; Roma2007: 266), which are known to be less embedded than restrictive relatives: They do not define their antecedent but add extra information, and are less integrated into the structure of the higher clause from a syntactic, prosodic, and informational point of view (HuddlestonPullum2002). Non-restrictive relative clauses in PDE are well-known to exhibit main clause behavior, such as verb phrase preposing, negative constituent preposing and topicalization (Emonds1970, HooperHooper1973, Green1976, deVries2012, Heycock2017). In terms of discourse function, they are often used for expressing the next event in a plotline as in (6) (foregrounding) but also for explanations, reasons and causes (backgrounding), as in (7):

- (6) *She gave the letter to the clerk, who copied it.* (Depraetere1996)
 (7) *If Peter called the dean, who hates me, I would be in trouble.*
 (JasinskajaPoschmann2020)

Another interesting use of relative clauses in Dutch is noted by deVries2012 (relative clause in bold):

- (8) *Joop, **die nu zijn kamer gaat opruimen!**, komt straks buiten spelen.*
 Joop who now his room goes clean comes later outside play
 ‘Joop, who is now going to clean his room!, will come play outside later.’

A sentence like (8) might be uttered by Joop’s mother to a friend of Joop’s who has presented himself at the door, importantly, within earshot of Joop; the non-restrictive relative clause is, in fact, an order to Joop.

The most natural way to translate *se*-relatives into PDE is with a relative clause, particularly, if they have human antecedents, as in (5). Although singular animate antecedents can be relativized by *who* or *that*, they no longer allow a translation by an independent demonstrative, unless they are in the plural, witness the contrast between (9a) and (9b) below. Example (9c) shows “a contrast in animacy, *he* referring to a person or animal, *that* to an inanimate” (HuddlestonPullum2002). Compare also the anaphoric uses in (9d–9e).

- (9) a. *Those who obtain a score of 90% will win a prize.*
 b. **That who obtains the highest score will win a prize.*
 c. *He/*That saved my life.*
 d. *The population of Victoria far exceeds that of Queensland.*
 e. **The premier of Victoria will be meeting with that of Queensland.*
 (HuddlestonPullum2002)

In Dutch and German, main clauses headed by demonstrative pronouns coerce topic shift (Comrie1997): The implication of using a demonstrative rather than a personal pronoun coerces interpreting the referent as a newly introduced one rather than as a continuing topic. This is no longer the case in PDE, so that the referent of *she* underlined in (10) is ambiguous – this is what makes the exchange interesting:

- (10) Columbo: *No, my wife is not here. She had to go to Chicago to look after her mother. She had a fall and broke her hip.*
 Woman at party: *Oh, your wife broke her hip? How terrible.*
 Columbo: *No, her mother.*

(Columbo, series 10.1, episode *No time to die*;
 (LosvanKemenade2018)

The natural choice to disambiguate the two possible referents in Dutch and German would be the personal pronoun for the continued topic (the wife) and the demonstrative for the new topic (the mother). In PDE, a relative clause would have disambiguated the referent; “who had a fall and broke her hip” could only have referred to the mother. This suggests that (non-restrictive) relative clauses may have stepped in as a compensatory strategy once the topic shifting function of demonstratives was lost; it also explains why a main clause headed by a “standalone” demonstrative referring to a singular human antecedent is much more likely to be annotated as a relative clause than as a main clause by annotators who, as in the case of YCOE, tended to be native speakers of English.

This loss in functionality of the singular demonstrative is already evident in Early Middle English, where one finds innovative periphrases – which did not catch on – *he this, she this* (Mustanoja1960: 137; Miura2016) as translations for topic-shifting Latin *iste*, indicating a need to restore the earlier functionality. The account provided in Allen2022 shows that *se*-relatives – reduced to invariant *þe* after the loss of the paradigm of Table 1 – are no longer found after 1150 with antecedents in the previous discourse, only with generic referents in free relatives, corresponding to Present-Day English *he who*; the loss of gender is tentatively identified as being implicated, supporting a suggestion made in LosvanKemenade2018, who point to the fact that the contrast in acceptability between singular and plural demonstratives in PDE as in (9a) versus (9b) makes sense in the context of the OE demonstrative paradigm, which was gendered in the singular but not in the plural.

Cases like (5), and the *ðone*-clause in (3), ambiguous between main clause and subclause, are the focus of the remainder of this paper. This ambiguity extends to standalone demonstratives as complements of a preposition in a clause-initial PP, as in (11):

- (11) *His wif wæs gecyged claudia be þære he gestrynde twægen suna ·*
his wife was called Claudia by SE.DAT.F he had two sons
auitum · 7 særgium · 7 ane dohtor eugenian þe we embe spræcaþ.
Avitus and Sergius and one daughter Eugenia þe we about talked
‘His wife was called Claudia, by whom he had two sons, Avitus and
Sergius, and one daughter Eugenia who we mentioned earlier.’
[ÆLS_[Eugenia]:15.202]

The fact that they do not allow preposition stranding, and thus always show “pied piping” as in (11), has been argued to be evidence that *se*-relatives involve movement or copying (Allen1977). Note that the absence of stranding could

equally well serve as evidence of main clause “Topicalization” (movement to the pre-field); PP-initial adjuncts linking to previous discourse are much more frequent in OE than in later periods, particularly PPs that contain standalone demonstratives as in (11); see [los2023decline](#) for the decline of such “local anchors” in the history of English. PDE favors subjects as the default expression of given information, so that a translation like *by/with her he had two sons* is less felicitous in PDE than a relative clause *by/with whom he had two sons*. This is why such PPs, too, are much more likely to be annotated as relative clauses than as PP-initial main clauses.

Note that the surface V3-order of (11) is part of the manner in which “V2” plays out in OE: The finite verb will move to C in “focus” contexts (interrogatives, negative initial constituents as well as the “narrative operator” *þa* ‘then’ as the first element). It will move to a lower head, however, if some other initial phrase, such as a PP, is occupying the prefield, which leads to surface V3 if the subject is a pronoun, as in (11) ([vanKemenade1987](#)). This means that the sequence *be þære he gesitynde* is a possible main clause order.

In YCOE, *se*-relative clauses receive a syntactic annotation in which the demonstrative is in Spec,CP and C is empty (marked by “0”); as a relative, the clause-label indicates a subclause (IP-SUB). As an example, (12) shows the annotation of the *ðone*-clause of (3); (5) and (11) are annotated along similar lines.

- (12) (CP-REL-1 (WNP-ACC-4 (D^A +done))
 (C 0)
 (IP-SUB (NP-ACC *T*-4)
 (VBD asende)
 (NP-NOM (D^N se) (N^N casere))

An example of a standalone demonstrative coded as heading a main clause is (13), with the demonstrative as a complement of a preposition:

- (13) & *þa sealde he heom mid his agenre handa ane trywene flascan*
 and then gave he them with his own hands a wooden bottle
 wines fulle, to þon þæt hi mihton heom þa on
 of-wine full to that that they might for-themselves SE-ACC.SG.F in
 heora færelde to underngeweorce habban. Of þære hi
 their journey for breakfast have from SE-DAT.SG.F they
 druncon, oð þæt hi to rauennam becomon
 drank, until that they to Ravenna came
 ‘and then, with his own hands, he gave them a wooden bottle full of wine,
 so that they might have that for breakfast during their journey. They

drank from it, until they arrived in Ravenna.’ (i.e. it lasted them all the way to Ravenna) [GD_1_[C]:9.66.12.742–743]

In (13), a referent (a feminine noun meaning ‘bottle’, underlined) is introduced and then referred back to by two instances of *se* (in bold). The second instance, of *þære*, receives the following syntactic annotation in YCOE, with the clause-label indicating a main clause (IP-MAT):

- (14) (IP-MAT (PP (P Of)
(NP-DAT (D[^]D +t+are)))
(NP-NOM (PRO[^]N hi))
(VBDI druncon))

What makes an analysis like (14) more likely for an example like (13) is the distance between the antecedent and the demonstrative. Faced with a similar dilemma – demonstrative-headed main clause or embedded relative? – in her Old High German data, *Schlachter*2012 also uses distance between antecedent and demonstrative to support a main clause analysis of analytically ambiguous demonstrative-headed V2 constructions.

The question is whether (3) (the *ðone*-clause), (5) and (11) should receive a syntactic annotation along the lines of (14), with the demonstrative (whether as part of a PP or as a subject or object) heading an independent main clause, rather than (12), with the demonstrative heading a subclause.

3 V2-relatives in Dutch and German

3.1 Introduction

The dilemma posed by *se*-headed structures in OE is presented in the literature as the choice of a main clause headed by a standalone demonstrative versus an embedded relative clause. The emphasis of the literature on V2-relatives in Dutch and German, however, is not so much on the status of main versus relative clause but on the difference between embedded and less embedded relative clauses. The analysis of main clauses headed by demonstratives as separate sentences never seems to be in doubt, presumably given the clear markings of sentences in written texts by means of punctuation, and their intonational phrasing in speech. The discussions in *Gärtner*2001 and *Zwart*2003 among others focus on V2-clauses that are part of a sentence that includes a previous clause where the referent,

typically, is newly-introduced, of the type of (15b–15c). The contrast, then, is between (15b) and the embedded, verb-final relative clause as in (15c).

- (15) a. *iemand had twee zonen.*
 someone had two sons
 [main clause with V2]
- b. *Er was iemand die had twee zonen.*
 there was someone that had two sons
 [main clause + V2-relative]
- c. *Er was iemand die twee zonen had.*
 there was someone that two sons had
 [main clause + embedded relative]
 ‘There was someone who had two sons.’

Both Gärtner2001 and DenDikken2003 discuss a third structure as in (16), on the basis of an observation made by Zwart2003 about responses to a 19th-century dialect questionnaire. This questionnaire presented the sentence in (15a) as the start of a story, its function being to introduce the father. The respondents subsequently gave (15b), the V2-relative, but also, unexpectedly, (16):

- (16) *Er was iemand en die had twee zonen.*
 there was someone and that had two sons
 [main clause + main clause]
 ‘There was someone and he had two sons.’

Example (16) contains two main clauses, the second headed by a standalone demonstrative. These were apparently felt, by the respondents, to be equivalent, an indication that they both serve the same function. Gärtner2001, discussing a similar set of sentences (*Es gibt Philosophen (und) die stammen aus Griechenland* ‘there are philosophers who come from Greece’) notes that the main difference between the version with and without *und* ‘and’ is one of informational units – a relative clause, whether V2 or embedded, will form an informational unit, as also defined by their intonation/prosody (a rising tone before the relative clause). In the version with *und*, a “falling tone, drop, pause, or similar features” (p. 109) would be expected. This does not sufficiently explain the Dutch case of (16), however, where the first main clause is just as uninformative as in (15), and would not have a falling tone after *iemand*.

The different emphases in the various discussions can be visualized by the Venn diagram in Figure 1.

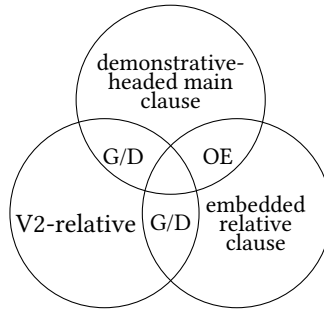


Figure 1: Focus of relevant investigations in the literature on OE and on German (G) and Dutch (D)

The existence of the V2-relative as a less embedded relative clause has not been mooted in the literature on OE – the dilemma is, as far as we are aware, always presented as being between a relative clause (assumed to be embedded by definition) and a main clause with a standalone, independent demonstrative. We will see in the following sections that the tests put forward in the literature about Dutch and German V2-relatives to untangle the three possibilities, that is, (15b)–(15c) and (16), will be too fine-grained for OE; but we will be able to say something about the embedded option versus the rest, at least in some OE texts, and hence identify verb position as a fairly clear marker for less embedded relative clauses.

3.2 Embedded relatives versus V2-relatives

Various syntactic tests demonstrate that V2-relatives behave like main clauses, unlike more embedded relatives. Some tests require manipulation of the data and native speaker judgements about the results that are not available to investigators of a dead language, particularly, if these judgements “are sometimes subtle” (Gärtner2001), and marked by “??” even by native speakers. The item that stands out as a potential feature to check in the OE instances is the fact that the V2-relative needs to be clause-final (p. 100) – an indication, perhaps, like the main-clause characteristics found in the German examples, that the V2-relative started out as a paratactic construction.

Note that Gärtner2001’s (Gärtner2001) paper provides a single PDE translation for his examples comparing V2-relatives with their more embedded counterparts, like (17a) and (17b), as their meanings do not differ; Gärtner’s notation (/) denotes that the V2-relative in (17a) is intonationally integrated into the previous clause, just like its more embedded relative clause counterpart in (17b); it

forms a single informational unit with the matrix clause, just like an embedded relative (p. 109). V2-relatives like (17a) are, in fact, presented as variants that are felicitous only when certain conditions are met, unlike the embedded versions – in a V2-relative, the antecedent (*eine Seite* in (17)) must be presentational, specific and indefinite:

- (17) a. *Das Blatt hat eine Seite, (/) **die** ist ganz schwarz.*
 b. *Das Blatt hat eine Seite, (/) **die** ganz schwarz ist.*
 ‘The sheet has a side that is completely black.’ (Gärtner2001)

The specificity-requirement is best illustrated by an example from ConiglioHinterhölzl2020:

- (18) a. *Hans sucht eine Frau, **die** hat blaue Augen* (*de dicto*)
 b. *Hans sucht eine Frau, **die** blaue Augen hat* (*de re, de dicto*)
 ‘Hans is looking for a woman who has blue eyes.’

In (18a), with the V2-relative, Hans is looking for a specific individual (*de dicto*), who is known to him – either directly or by hearsay – and who happens to have blue eyes; (18b), the embedded relative, can also refer to a generic entity (*de re*) – Hans might be looking for a future wife who he hasn’t met yet, and who he wants to be blue-eyed. See also Gärtner2001 for examples showing a similar contrast. The upshot of the discussion is that V2-relatives are main-clause-like, but show a degree of integration into a higher clause that is reminiscent of an embedded relative. The pivot is the demonstrative itself – and how it differs from a relative pronoun. We will return to this below.

3.3 Demonstrative-headed main clauses versus V2-relatives

A comparison between V2-relatives as in (19a) and demonstrative-headed main clauses as in (19b) (examples from (Gärtner2001) reveals subtle semantic differences in terms of the interpretation of quantification (ConiglioHinterhölzl2020):

- (19) a. *Apfeldorf hat viele Häuser, **die** stehen leer.*
 ‘Apfeldorf has many houses that are empty.’
 b. *Apfeldorf hat viele Häuser. **Die** stehen leer.²*
 ‘Apfeldorf has many houses. Those are empty.’

²ConiglioHinterhölzl2020 have a proximal demonstrative (*Diese*) for Gärtner’s distal demonstrative. Note that proximal demonstratives, in German as well as in OE, do not have relative pronoun counterparts, so they do not present the same analytic dilemma.

The interpretational difference in German seems to carry over in PDE; in (19a), with its V2-relative, many of the houses in Apfeldorf are empty, but there is no implication that Apfeldorf has many houses. The V2-relative is like the embedded relative in this respect. In the demonstrative-headed main clause in (19b), there is an assertion that Apfeldorf has many houses and that all of them are empty. Such judgements cannot be replicated for a dead language like OE, however, although it may be relevant that (19b) can also be expressed by *and* (... *und die stehen leer*).

Interesting effects appear if one considers Gärtner2001's other famous example, (17), of which deVries2012 notes that *eine Seite* has specific reference, and that this leads to a pragmatically odd reading if the second clause is not integrated into the first one (if it was, for example, separated from the first clause by *and*), as sheets tend to have two sides; de Vries concludes that the existential quantifier set up by *eine Seite* has high scope, which includes the second clause in a V2-relative (see also Gärtner2001: "one function of [V2-relatives] would be to signal wide scope of the indefinite it modifies"). The first clause ("the sheet has a side") can stand on its own if the indefinite has a non-specific reading, but this reading is unavailable because the demonstrative that heads the second clause requires a specific referent (deVries2012). Although this may not be helpful as a diagnostic for our OE examples, what one can take away here is that the close relationship between the V2-relative and the demonstrative-headed main clause has to do with the demonstrative and its interaction with the Spec,CP position, which is its most likely position in a V2 language (see also the YCOE-coding of the OE relatives presented in Section 2.4). This also explains why the V2-relatives do not show main-clause-like behavior with respect to some of de Vries' tests, which topicalize various constituents from the clause (deVries2012): The pre-field position is already filled (by the demonstrative).

Recall Gärtner2001 remark about intonation (see Section 3.2) – one would expect a demonstrative-headed main clause like (19b) to be immediately preceded by a final boundary marking (Ladd1986, cited in Gärtner2001). Final boundaries are marked in modern texts by punctuation, following the idea that a sentence, from full stop to full stop, is "[a]n utterance or complete rhetorical structure which expresses a single idea or *sententia*" (Parkes1992). The modern concept of "sentence" can, of course, contain several clauses, both main and embedded, demarcated by commas, colons or semi-colons as well as adverbs and complementizers, but one cannot rely on such meta-signs in earlier texts.

The amount of information provided by the previous clause appears to be key to the selection of V2-relatives over demonstrative-headed main clauses. The fact that, in the case of V2-relatives, the previous clause (which is the higher clause in the same sentence) appears to be presentational in nature may be the

reason why it is felt to be incomplete in terms of how informative it is, so that the subsequent elaboration pinpointing the referent triggers a more integrated structure.³ Note that observations that the referent needs to be specific but that the antecedent is often indefinite, seem to point to the denotational link being one of restriction. This is also noted by Gärtner2001: “in terms of interpretation [V2-relatives] should be treated on a par with restrictive modification.” Note that many of the V2-relatives seem to do more than restriction, and convey, in addition, a high degree of salience. In (15b), the expectation is that the sons will be the next topic in the discourse – this is part of the “topic shifting” property of the standalone demonstrative (see again example (10) in Section 2.4). This conforms to Brandt1990 (cited in Gärtner2001) observation that V2-relatives are restricted to presentational constructions that introduce an individual into the discourse. This is testable in a historical corpus.

3.4 Conclusions

Although the focus of the literature about V2-relatives in Dutch and German is not quite the same as that on the long-standing ambiguity of these structures in OE, and much of the diagnostic apparatus involved requires native speaker judgements which are not available to us, there is one aspect that seems testable: The fact that the V2-relative takes as its antecedent a referent that is newly presented to the discourse. It follows from this function that the antecedent is likely to be specific, and indefinite. In what follows, we are working from the hypothesis that these demonstrative-headed structures are typically used for what we will call “second mentions”: A protagonist has been introduced, and the *se*-clause “takes up” the story – either by providing relevant background information or a foregrounded event. This may not lead to an answer to the question of the exact clausal status – main or embedded? – but it might go some way to explain why we find V2-like word orders in a set of clauses that have received the label “IP-SUB” in YCOE, what function V2 serves, and what a paratactic construction with a standalone demonstrative can tell us about the genesis of embedded relative clauses.

³Gärtner2001 refers to one of his examples of V2-relatives “modifying a preposterously vacuous matrix clause”.

4 Methodology

4.1 Texts

Based on the discussion of Section 3, we will focus on the presentational function of the referent of V2-relatives. If the V2-relative picks up a newly introduced, specific entity in the modern West Germanic languages, this might also be its function in OE. As introducing and maintaining protagonists is very much a feature of narrative texts, we will be investigating these “second mentions” in Ælfric’s *Lives of Saints* (Skeat1966, 1881–1900) as our primary text. The *Lives of Saints* is a collection of 37 (mostly) narrative texts in Old English, the only narrative collection of some length (100,193 words) in YCOE that is not a direct translation from Latin. Although Ælfric must have relied on Latin works for their contents, his views on the translation process, explicitly stated in his preface to the translation of Genesis, make it unlikely that he would opt to stay close to the text of his Latin or Greek *Vorlage*, and follow it slavishly (Koopman1992). The narratives in the *Lives of Saints* are not a random collection but “the product of systematic thought” (Hurst1972). Ælfric’s linguistic awareness is evident from the prefaces to his works, where he emphasizes that he intends to write in plain language, and it is clear that he thought of his work as a “consciously ‘literary’ act” (Clemoes1956).

Although the main focus of the investigation lies on Ælfric’s *Lives of Saints*, one other text in YCOE was investigated in order to see whether the findings from *Lives of Saints* could be replicated elsewhere. This second text was a translation from Latin (unavoidable as Ælfric’s *Lives of Saints* is the only untranslated narrative text in OE): *Gregory’s Dialogues* (ed. Hecht, 1900–1907). This is a collection of legends about early Italian saints, which makes its subject matter comparable to Ælfric’s *Lives of Saints*. Bishop Wærferth’s original translation of the text was undertaken some time between 870–890, so at least a century before Ælfric’s *Lives of Saints*, but we used the H ms, which, comprising 25,593 words, is shorter than the earlier C manuscript, and also, very unusually for the period, not a straightforward copy of the C text but a thorough re-working which alters the syntax; clauses are grouped together into *sententiae* in a more coherent fashion (Yerkes1982), and the position of finite verbs is often changed. This makes it likely that the H ms is less dependent on its Latin *Vorlage* than C. This revision is also more likely to be contemporary with Ælfric’s work, as it was probably produced a century or a century and a half after Wærferth’s translation (Yerkes1982).

4.2 Analysis

The various types of relative clauses (see Section 2) were extracted by means of CorpusSearch 2 (Randall2005); cases where the relative clause provides an explanation of a term or phrase (usually in Latin), somewhat akin to the equivalent of *id est*, like (20), were noted but excluded from the analysis, as the relative pronoun in such cases is invariably *þæt* ‘that’ and is more like a meta-comment than a relative clause:

- (20) *sum wæs arwurðes lifes wer, se wæs mid Godes gife*
 some was honourable.GEN life.GEN man SE was with God.GEN gift
gebledsod & be naman genemned Benedictus, þæt is on Englisc se
 blessed and by name called Benedictus SE.NOM.N is in English the
gebletsoda.
 blessed
 ‘There was a man of honourable life who was blessed with God’s gift and
 called by the name of Benedictus, which is in English “the blessed one”.’
 [cogregdH,GDPref_2_[H]:94.23.971]

Also excluded from the analysis were demonstratives as complements of prepositions, like (11) above.⁴ The remaining relative clauses were annotated for the presence of a D-head in Spec,CP with an empty (“zero”) C, following the coding in YCOE. When we discovered that no human personages were referred to by the neuter pronoun *þæt*, in spite of what we might expect of a language like OE with grammatical rather than natural gender where *wīf* ‘woman’ and *mægden* ‘girl’ are neuter nouns, the D-head was split up according to masculine/feminine forms (denoted by “SE” in the analysis) versus neuter (“THAT”).⁵

We also coded for second mention, early verbs (i.e. V2, main clause word order), and the presence/absence of *þe*. The definition of “early verb” was a fairly restricted one, to exclude cases where there was too little material in the clause

⁴This was mainly done in an effort to keep numbers manageable; our impression of the ca. 40 instances in Ælfric’s *Lives of Saints* is that these PPs conform to the *se*-relatives we investigated in every respect.

⁵This is not due to a lack of female protagonists, as there are many. The main reason appears to be the way protagonists are introduced – the *se*-relative often follows a name:

- (i) *Sum wif hatte Sintice, seo wæs six gear blind*
 Some woman was.called Sintice SE was six year blind
 ‘There was a woman called Sintice, who had been blind for six years.’
 [ÆLS_[Thomas]:263.7705]

to make a judgement, or where the early position of the verb could be due to extraposition of a heavy constituent like a clause. Some other variables were also tracked, namely, whether the relative clause was part of a set of relative clauses in a single sentence, which might have influenced their word order, or whether the antecedent was a divine being that would not require an introduction.

5 Findings

5.1 *Se*-relatives in in Ælfric’s *Lives of Saints* and Gregory’s *Dialogues*

The numbers of Ælfric’s *Lives of Saints* and the OE translation of Gregory’s *Dialogues* (Ms H) are provided in Table 2. It is clear that the proportion of *se*-relatives is much larger (28% of the total) in Gregory’s *Dialogues* than in Ælfric’s *Lives of Saints* (17%); we do not want to make too much of this difference, although an appeal to chronology might fit a narrative in which Gregory’s *Dialogues*, the earlier text, relies more on *se*-relatives not marked by the universal embedder *þe* than the later *Lives of Saints* because the *þe*-construction is more grammaticalized and a more recent development; nailing this point would require a much more extensive investigation than the present one. The purpose of this paper is to compare the *se*-relatives with “early verbs” to what has been said about the V2-relatives in Section 3.

Table 2: *Se*-relatives in Ælfric’s *Lives of Saints* and the OE translation of Gregory’s *Dialogues*, ms H

	Ælfric’s Lives	Gregory’s Dialogues, H
all relatives	1261	373
of which <i>se</i> -relatives (excl. <i>þæt</i>)	215	106
of which no <i>þe</i>	108	80
of which 2 nd mention	100	60
of which also early V	94	39

As an illustration of the “second mention” function, Table 3 presents the first 10 instances in the data where the conditions “D=SE”, “second mention” and “early verb” are all met; for reasons of accessibility, the data are presented in Skeat’s PDE translation, while following the YCOE sentence boundaries. The first sentence has the *who* or *which* in bold that translates an OE *se*-form heading a V2-relative.

Note that they all follow the introduction of a new referent, which is why they were coded as “second mentions”. The second sentence illustrates the referent’s persistence in the discourse. The only referent who does not “persist” is Hilarion (#7); Paschasius in #5 remains activated and persists in the discourse, although he happens to be absent in the second sentence.

It became very clear early on that there is a difference between divine referents (God, Christ) and other referents – the divine referents are rarely “presented” but assumed to be known, and hence rarely qualify as “second mentions”. When referred to by a *se*-relative rather than a *þe*-relative, they do show higher rates of “early verbs”, however, so that in this they follow the trend of *se* being associated with V2-like word order; these instances have been marked “Divinity” in Figure 2 below. Figure 2 clearly shows that the triad “*se*-relative/second mention/early verb” is a coherent package in Ælfric’s *Lives of Saints* (dark purple) if we keep “divine referents” separate.

A second look at the items that are “second mention” but do not have an “early verb” (light purple in Figure 2) reveals that all six of these in Ælfric’s *Lives of*

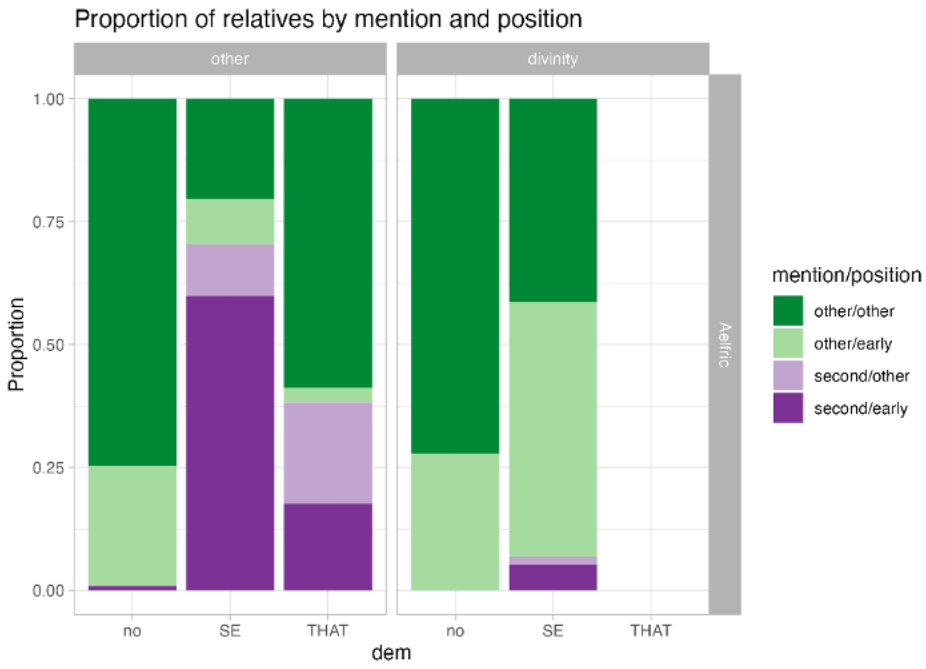


Figure 2: Proportion of demonstratives by mention and by position in Ælfric

Table 3: A random sample of the data coded as “D=SE”, “second mention” and “early verb” in Skeat’s PDE translation. The *se*-form (expressed by *who*, *which*, *he* and *this* in this translation from Skeat, in bold) with V2 encodes a second mention of a newly introduced referent in the first sentence; their degree of persistence can be gleaned from the referent re-appearing in the second sentence (also in bold).

#	first sentence	next sentence
1	A certain nobly-born thane was named Philip, whom the emperor Commodus sent [...] to the city which is named Alexandria [Eugenia 5.198]	This thane Philip was not baptized unto God
2	Another man was also blind for seven full years; he had a guide who led him everywhere. [Swithun 202.4349]	One day he went out as he often did, and the guide became angry, and left the blind man, and ran away, and the other knew not how he should come home
3	There was a certain woman called Syntyche, who had been six years blind, and was then healed by the holy apostle, and came, seeing clearly, unto her kinswoman named Migdonia, who had left her blind. [Thomas 263.7705]	Then said Migdonia: ‘This man is God Himself, or God’s angel, who hath enlightened thine eyes thus without leechcraft.’
4	Likewise Athelwold, the holy bishop, who now worketh miracles through God, often told us, that he knew a man with bishop Ælfheah, who would drink in Lent whenever it pleased him. [Ash_Wed 65.2738]	Then one day he prayed the bishop Aelfheah to bless his cup; he would not, and the fool drank without blessing, and went out.
5	This came to the ears of the nobly-born youth who was wooing Lucy, who was named Paschasius, an impious idolater, who enticed the holy maid to make offerings to devils; [Lucy 57.2202]	but the Lord’s virgin said, ‘A pure offering is this, and acceptable to God, that one should visit widows, and comfort exiles, and help orphan children in their affliction.

#	first sentence	next sentence
6	How there are eight Chief Sins, which sorely fight against us: One is called Gula, that is Gluttony in English, which maketh a man eat and drink before the time, or again to take too much in food or in drink. [Memory_of_Saints 268.3471]	This destroyeth both soul and body, because it bringeth upon the man much sickness, and bringeth him to death through immoderate drinking;
7	and turned him thence to a certain holy man [who was] called Hilarion, with some of his men who would not leave him. [Agnes 381.1985]	Four estates he gave up entirely, together with himself, for the reception of strangers and for alms-deeds.
8	and the Savior provided, after He had ascended to Heaven, that He should send to the king, as He had before spoken, one of the seventy whom He had chosen to preach, who was called Thaddeus, that he might heal the king. [Abdon_and_Sennes 124.4795]	He came then, by God's commission, to the aforesaid city, and healed the afflicted king in the Savior's might, so that the citizens greatly wondered thereat.
9	Again there was a leader, named Seron, in the land of Syria, who quoth to his people, 'I will get me a name and overcome Judas, and them that are with him, who despised the king.' [Maccabees 298.5031]	He gathered then his host, and went with great array to Judea-land, and many people with him.
10	Again, his son Jove, whom ye worship as a god, who desired to kill his unclean father that devoured his brothers as soon as they were born, this Jove was filled with foul lust, [Sebastian 172.1310]	and [Ø] took his own sister to his unclean wedlock, even as ye read in your histories.

Saints were not coded as such because of the restricted definition of what we coded as “early verb” (see Section 4.2); they all have the finite verb in second place, but either do not contain enough constituents to rule out a different analysis, or contain an object of such length that it can be argued to have been extraposed, making the position of the finite verb analytically ambiguous.

The situation with respect to verb-placement is not as clear-cut in *Gregory’s Dialogues*. There are 21 cases of *se*-relatives that did not meet our criteria for “early verb”; three of them have pronominal subjects and surface V3, so can be claimed to have main clause order (see Section 2.4). Of the remainder, there are only three cases which have been excluded because of the possibility of extraposition; the rest all have material intervening between the D-pronoun and the finite verb, one of them an anacoluthon, where a *se*- is started but not finished, so that there is a subject *se* but no finite verb:

- (21) *On þam þinge ic oncnawe, þæt Benedictus hæfde Paulus gewrixle,*
 In that thing I understand that Benedictus had Paul to.example,
se þa þa he þolode scipes forwyrð & lyre eallra
 SE.NOM.M then then he suffered ship.GEN wreck and loss all.GEN
þara þinga, þe þæar on wæron, þa onfeng he sylf him to
 the.GEN things.GEN that there on were then received he self him as
frofre eallra þara lif, þe him mid ferdon.
 comfort all.GEN SE.GEN.PL lives that him with travelled
 ‘In which thing I see that Benedict imitated St. Paul: who, when he
 suffered the wreck of a ship and the loss of all the goods that were on it,
 then he himself received, for his comfort, the lives of all that were in his
 company.’ [cogregdH,GD_2_[H]:17.141.10.1370]

In terms of the observation that a German V2-relative needs to be clause-final (Gärtner2001), this was also a consistent finding in Ælfric’s *Lives of Saints* – all V2-*se*-relatives were clause-final. In the OE translation of *Gregory’s Dialogues*, cases like (22) occasionally occur:

- (22) *& he þa on þære ilcan stowe, þe ic bufan ymbe spræc, seo is*
 and he then in the same place that I above about spoke SE.NOM.F is
geciged Subpentoma, manega gear his lif adreah on halgum dædum
 called Suppentonia many years his life conducted in holy deeds
 ‘and he then in the same place that I mentioned above, that-one is called
 Suppentonia, for many years conducted his life by holy deeds.’
 [cogregdH,GD_1_[H]:8.52.9.494]

The *se*-clause is not clause-final; it appears to resemble a PDE deaccented parenthesis (Gärtner2001), and it is quite likely that these *X is called*-phrases are formulaic chunks. Note that this case was not coded as a “second mention”, as the place in question has already been presented in the previous discourse; the most we can say is that it is reactivated here. There is another case in the same text of a *se*-relative that is not clause-final. This one is a “second mention”, but note that it is also anacoluthic; the subject of the higher clause, *Exhilaratus ure gefera* ‘Exhilaratus our companion’ is the antecedent of the *se*-relative, which explains why the *se*-relative immediately follows it; but this apparently causes such a disruption to the flow of the discourse that this subject is restarted after the *se*-relative by means of a resumptive pronoun *he* (also in bold):

- (23) *Eac hit gelamp on sume tid, þæt Exhilaratus ure gefera, þone*
 also it happened at one time that Exhilaratus our companion SE.ACC.M
þu canst gecyrredne to rihtre drohtnunge, he wæs fram his hlaforde
 you know converted to right conduct he was by his master
asended to þam mynstre, þæt he sceolde þam Godes were
 sent to the monastery that he should the.DAT God.GEN man.DAT
Benedicte beran twa treowene fatu wines fulle, þa syndon on
 Benedict bear two wooden bottles wine.GEN full SE are in
folkisc flaxan gehatene.
 vernacular flagons called
 ‘Upon a certain time, Exhilaratus our companion, whom you know as
 having been converted to the right conduct, was sent by his master to the
 monastery of the man of God, to carry him two wooden bottles,
 commonly called flagons, full of wine.’
 [cogregdH,GD_2_[H]:18.141.19.1372]

It is clear from Table 2 that a minority of *se*-relatives in these texts did not “qualify” as second mentions. One of them is (24), which is a reactivation of a referent who has been mentioned earlier, *se ylca broðor* ‘that same brother’:

- (24) *þa ær þæs Anastasies seofeðan dæge, wæs eac forðfered se*
 Then before that.GEN Anastasius’ seventh day was also departed that
ylca broðor, se swapeah næs na geciged on þære nihte
 same brother SE.NOM.M however not-was not named in that night
betweoh þa oðre broðru
 among the other brothers
 ‘Then before that Anastasius’ seventh day was also departed that same

brother, that one, however, was not named that night among the other brothers.’ [cogregdH,GD_1_[H]:8.53.28.510]

Example (24) has not been coded as an “early verb” because *swaþeah* ‘however, nevertheless’ intervenes. Reactivation after a diversion is an established function of *se* (also for names: *se Cynewulf* ‘This Cynewulf, i.e. the one we were talking about earlier’ in the *Cynewulf and Cyneheard* interpolation in the Anglo-Saxon Chronicle [ChronA, Plummer 755.1–38]) and of demonstratives in general. (26) is another example of a referent already having been mentioned – in this case, many times before, seeing that the referent is Martin, the protagonist of this “life”. This example does not have an “early verb” either:

- (25) *and Godes sacerdas synd gewuldrode mid þære onwrigennysse*
 and God.GEN priests are glorified by the revelation
Martines forðsiðes, þonne se halga Michahel mid englum
 Martin.GEN departure.GEN SE.ACC.M the holy Michael with angels
underfeng and Maria seo eadiga mid mædenlicum werodum, and
 received and Maria the blessed with virginal companies and
neorxnewang gehylt bliðne mid halgum.
 paradise holds happy among saints
 ‘and the priests of God are glorified by the revelation of Martin’s
 departure, whom the holy Michael with his angels and blessed Mary with
 companies of virgins received; whom paradise holdeth, happy among
 saints.’ [ÆLS_[Martin]:1435.6918]

Of the five *se*-relatives in Ælfric’s *Lives of Saints* that have “early verbs” but do not qualify as second mentions, four form a coherent group as a favourite ending of a saint’s life, referring back to a divine antecedent with variations on the phrase *þam is wuldor and wurðmynt on ealra worulda woruld AMEN* ‘to him is glory and honour in world without end, amen’ (for example, [ÆLS [Peter’s_Chair]: 292.2467]; these formulae have a Latin model.⁶ The fifth is a case of reactivation, with *þone foresædan Tattheum* ‘the aforementioned Tatheus’ as its referent ([ÆLS [Abdon and Sennes]:132.4799]); see (21) above.

Others do not qualify because although they are part of a follow-up to a presentational clause, they are “third” rather than “second” mentions, as the “second mention proper” is the preceding *se*-relative that gives the priest’s name:

⁶We thank one of the reviewers of this paper for pointing this out to us.

- (26) *Ða wæs þær sumre neah cyricean mæssepreost, þam wæs
 then was there some.GEN nearby church masspriest SE.DAT.M was
 nama Florentius, se wæs Florenties ures subdiacones yldra
 name Florentius SE.NOM.M was Florentius' our subdeacon.GEN older
 fæder.
 father*

‘Then was there of a nearby church a masspriest who was by name
 Florentius. He was the father of our subdeacon Florentius.’

[cogregdH,GD_2_[H]: 8.117.2.1151]

For the statistical analysis, we fitted a Bayesian linear model using a Poisson distribution for the outcome variable (number of relatives), with demonstrative, second mention and early verb as predictors (including interactions thereof); visualizations of the model results are provided for *Ælfric's Lives of Saints* (Figure 3) and *Gregory's Dialogues* (Figure 4). The figures show the predicted percentage of occurrence of different demonstratives depending on mention (other vs second) and verb position (early vs other). Different Credible Intervals (CrIs) are indicated in the figures by means of different shades of blue: from lighter to darker, 50, 70, 80 and 95% CrIs. A CrI of, say, 90% indicates that the predicted percentage is within that interval at 90% certainty, based on the data and model. Larger intervals have greater certainty than smaller intervals (with greater precision in the predicted percentage there is greater uncertainty).

In general, *se* with second mention and early verbs tends to occur at the same rate in *Lives of Saints* and *Gregory's Dialogues* but, while this stands out in *Ælfric*, this is not so much the case in *Gregory's Dialogues* – *se* also occurs more or less as frequently with verbs in non-early positions, as shown above, and in cases where they do not qualify, strictly speaking, as second mentions. The difference between *Ælfric's Lives of Saints* and the H ms of *Gregory's Dialogues* in how consistently they use *se*-relatives may be due to the fact that the former reflects a more sophisticated style by a very diligent author (*Ælfric* is labelled a “conscious stylist” in Hurst1972) while the latter is a fairly pedestrian translation of a Latin text, which shows its origins in spite of the fairly substantial editing of the H ms.

5.2 Demonstrative-headed main clauses in *Ælfric's Lives of Saints*

To put these findings in perspective, we looked at *se*-clauses in *Ælfric's Lives of Saints* that are coded as demonstrative-headed main clauses rather than

relative clauses in YCOE.⁷ There are 43 cases of demonstrative-headed main clauses that might possibly qualify as meeting most of the conditions for relative-clause-status in that they are not connected to the preceding clause by *and*, and the demonstrative is the first element in the clause. Following **Schlachter2012**, who found that capital letters were systematically used in Old High German manuscripts to signal sentence beginnings, including demonstrative-headed main clauses, we investigated whether capitalization might provide any clues to clausal status. Comparing images of the ms made available on the website of the British Library to the edition of *Ælfric's Lives of Saints*, on which the text in YCOE is based (Skeat), it appears that Skeat follows the capitalization of the ms quite faithfully, using it to demarcate the text into what are at times quite large stretches of discourse, paragraphs rather than sentences. The YCOE sentential coding follows this system. An example is (27), where both *ses* are

⁷Note that here, too, *se*-clauses refer to masculine, feminine and plural demonstratives, and exclude neuter forms of *this* and *that*, and demonstratives as complements of prepositions.

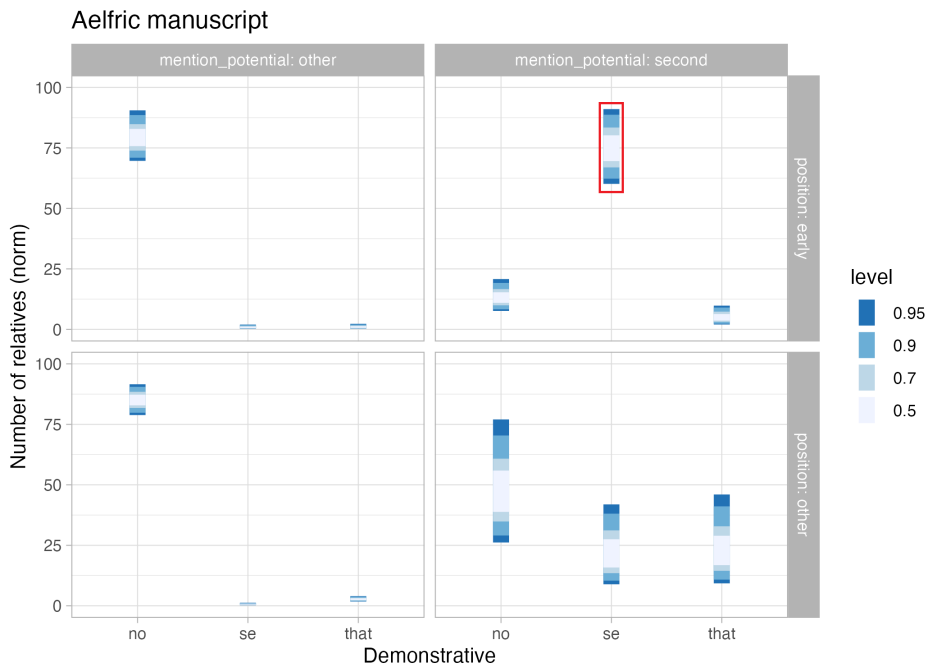


Figure 3: *Ælfric's Lives of Saints*, Predicted Credible Intervals (50, 70, 90, 95%) of the posterior probability distribution from a Bayesian model fitted to number of occurrences of demonstratives

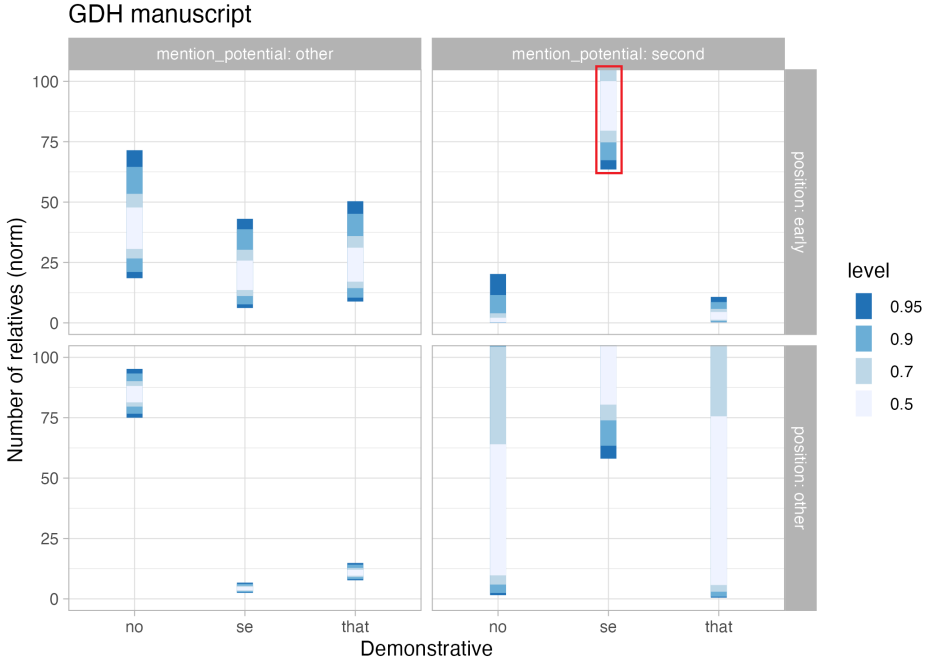


Figure 4: Gregory's Dialogues ms H, Predicted Credible Intervals (50, 70, 90 and 95%) of the posterior probability distribution from a Bayesian model fitted to number of occurrences of demonstratives

capitalized, and follow a sentence-code; note that the second *se* would not qualify as a “second mention”; instead, it is reactivating the referent of the first *se* after the narrative is interrupted by background information – reminiscent of Schlachter2012's (Schlachter2012) finding that antecedents of these clauses can be quite distant, and that the demonstrative-headed clauses can apparently select and reactivate topics, which cannot be achieved by ordinary relative clauses.

- (27) *Saul hatte se forma cyning þe ofer Godes folc rixode.*
 Saul was.called the first king who over God.GEN people ruled
Se wæs to cynincge ahafen swyðor for folces gecorennysse
 SE.NOM.M was to king raised rather for people.GEN choice
þonne ðurh Godes ræd. Fela oðre cynincgas rixodon ær
 than by God.GEN counsel many other kings ruled earlier
geond ealne middaneard ofer hæðenum leodum, ac ofer
 throughout all middle-earth over heathen peoples but over

Israhela folc þe on God belyfde næs nan eorðlic cynincg
Israel.GEN people who in God believed not.was no earthly king
ærðan þe Saul, swa swa hi sylfe gecuron, ofer hi cynerice
before that Saul as as they self chose over them kingdom
underfencg. Se beah hrædlice fram þæs ælmihtigan Godes
received SE.NOM.M bent soon from the almighty God.GEN
willan and nolde be his wissunge and be his witegan lare
will and not.wanted by his guidance and by his prophets.GEN teaching
faran, and se yfela gast hine drehte mid deofollicum sticelsum, and
go and the evil spirit him troubled with diabolic instigations and
on ungewitte his mod awende.
into madness his mind turned

‘Saul was the name of the first king who reigned over God’s people. He was raised to be king rather by the people’s choice than by God’s counsel. Many other kings had reigned before throughout the whole world over heathen nations; but over the people of Israel, who believed in God, there was no earthly king before that Saul (as they had themselves chosen) assumed the dominion over them. He turned quickly aside from the will of Almighty God, and would not walk by His instruction and by the teaching of His prophets, and the evil Spirit troubled him with diabolic instigations, and turned his reason into madness.’ (tr. Skeat).

[ÆLS (Book of Kings) 1–8]

All of the 23 cases of capitalized *se* in *Lives of Saints* have early verbs; all but two are the first word of a “sententia”, which is often close to the beginning of the chapter, as in (27). This makes sense as the protagonist is likely to be introduced there, and once introduced, will be referred to by a form of *se*. This completely parallels the “second mention” *se*-relatives with early verbs; the difference is that the preceding chunk in the case of these demonstrative-headed main clauses has been informative enough to warrant “sentence-hood” (cf. Gärtner’s observation that the choice between a V2-relative and a demonstrative-headed main clause appears to depend on “information chunking”, see Section 3.3).

Note that of the two cases where capitalized *se* appears inside a “sententia”, one is the first word of direct speech, coded as a standalone demonstrative, like the other capitalized standalone demonstratives, while the other one, (28) below, is coded as a relative clause. Example (28) is part of the following story: A widow realizes that she is “destroying herself by deadly sins” and writes all her wicked deeds on a piece of paper, which she takes to the holy man Basil, imploring him to blot her sins out. He keeps vigil that night to pray for her:

- (28) *Git þa Basilius gebæd for þæt wif. Waciende þa niht, and þæt*
 Yet then Basilius prayed for the woman waking the night and that
gewryt ageaf þam foresædan wife, and þa wæron
 writing gave.back the.DAT aforementioned woman.DAT and then were
þa synna ealle adilegode butan anre synna, Seo wæs seo mæste.
 the sins all blotted.out except one sin SE.NOM.F was the greatest
 ‘Still Basil prayed for the woman, keeping vigil that night, and gave back
 the writing to the aforesaid woman, and then were the sins all blotted
 out, save one of the sins, which was the greatest.’
 [ÆLS [Basil]:549.848]

Note that we have arrived at the “central reportable event” (Labov1972) of the story. In a PDE translation, this important turn would have been marked by “However”, to mark the contrast that at first, the saint seems to have succeeded: The slate of the woman’s sins has been wiped clean – but it turns out that there is still one sin left. The start of a new sentence rather than just a new clause inserts a pause before the devastating turn that the one sin that is not blotted out also happens to be the most reprehensible one. A demonstrative-headed main clause interpretation seems warranted here, even though it is not coded as such in YCOE.

Two of the remaining 19 clauses are preceded by left dislocations, as part of a correlative generic expression: *þa ðe fram gode bugon to bysmor-fullum hæðen-scype þa wurdon gescynde* ‘as for those who turn from God to shameful idolatry, those were put to shame’ [ÆLS [Book of Kings]:43.3687]. The remaining 17 are all “second mentions” of referents that have just been introduced – three are expressions like *sum þara* ‘some of those’ or *an þara* ‘one of those’ which find their referents in the immediately preceding discourse; eight follow names, four follow indefinite expressions (*sum earm ceorl* ‘some poor man’ [ÆLS [Swithun]:97.4274]; *ænne blindne mann* ‘a blind man’ [ÆLS [Denis]:51.5817]; *micel folc manna* ‘a great multitude of people’ [ÆLS [Martin]:1015:6626] and two follow what we would label definite expressions, although they refer to entities newly brought onto the scene: Christ’s holy angel in *Asende ure Hælend Crist his halgan engel mid þe* ‘Our Savior will send his holy angel with you’ [ÆLS [Apollinaris]:25.4544] and *an God* ‘one God’, who is introduced in direct speech in a successful attempt to convert an unbeliever [ÆLS [Cecilia]:163.7213]. These new referents all persist in the subsequent discourse.

These *se*-clauses, then, are clearly of the same form (distal demonstrative as initial element in a V2-clause) and the same function (second mentions, as well as, in a few cases, reactivations) as the *se*-clauses that were coded as relatives.

Half of them appear in separate “sententiae” and hence offered clear signs to the annotator of their main clause status, while the other half were preceded by clauses that were apparently informative enough to prompt an analysis of “complete sentence”.

6 Discussion

One of the interesting outcomes of the comparisons of V2 relatives with embedded relative clauses is the possible light they cast on the diachrony of how the latter developed out of demonstrative-headed main clauses: The embedded relative shows signs of grammaticalization in that they are more versatile than V2-relatives and demonstrative-headed main clauses in terms of the *de re/de dicto* contrast reported in Section 3.2 (examples 18a–18b). The pivot here is the change from demonstrative to relative pronoun: The demonstrative has to refer to an entity that is specific, unlike the relative pronoun. If one follows the observation in RobertsRoussou2003 about the kind of semantic bleaching which accompanies the grammaticalization process, it can be argued that the more general use of the relative pronoun indicates that it is more bleached than the demonstrative. Roberts & Roussou argue that bleaching, often described as the loss of lexical content, would be more accurately described as the loss of non-logical content. Logical content is “insensitive to specific facts about the world” (Fintel1995, in RobertsRoussou2003); the meaning of a functional element is predictable and stable, that is, independent of the (lexical) items it scopes over. It follows that functional elements have more general applications than lexical elements; see also Lehmann1995 (Lehmann1995 [1982])’s parameter of “paradigmatic variability” as a metric for quantifying the degree of grammaticalization. The value of the measure increases when restrictions imposed on the contexts in which the functional element can appear are dropped (Lehmann1995 [1982]: 141), or if there is an increase in its distribution (p. 142). The relative pronoun of a properly embedded relative clause, then, can be argued to be more grammaticalized than the demonstrative.

Another difference between embedded relative clauses and V2 relatives indicating that the former are more grammaticalized is their ability to express backgrounded material, that is, information that has already been expressed explicitly in the previous discourse (from Gärtner2001):

- (29) a. A: Kennst du jemanden, [*der* ein Fahrrad besitzt]?
b. ??B: Ja, ich KENNE jemanden, (/) [*der* besitzt ein Fahrrad].

(29b) is fine for embedded relative clauses, with the answer just echoing the question, but degraded for V2-relatives (Gärtner2001). This observation also points to the greater versatility of embedded versus V2-relatives (and demonstrative-headed main clauses).

The embedded relative clause is part of a general development along a cline of clause linking, with two maximally elaborated paratactic clauses at one end and a single clause a compressed, embedded complement at the other end (Lehmann1988). Movement along this cline is not necessarily unidirectional: Although

Lehmann1988's examples show movement towards compression (involving his parameters of hierarchical downgrading, the main clause syntactic level, desententialization and explicitness of linking), nonfinite/infinitival clauses, which are positioned towards the compressed end of the cline on all parameters, are much more likely to represent "upgraded" nominalizations (Los2007). In the case of relative clauses, there is a consensus in the literature that "downgrading" is at play. For OE, Mitchell1985 specifically points to the V2-position of the verb as a main-clause feature inherited from the paratactic origin of the construction. Parataxis also accounts for the requirement that the V2-relative must be in clause-final position (see Section 3.2). The frequency of such "extraposed" relatives shows a diachronic decline in a number of European languages (Wallenberg2016), suggesting that relative clauses, earlier paratactically linked to the preceding clause, become more integrated into that clause over time; the V2-relatives in Dutch, German and OE might reflect an intermediary stage in the process of parataxis to hypotaxis, poised between restriction (an embedded function) and foregrounding (a main clause function). It is typical for grammaticalization that both the original construction as well as any intermediary stages may coexist with the maximally grammaticalized construction ("layering", HopperTraugott2003).

Cases like (29) above are discussed by Gärtner as part of a discussion of "assertionality" (Gärtner2001) – as is well-known, main and subclauses in language generally tend to align along the assertion/foregrounding versus the presupposition/backgrounding axis (see, for example, Cristofaro2003). Whether, and how, the alignment is marked overtly (by word order like V2, by indicative versus subjunctive, or by other means) will depend on diachrony, and may be the result of competition between which function wins out as the first item in the construction of utterances – assertion (marking realis versus irrealis) or foregrounding (marking the main event of the plotline out from any embedded events). Assertion-first may provide reasons or explanations for a foregrounded event embedded under the assertion, even though reasons and explanations are background in terms of discourse structure; and the natural rapid churn of narrative

devices to signal peak marking, episode boundary marking, or the creation of suspense are likely to lead to the development and subsequent decline of a multitude of expressions that prioritize one alignment over the other, so that either foregrounding or assertion ends up being relegated to syntactically embedded positions (Los2015).

If main clause assertion is used as a frame for a particular narrative device, the foregrounded event follows in an embedded clause. Examples are episode boundary marking in Middle English by means of the frame *It befell that* (Brinton1996), see also *hit gelamp* ‘it happened’ in (23) above; or the use of the *it*-cleft, specifically in the frame *It was then that*, to mark “the central reportable event” in a story (Los2015: 219, 242). A third example is sentential relatives being used as foregrounding devices (HuddlestonPullum2002: 1148). A preponderance of one alignment over the other might in time lead to a particular kind of marking (see, for example, GärtnerEythórsson2020, reported in ConiglioHinterhölzl2020: 227).

Syntactic developments may obscure established patterns of marking main and subclause asymmetries from one period to the next – in the history of English, for example, finite subjunctive clauses expressing “dependent desires” (the complements of verbs with meanings of fear, promise, order, hope, expect and the like) were ousted by non-finite (*to*-infinitival) clauses (Los2005), obscuring any alignments that may have been established between assertions and indicatives or V2, versus irrealis and subjunctives or verb-final. An example of a pocket of misalignment in PDE are conditional clauses where C can be expressed not only by a “bespoke” element like *if* but also by finite verb movement to C, now restricted to a dwindling set of lexemes (the auxiliaries *had* or *should*, RobertsBiberauer2017) – a case where finite verb movement to C marks an embedded clause.

It is also a common finding that subclauses tend to preserve older orders, whereas main clauses tend to innovate; the reason appears to be that main clauses have to satisfy various communicative requirements, the positioning of focus and discourse-old or discourse-new material, and they therefore tend to develop special constructions not found in the subclause (see Bybee2001). The finite verb may possibly have functioned as a focus marker first, as it still does in Hungarian (Comrie1997), and may later have become entrenched as a clause-typing device. Any asymmetry developing between main and subclause in terms of whether or how the clause-typing head, say C, is expressed is probably a secondary or epiphenomenal development.

7 Conclusion

This paper has investigated whether more can be made of the observation in the literature that a certain type of relative clause, specifically the *se*-relative, is found more often with main clause word order (in terms of finite verb movement to C or to a lower head) than other types, particularly the *þe*-relative. The latter also shows other signs of being more integrated into the previous clause than the *se*-relative. We have argued that the *se*-relative in OE exhibits similarities with the “V2-relatives” in modern Dutch and German, although some of the clauses coded as “relative” in the OE corpus can also be analysed as demonstrative-headed main clauses. As verb position is not a clear diagnostic of clausal status in OE, we looked at a discourse function, that is, whether or not the *se*-clause expresses a “second mention” of a main protagonist. Both texts showed the same configuration of *se*-clause, second mention and V2-order as a coherent package, although the use of the V2 *se*-relatives as a “second mention” device is most consistent in Ælfric’s *Lives of Saints*, rather than in the OE translation of Gregory’s *Dialogues* (ms H).

There are many issues that could not be addressed in the scope of this paper; as avenues for further research, the analysis of V2-*se*-relatives in other OE texts would be a next step. Another outstanding issue is how the function of restriction operates in *se*-relatives and in *se*-headed main clauses, and how it interacts with the foregrounding function of the *se*-clause. In both *se*-relative and in *se*-headed main clause, the topic-shifting function of *se* coerces identification with a new referent who has been introduced in the previous clause; although indefinite, the referent is already specific when first introduced, unlike the type of antecedent usually associated with restrictive relative clauses; there is no sense of the *se*-clause identifying a subset from a larger set. In Section 3.3, it was shown that Gärtner2001 identifies the function of V2-relatives as restriction, although the *se*-relative in OE is said to be mainly non-restrictive in the literature (see Section 2.4). This observation might be explained by the fact that they often follow proper names in OE (as in (i) in footnote 5), even if those names cannot be said to be the antecedent (which in (i) would be *sum wif* ‘a woman’), although the antecedent in all of these cases is a specific entity. See also the discussion in DenDikken2003 about a Dutch example where the first clause translates to *There were two boys on the beach*, where the question is whether this entails that there were only two boys on the beach, or more, and the same question might be asked about the two sons in (15). The PDE counterparts of the “second mention” relatives capable of performing the required topic-shifting function would be non-restrictive relative clauses (see, for instance, the discussion of the Columbo-

example (10) above), which are, in fact, typical foregrounding devices in PDE. A provisional conclusion might be that restriction is just not relevant here (see also **DenisonHundt2013**), and that the primary function of V2-*se*-relatives is presentational.