

Which-hunting in Medieval England

ROBERT TRUSWELL
University of Edinburgh
rob.truswell@ed.ac.uk

and

NIKOLAS GISBORNE
University of Edinburgh
n.gisborne@ed.ac.uk

Abstract

In many of the first English headed *which*-relatives, *which* has an NP complement. Using distributional tests grounded in contrasts revealed by research in formal semantics, we demonstrate that the presence of an NP complement forces a nonrestrictive interpretation of the relative, while ‘bare’ *which*-relatives may be restrictive or nonrestrictive. We situate this finding in relation to both the formal semantics of relative clauses, and the history of *wh*-relatives in English.

Keywords: relative clauses, Middle English, discourse anaphora, *wh*-phrases, parsed corpora

Résumé

Dans les textes du moyen anglais, on constate que pour les phrases relatives avec antécédent qui contiennent le pronom relatif *which* (*Qu*-), ce pronom a souvent un complément SN. À l’aide de tests distributionnels basés sur les contrastes révélés par la sémantique formelle, nous démontrons que la présence de ce complément SN force une interprétation non déterminative de ces phrases relatives, alors que des phrases relatives avec *which* mais sans complément peuvent avoir des interprétations soit déterminatives, soit non déterminatives. Nous situons ce constat par rapport d’une part aux études sur la sémantique formelle des phrases relatives, et d’autre part à l’histoire des phrases relatives-*Qu* en anglais.

Mots-clés: phrase relative, moyen anglais, anaphore discursive, syntagme-*Qu*, corpus annoté

Portions of this work were presented at the UCL workshop on movement (2015), a linguistics colloquium at the University of York (2016), and the first Formal Diachronic Semantics workshop (Konstanz, 2016). Thanks to those audiences, to the guest editors of this volume (particularly Igor Yanovich for copious comments on a prefinal version), and to two anonymous reviewers.

1. INTRODUCTION

There is a peculiar disconnect between formal semantics and diachronic semantics. Formal semantics, like other areas of theoretical linguistics, is primarily concerned with ‘hidden’ aspects of grammatical representations: everyday discourse doesn’t immediately reveal constraints on scope relations, or anaphora, or other core semantic topics, so our theoretical understanding is advanced through the painstaking elaboration of a model of meaning that is constructed on the basis of systematic, controlled manipulation of crucial test sentences, judgements of acceptability, and intuitions about valid and invalid inferences. Direct negative evidence is crucial, and freely available: we know when a given utterance is infelicitous in context, or when a sentence *S* cannot assert a proposition *P*.

As an example, the antecedents of nonrestrictive relatives cannot be nonreferential quantifiers, as shown in (1a). This has been taken (for instance by Sells 1985) to indicate that the *wh*-phrase in a nonrestrictive relative is a discourse anaphor, as discourse anaphors require accessible antecedents.¹ On the other hand, *some person* in (1b) makes a perfectly good antecedent for a discourse anaphor such as a nonrestrictive relative: although classically considered to be a quantifier, *some* introduces a discourse referent which can serve as antecedent (Kamp 1981, Heim 1982). None of this is obvious, and carefully constructed contrasts like those in (1) are central to our understanding of these topics.

- (1) a. **no person*, [**who** left]
 b. *some person*, [**who** left]

Diachronic semantics, for the most part, has been different: as a discipline, it has no choice but to rely on observation of naturalistic data. The various kinds of introspective judgement available to synchronic formal semanticists are unavailable to diachronic semanticists, and negative evidence has to be inferred from absence of positive evidence. This means that the weapon of choice for classical diachronic semantics is the collocation, and diachronic semantics is typically practiced as a form of distributional semantics. If a word is characterized by the company it keeps, then changes in word meaning are characterized by changes in the company a word keeps. For instance, the grammaticalization literature (e.g., Traugott and Dasher 2002) contains several examples like (2), which demonstrate the development of *going to*, from a verb of directed motion into an expression of futurity.

- (2) a. I am going to London (to marry Bill).
 b. I am going to marry Bill.
 c. If interest rates are going to climb, we’ll have to change our plans.
 (Hopper and Traugott 2003)

¹In all examples in this paper, we enclose the relative clause in brackets, format the relative clause’s antecedent or external head (if there is one), in italics, and format the relativizer (including any complement of a *wh*-word) in boldface.

In (2b), *marry Bill* is not a place you can go to; and in (2c) *interest rates* are not the kind of things that can go. From collocational changes like these, we can infer a change in denotation: the meaning of *go* is no longer restricted to literal motion.

A consequence of this is that formal semantics and diachronic semantics often simply talk past each other. The different methods available favour different approaches to what is surely a single underlying phenomenon. Fortunately, though, the two approaches are usefully complementary. The virtues of a formal approach extend beyond precision and objectivity, the usual benefits attributed to it. Approaching semantic change through the lens of synchronic formal theories can *tell us where to look*.

Take the explanation just given for the contrast in (1): insights like this from formal semantics allow us to make precise statements about possible distributions, which in turn allow us to draw nonobvious distributional predictions. These predictions can be leveraged to provide insight into distributional changes in the historical record.

In this article, we develop an in-depth example of this kind of formal, hypothesis-led investigation of semantic change, concerning the emergence of headed relative clauses with *which* in Middle English. *Which* appears in two types of relative in Present-Day English (PDE): nonrestrictive relatives like (3a), and restrictive relatives like (3b).

- (3) a. *the University of Edinburgh*, [**which** is in Scotland]
 b. *The jewellery* [**which** he chose] was always vulgar.

There are also some cases of *which* in free relatives, such as (4a). However, bare *which* cannot appear in free relatives (see (4b)), and in most cases, both *-ever* and an NP complement are required in free *which*-relatives. Bare *what*, however, can appear in free relatives, as in (4c).

- (4) a. I ate [**whichever dish** he cooked].
 b. *I ate [**which** he cooked].
 c. I ate [**what** he cooked].

Free relatives can often be straightforwardly distinguished from headed relatives, because they do not have an external head or overt antecedent. The distributional differences between nonrestrictive and restrictive headed relatives are more subtle. There are some clear syntactic distinctions (for instance, only nonrestrictive relatives can modify clauses), but examples like (5) are structurally ambiguous between restrictive and nonrestrictive analyses. A restrictive analysis of *which I enjoyed* restricts the set of books to a subset of books which I enjoyed, while a nonrestrictive analysis adds a parenthetical remark that I enjoyed the relevant member of the set of books. Either way, (5) could be talking about the same book.

- (5) *a book*(,) [**which** I enjoyed]

In PDE, the most robust cue to the restrictive/nonrestrictive distinction is arguably prosodic: comma intonation in (5) indicates a nonrestrictive relative, and its absence indicates a restrictive relative. This correlates with a semantic (and perhaps a syntactic) distinction, but there are many cases, like (5), in which the semantic distinction is neutralized.

In Old English and Early Middle English, *which* was only used in free relatives. Headed *which*-relatives are first robustly attested in the mid-14th century. In Truswell and Gisborne (2015), we proposed that this spread of *which*-relatives followed a pathway from free relative in apposition, to nonrestrictive relative, to restrictive relative, a gradual and incremental increase in syntactic and semantic integration into the host clause. This built on a long-established literature (see Curme 1912, Johnsen 1913) demonstrating a semantic overlap between free relatives and nonrestrictive relatives, in that both constructions crucially involve definiteness.² More precisely, free relatives just are definite descriptions (Jacobson 1995), while the *wh*-phrase in a nonrestrictive relative is a (definite) discourse anaphor (Sells 1985). In contrast, the *wh*-phrase in a restrictive relative is just a λ -abstractor over a variable in the corresponding gap position. In Truswell and Gisborne (2015) we described contexts in which this semantic similarity could in principle facilitate reanalysis of free relatives as nonrestrictive relatives.

The problem with this hypothesized pathway, and the starting point for this article, is that it just doesn't work. To demonstrate this, we adapt ideas from Sells (1985) to recast the denotational differences between restrictive and nonrestrictive relatives in distributional terms. The crucial test is that the *wh*-phrase in a nonrestrictive relative is a discourse anaphor, and discourse anaphors can take certain types of referential DP as antecedents (for instance, indefinites), but not nonreferential DPs (for instance, universals).³ If we find a *which*-relative modifying a nonreferential DP, we know it's restrictive.

Using this test, we uncover a split in the behaviour of headed *which*-relatives, depending on whether determiner *which* takes an NP complement. Restrictive and nonrestrictive 'bare' *which*-relatives (with no NP complement) emerge simultaneously, as far as we can see in the textual record. As for *which*-relatives with an NP complement, like (6), they are always nonrestrictive. That is to say, they always modify referential antecedents, so there is no distributional evidence that they are restrictive, and there is enough data to make this absence statistically highly significant. In neither subcase is there a gradual progression from free to nonrestrictive to restrictive.

- (6) *the bifore knowing of God, [which bifore knowing of God bihooldith so the before knowing of God which before knowing of God beholds so without fayling thingis to comynge] without failing things to come 'the foresight of God, which beholds so infallibly things to come'*
(Late 14th century, PPCME2, cmpurvey-m3,I,55.2216)⁴

²De Vries (2002, 2006), among others, has claimed that they are also *syntactically* similar in that nonrestrictive relatives are syntactically a type of free relative. It turns out that the Middle English data actually argue against this claim, but we won't go into the details here.

³In this article we adopt the DP hypothesis, that noun phrases are DPs and NPs are complements of D, for terminological consistency with the literature that we build on. Nothing important rests on this decision.

⁴For corpus examples like (6), we give the text as it appears in the corpus, a gloss and idiomatic transcription, an approximate date, the acronym for the corpus from which the

We don't know why any of this should be the case.⁵ Our narrower aims in this article are to demonstrate that it is robust across several centuries of the history of English, and to examine the emergence of this system in Early Middle English. In doing this, we show that it is possible to give distributional historical evidence supporting precise, formally statable semantic claims.

The article is structured as follows. Section 2 gives a brief review of the diachrony of English *which*-relatives, and *wh*-relatives more broadly. Section 3 introduces the synchronic semantic analysis, and outlines the diachronic hypotheses it implies. Finally, Section 4 revisits the diachrony of *which*-relatives in the light of these hypotheses.

2. THE DIACHRONY OF *WH*-RELATIVES

The diachrony of *which*-relatives can be viewed as a special case of the diachrony of *wh*-relatives. In Old English, the major strategies for forming both headed and free relatives involved not *wh*-phrases, but instead the complementizer *þe* and the *se* series of demonstrative phrases, as in (7) (see Allen 1977, 1980 for a comprehensive description).

(7) a. ***þe*-relative:**

Gemyne he *ðæs yfeles* [*ðe* he worhte].
remember he the evil that he wrought
'Let him remember the evil that he wrought.'

b. ***se*-relative:**

Ac ge onfoð *ðæm mægene Halges Gastes* [*se* cymeð ofor
but you receive the power holy.GEN ghost.GEN DEM comes over
eow].
you
'But you shall receive the power of the Holy Ghost, who will come over you.'

c. ***se þe*-relative:**

Ic wat wytdlice *ðæt ge secað ðone Hælend* [*ðone ðe* on rode
I know truly that you seek the saviour DEM that on cross
ahangen wæs].
Hanged was
'I know truly that you seek the Saviour, who was hanged on the cross.'

(Allen 1980: 266, 269, 271)

The only *wh*-relatives were free relatives. We refer the reader to Truswell and Gisborne (2015) for a full account of Old English free *wh*-relatives; for the purposes of this article, the main points are: (1) free *wh*-relatives could occur either clause-initially or clause-finally (modulo other elements in the left and right peripheries of the

example was taken, and the ID of the sentence token. For an explanation of the latter two pieces of information, see Section 4.1.

⁵Thanks to the reviewers for critical discussion of an earlier attempt to explain the generalization.

clause); and (2) a clause initial *wh*-relative obligatorily occurred with *swa* on either side of the *wh*-phrase (as in (8a)), while this braced *swa ... swa* was optional in clause-final position (as in (8b–c)).

A reviewer asks whether the clause-initial free *wh*-relatives should instead be classed as correlatives. Indeed they are often treated as such in the descriptive literature. However, Gisborne and Truswell (2017a) give several arguments that they are instead what are sometimes called *hanging free relatives*. The simplest argument is that there are no instances of multiple *wh*-phrases in this construction at any point in the history of English, which is surprising if these are genuine correlatives but expected if they are free relatives.⁶

- (8) a. [Swa hwylc eower swa næfð nane synne on him],
 So which you.GEN.PL SO NEG.have no sin in him,
 awyrpe se ærest ænne stan on hy
 cast.out.SBJ he first one stone on her
 ‘He that is without sin among you, let him first cast a stone at her.’
 (c.1000, YCOE, coaelhom, +AHom_14:214.2117)
- b. he him aþas swor & gislas salde, þæt he him gearo wære [swa
 he them oaths swore and pledges gave that he them ready be.SBJ so
hwelce dæge swa hie hit habban wolden]
 which day so they it have want.PST.SBJ
 ‘He swore oaths and pledged to them that he would be ready whenever they
 wanted it.’ (c.900, YCOE, cochronA-1, ChronA_[Plummer]:874.5.844)
- c. Sunnandagum rædan þa gebroðra halige bec, butan [hwylcum
 Sunday.DAT read the brethren holy book except which
þenuncg betæht sy].
 service delivered be.SBJ
 On Sunday, the brethren read the holy book, except for the service that is delivered
 (that day). (c.1000, YCOE, cobenrul, BenR:48.75.1.904)

In Truswell and Gisborne (2015), we claimed that *swa ... swa* was semantically equivalent to PDE *-ever*, and adopted an analysis of free relatives with *swa hw... swa* as modal definite descriptions, based directly on the analysis for PDE developed in Jacobson (1995), Dayal (1997), von Stechow (2000). In this article, little hinges on the accuracy of that claim. The more important (and less controversial) claim is that bare free *wh*-relatives are straightforward definite descriptions.

Early Middle English saw a breakdown of the Old English free *hw*-relative system. There was a gradual erosion of the *swa ... swa* marker: the initial *swa* quickly disappeared, and the final *swa* was most often realized as *se* or *sum* (later *so*). This was later reinforced by *-ever*, giving the *what(so)ever* forms that survive today.

⁶A reviewer also noted that, on some definitions, a free relative containing an overt NP within a *wh*-phrase, as in (8b–c) should instead be classed as an internally headed relative, and that the existence of such free relatives would be problematic for theories such as that of Cecchetto and Donati (2015). This strikes us as a terminological matter: we are interested in these structures, whatever they are called.

- (9) a. *te33 inn heoffness blisse A foll3henn ure Laferrd Crist [Whatt*
 they in Heaven's bliss forever follow our lord Christ what
gate *sum he ganngēþþ];*
 way so he goes
 'They follow our lord Christ in Heaven's bliss forever, whichever way he goes.'
 (c.1200, PPCME2, cmorm-m1,I,285.2358)
- b. and [**what so euere** þu do or þenke], hit is open biforn his e3en.
 and what so ever thou do or think it is open before his eyes
 'And whatsoever you do or think, it is open before his eyes.'
 (c.1400, PPCME2, cmaelr3-m23,29.79)

At the same time, the positional conditioning of *swa ... swa* became weaker. Specifically, bare free *wh*-relatives began to be found in the left periphery, in some cases apparently with the kinds of interpretations previously associated with *swa ... swa*. For instance, (10) is a translation of the same bible passage as Old English (8a), but only the earlier translation has *swa ... swa*.⁷

- (10) [**wuch of eou echon** \ Is clene withoute sunne] ʒ þrowe þene furste ston
 which of you each.one is clean without sin throw the first stone
 'Which of you is clean without sin may cast the first stone.'
 (c.1300, PLAEME, laud108alife.473)

Concurrently, *wh*-phrases began to appear in headed relatives. Romaine (1982) showed that the first headed *wh*-relatives were confined to the bottom of the Keenan and Comrie (1977) noun phrase Accessibility Hierarchy, for instance in adjuncts and obliques rather than direct arguments. A fuller account would make reference to the fact that Early Middle English headed *wh*-relatives typically relativize PPs or adverbials rather than DPs, but these very early headed *wh*-relatives are complex to analyse, and data is scarce. For instance, there are several apparently semantically equivalent forms for a PP-gap relative, with *through what*, *through which*, and *wherethrough* all attested in different texts at roughly the same time.

- (11) a. *he sei auair welle* \ [**Of wan** al þe wat(er)es comþ an eorþe]
 he saw a.fair well of what all the waters come on Earth
 'He saw a fair well from which all the waters on Earth come.'
 (Early 14th century, PLAEME, corp145selt.70)

⁷A reviewer asks us to clarify the status of *echon* in (10), and asks specifically whether it has a similar role to *swa ... swa* or *-ever*. There is no evidence in PLAEME (from where (10) is taken) to support this conjecture. Almost all examples of *echon* occur in regular declarative clauses. For instance, (i) is from the biblical story of the feeding of the 5,000, as told in the same text as (10).

- (i) *Inou3 heo hadden ecchone* On þis o lof in þe se ʒ þat heo ne maden
 enough they had each.one in this one loaf in the sea that they NEG made
 noumore mone
 no.more moan
 'They each had enough in this one loaf in the sea that they complained no more.'
 (c.1300, PLAEME, laud108alife.35)

- b. Ne let vs no lenger(e) *þis peyne se* \ [In weche we hauen longe ybe]
 NEG let us no longer this pain see in which we have long been
 ‘Do not let us see this pain any more, in which we have been for a long time.’
 (Late 13th century, PLAEME, adde6bxvsigns.75)
- c. For *þe eareste Pilunge* [hwer of al þis uuelis] nis buten of prude.
 for the first stripping where of all this evil is NEG.is but of pride
 ‘For the first stripping, where all this evil comes from, is but of pride.’
 (Early 13th century, PPCME2, cmancriw-1-m1.II.119.1513)

Although such examples clearly form part of the story concerning the rise of headed *wh*-relatives, we put them aside in this article and concentrate on DP-gap relatives. Gisborne and Truswell (2017b) demonstrate that DP-gap headed relatives spread from *wh*-lexeme to *wh*-lexeme: *which*-relatives emerge in the mid-14th century (initially with both animate and inanimate antecedents), followed by *whom*-relatives and then *who*-relatives in the 15th century.⁸

- (12) a. he is emperor of him-zelue. Ðet is of his bodye: and of *his herte*.
 he is emperor of himself that is of his body and of his heart
 [huiche he demþ and halt ine guode payse] huerof he dep his
 which he deems and holds in good weight whereof he does his
 wyl.
 will
 ‘He is emperor of himself, that is, of his body and of his heart, which he judges and holds (to be) in good weight, whereby he does his will.’
 (1340, PPCME2, cmayenbi-m2,85.1658, 1340)
- b. But *he* [whom God hath sent], spekith the wordis of God
 but he whom God hath sent speaks the words of God
 ‘But he whom God hath sent speaks the words of God.’
 (Late 14th century, PPCME2, cmntest-m3,3,20J.234)
- c. This declaryth *the Mayster of the storyes* [who so lyste to se it].
 this declares the master of the stories who so wants to see it
 ‘The master of the stories, who so wants to see it, declares this.’
 (Late 15th century, PPCME2, cmfitzja-m4,A5R.71, 1495)

Early headed *which*-relatives do not have the same syntax as they do today. They can occur with *the*, and more importantly for this article, they can also take NP complements.

- (13) a. Þis sinful wrecche hadde *remissioun of his synnes*, [þe **whiche**
 this sinful wretch had remission of his sinnes the which
 outwardly he ne askedenouȝt, ne duely ne haddenotdeseruyd];
 outwardly he NEG asked not nor duly NEG had notdeserved
 ‘This sinful wretch was absolved of his sins, which he had not outwardly asked for nor duly deserved.’
 (c.1400, PPCME2, cmaelr3-m23,43.508)

⁸Although forms like *through what* appear in Early Middle English PP-gap relatives, bare *what* in DP-gap headed relatives is infrequent throughout the history of English.

- b. More grater uayrhede ne may by: þanne to by him ariȝt ylich.
 more greater fairhood NEG may be than to be him truly alike
 [Huych uayrhede is zuo grat: þet hit paseþ þoȝt of man / and of angle].
 which fairhood is so great that it passes thought of man and of angel
 ‘There may be no greater splendour than to be truly alike to him, which splendour
 is so great that it surpasses the thought of man and of angel.’
 (1340, PPCME2, cmayenbi-m2,100.1965)

The example in (13b) is representative of the Middle English norm, where the NP inside the relative clause is identical to an external NP within the same sentence (typically, but not always, the antecedent). From the late 15th century, it became more common for the internal and external NPs to be different. We cannot go into the syntactic implications of this here (see Truswell 2016 for discussion), but the semantic status of these NP complements will be a major focus in Section 4.

So far, we have described the reasonably well-known spread of *wh*-forms from free to headed relatives, with *which* having a special status as the first *wh*-form to spread in this way. However, the picture is incomplete in that ‘headed relative’ is a cover term for two constructions, namely restrictive and nonrestrictive relatives. These are uncontroversially semantically distinct, and since at least Jackendoff (1977) have often been taken to be syntactically distinct, too. This raises immediate questions about the diachrony of *wh*-relatives, the most basic of which is whether both restrictive and nonrestrictive relatives emerged at once or in series. Section 3 will sharpen this question, before we return to corpus data in Section 4.

3. SEMANTICS OF RELATIVE CLAUSES

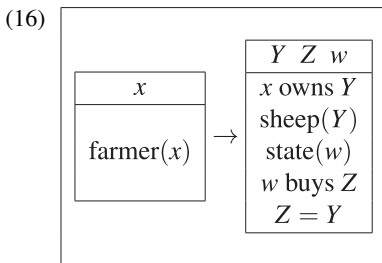
We adopt standard models of the semantics of restrictive and nonrestrictive relatives. Specifically, we assume with Heim and Kratzer (1998) and many others that a restrictive relative denotes a 1-place predicate. The restrictive relative composes with its NP sister by conjoining the predicates that each denote. The relative restricts the extension of the NP, in that $|\{x : P(x) \wedge Q(x)\}| \leq |\{x : P(x)\}|$. Compositionally, the restrictive relative is *transparent*: adding a restrictive relative does not affect the type of NP, so the constituent derived can combine with any determiner or other material that NP can normally combine with.

- (14) a. $\llbracket \text{book} \rrbracket = \lambda x. \text{book}'(x)$
 b. $\llbracket \text{which Sally wrote} \rrbracket = \lambda x. \text{write}'(s, x)$
 c. $\llbracket \text{book which Sally wrote} \rrbracket = \lambda x. \text{book}'(x) \wedge \text{write}'(s, x)$
 d. $\llbracket \text{the book which Sally wrote} \rrbracket = \iota x. \text{book}'(x) \wedge \text{write}'(s, x)$

For nonrestrictive relatives, we adopt the analysis of Sells (1985). According to Sells, a nonrestrictive relative is propositional, with the *wh*-phrase interpreted as an E-type anaphor (that is, an anaphor functioning semantically like a definite description, as in Evans 1980). Sells’ analysis is supported by the fact that *wh*-phrases in nonrestrictive relatives are maximizing, like other E-type anaphors (Evans 1980). In (15a) but not (15b), the state necessarily buys all the sheep that each farmer owns.

- (15) a. Each farmer owns *some sheep*, [**which** the state buys in the Spring].
 b. Each farmer owns *some sheep* [**that** the state buys in the Spring].

Within the framework of Discourse Representation Theory, Sells (p. 26) proposes the following representation of (15a). The essential points of this representation for our purposes are firstly that *some sheep* introduces the plural discourse referent Y , along with the information that Y is a group of sheep; secondly that *which* introduces a second discourse referent Z , and finally that the condition $Z = Y$ expresses the anaphoric relation between *which* and *some sheep*. This representation of anaphora also captures the maximizing effect of nonrestrictive *which*, as the referent Z is identical to the group of sheep which *some sheep* picks out ($Z = Y$), and cannot pick out some subset of that group.



Sells' analysis implies a first distributional test. The antecedent of a nonrestrictive relative, like any other E-type anaphor, must be referential, in a sense that includes those indefinites that introduce discourse referents (Kamp 1981, Heim 1982). There is no such requirement for restrictive relatives. The contrasts in (17) show how these facts can give a distributional diagnostic of restrictiveness.

- (17) a. *The/some/few/no sheep* [**that** the state buys] are happy.
 b. *The/some/#few/#no sheep*, [**which** the state buys], are happy.

Therefore, searching for patterns of the form $Q NP \dots RC$,⁹ where Q is a quantifier such as *few* or *no* and RC is a *which*-relative modifying $Q NP$, can inform our understanding of the diachrony of *which* relatives: following Sells, we assume that the relatives in such strings simply cannot be nonrestrictive, because *which*, as an E-type anaphor, wouldn't have the antecedent that it needs.

In Section 4, we will investigate the interactions between this straightforward test and a second distributional property, namely the presence of an overt NP complement of D. In tracking these two distributional properties, we uncover several details of the diachrony of *which*-relatives. In particular we will see that such *which NP*-relatives force a nonrestrictive interpretation.

⁹Extrapolation of relative clauses is common in Middle English, so string-adjacency cannot be relied on to determine the antecedent of a relativizer. However, the parsed corpora used in this article indicate the antecedent of extraposed relatives.

4. BACK TO WHICH-RELATIVES

This section gives a corpus-based account of the diachrony of English relatives with *which* and *what*. The major developments that we will document are the following. In Early Middle English, *which* and *what* simultaneously develop different patterns of use in two distinct dimensions. *Which* specializes for headed relatives and *what* for free relatives, and at the same time, *which* comes to allow NP complements. This Early Middle English grammar is a transitional one: after this period, *which* occurs only in headed relatives and *what* only in free relatives, but either can take an NP complement.

When headed *which*-relatives emerge, both restrictive and nonrestrictive readings are simultaneously available. However, there is an interaction with the presence of an NP complement: in the absence of an NP complement, either reading is available, but *which* NP-relatives are always nonrestrictive.

4.1 Materials

We rely exclusively on data from parsed corpora in our analysis, because parsed corpora are the only tools available for this kind of fine-grained quantitative diachronic investigation. The corpora used in this article are as follows: YCOE (York–Toronto–Helsinki Parsed Corpus of Old English Prose, Taylor et al. 2003), PPCME2 (Penn–Helsinki Parsed Corpus of Middle English, 2nd edition, Kroch and Taylor 2000), PPCEME (Penn–Helsinki Parsed Corpus of Early Modern English, Kroch et al. 2004), PCMEP (Parsed Corpus of Middle English Poetry, Zimmermann 2015), and PLAEME (Parsed Linguistic Atlas of Early Middle English, Truswell et al. 2019). Of these, YCOE, PPCME2, and PPCEME are the major parsed corpora for the relevant periods of English. However, as will become apparent, a period of particular interest in the history of *wh*-relatives is the late 13th and early 14th centuries, the ‘M2’ period in PPCME2. This is the most poorly represented period in the above corpora, in part because of the scarcity of surviving written English from this period. Accordingly, we supplement the above resources with two smaller corpora, PCMEP and PLAEME. PLAEME, in particular, is designed to fill this gap in the textual record, being composed entirely of texts from 1250–1325. PCMEP and PLAEME are composed almost entirely of verse texts, while the three major corpora are almost entirely prose. We have made no attempt to control for this in what follows, because we do not see a clear reason why metre would affect the choice between monosyllabic *that*, *which*, and *what*.

As noted in footnote 4, for each corpus example in the article we give one of the above acronyms, as well as the ID of the relevant sentence token. This information, together with the documentation for the relevant corpora, can be used to locate the examples precisely. For instance, ‘cmayenbi’ in (13b) identifies that example as coming from the *Ayenbite of Inwyt*, ‘m2’ gives the period, and ‘100.1965’ locates the example within the text.

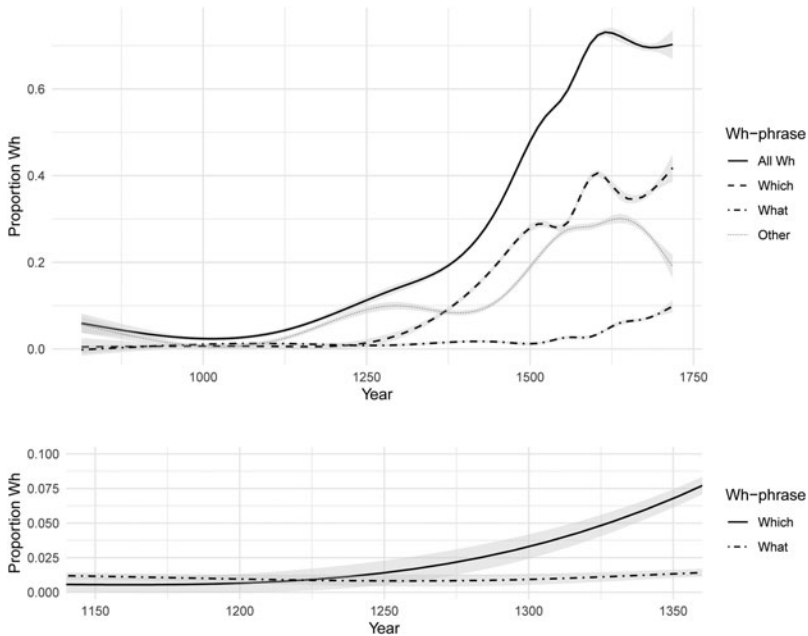


Figure 1: Frequency of *wh*-relatives over time, as a proportion of all relative clauses (top), and close-up of *which*- and *what*-relatives in Early Middle English (bottom).

4.2 Broad diachrony of *wh*-relatives

Figure 1 shows the change in global frequency of *wh*-relatives over time, as a proportion of all relative clauses (whether headed or free).¹⁰ Although *wh*-relatives are present throughout the history of English, they are very much a minority strategy in Old English: as mentioned above, they are confined to free relatives in Old English (which are much less frequent than headed relatives: only c.8% of relatives in the corpora are free relatives), and indeed they are a relatively infrequent form of free relative.

The top half of Figure 1 reveals that the spread of *wh*-relatives from this point occurs in three main bursts. A sharp increase in the frequency of ‘other’ *wh*-relatives (to c.10% of all relatives) occurs c.1150–1250, followed by an increase in *which*-relatives to c.30% of all relatives c.1250–1500, and a second increase in ‘other’ *wh*-relatives (also to c.30% of all relatives) c.1450–1550. Although our figure collapses all ‘other’ *wh*-relatives, the first of these increases is driven by use of *wh*-PPs in headed relatives, and the last by the use of *whom* and then *who* in headed NP-gap relatives.

¹⁰See the Supplementary Materials at <http://doi.org/10.1017/cnj.2020.11> for a description of our modelling and visualization choices. We believe that these choices represent an optimal trade-off between simplicity, transparency with respect to the data, and interpretability for these complex datasets, and we encourage readers to explore further visualization and modelling possibilities, using the scripts in the supplementary materials as a starting point.

The bottom half of [Figure 1](#) reveals that *which*- and *what*-relatives occurred with a frequency barely above zero throughout Early Middle English. The first point of interest in our story is the period c.1250–1350, during which the frequency of *which*-relatives began to move upwards, while that of *what*-relatives flatlined at just above zero.

4.3 *Which and what*

The increase of frequency of *which*-relatives reflects the emergence of headed *which*-relatives (recall that headed relatives are by far the more common type of relative). [Figure 2](#) shows the proportion of *which*- and *what*-relatives which are headed. The beginning of the increase in frequency of *which*-relatives in [Figure 1](#) corresponds closely to the point at which *which*-relatives become categorically associated with headed relatives, while *what*-relatives become categorically associated with free relatives.

At around the same time that *which* and *what* were specializing for headed and free relatives, respectively, a strong tendency was developing for *which* to take NP complements. This is shown in [Figure 3](#). We distinguish three broad stages in [Figure 3](#). Stages 1 and 3 are not of immediate interest: stage 1 is Old English (lasting until the mid-12th century), when no free *which*- or *what*-relative took an NP complement. Stage 3 begins in the mid-14th century and represents a stable system still largely visible in PDE. In stage 3, the primary distinction is that *which* is used almost exclusively in headed relatives, and *what* in free relatives, and choice of *which* or *what* is not directly conditioned by whether they take an NP complement.

Our interest is rather in the short-lived stage 2 (c.1150–1350), during which free *which*-relatives could take an NP-complement, and free *what*-relatives only rarely did.¹¹ In other words, examples like (18a–c) were found throughout Early Middle English, but examples like (18d) are a hallmark of later Middle English.

- (18) a. beo þe cnotte icnut eanes of wedlac. beo he cangun oðer crupel
 be the knot knitted once of wedlock be he fool or cripple
 beo he [hwuch-se he¹²eauer beo]; þu most to him halden.
 be he which.so he ever be thou must to him hold
 ‘If the knot of wedlock is knitted once, if he is a fool or a cripple, whichever he may
 be, you must remain with him.’

(Early 13th century, PPCME2, cmhali-m1,152.352)

¹¹Although stage 2 is short-lived and the distinction between *which* and *what* in that stage is not categorical, it is still clearly distinct from the better-attested grammars before and afterwards. In the Old English grammar, there are no *which* NP-relatives, and in the later grammar there are almost no free *which*-relatives. The reality of this distinct Early Middle English grammar is therefore not in doubt.

¹²This word is omitted from the version in PPCME2 but supplied on the basis of the transcription in the Linguistic Atlas of Early Middle English (Laing 2013).

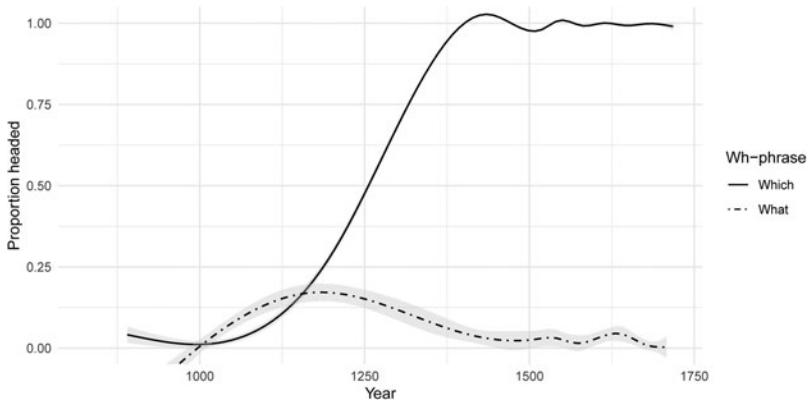


Figure 2: Proportion of *which*- and *what*-relatives which are headed, as opposed to free.

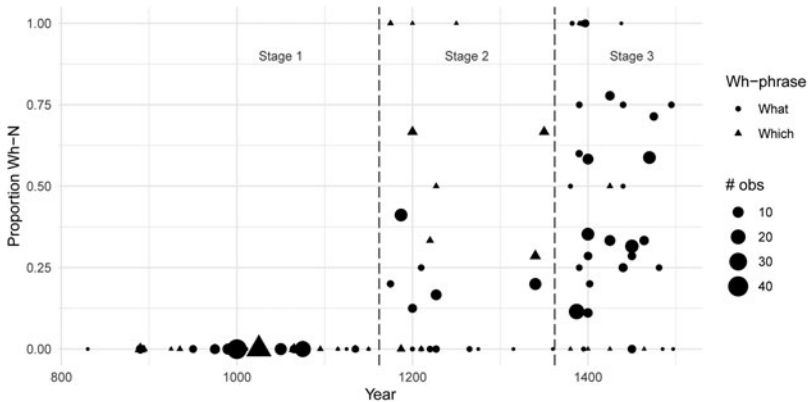


Figure 3: Proportion of free *which*- and *what*-relatives that had an NP complement.

- b. [Hwich saule. þe þer cumeþ to]. Naueþ heo neuer reste ne ro.
 which soul that there comes to NEG.have he never rest nor repose
 ‘Whoever soul that comes there will never have rest or repose.’
 (Mid-13th century, PCMEP, ElevenPains,148.52.28)
- c. leteð written on an scrouwe [hwetse ze ne cunnen].
 let write on a scroll what.so you NEG can
 ‘Let what you don’t know be written on a scroll.’
 (Early 13th century, PPCME2, cmancriw-1-m1,I.74.292)
- d. doo [what seruise þat þu canst];
 do what service that thou can
 ‘Do what service you can.’ (c.1400, PPCME2, cmaelr3-m23,40.418)

Figure 4 shows that the first headed *which*-relatives, like the last free *which*-relatives, optionally took an NP complement. Moreover, there is no evidence of a

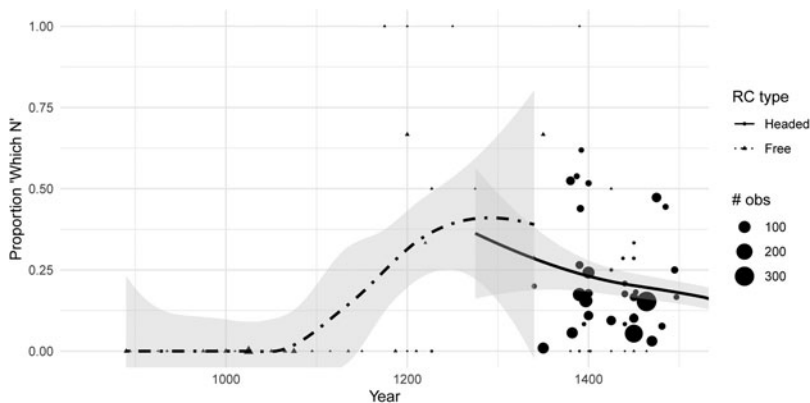


Figure 4: Proportion of free and headed *which*-relatives that had an NP complement. Loess smoothers are plotted for free relatives until 1350, and for headed relatives from 1250, because of absence of data at other times.

difference in the frequency of NP complement between headed and free *which*-relatives, although the sparsity of data c.1300 (visible in Figure 4 as very wide confidence intervals) limits our ability to interpret this absence of evidence.¹³ We take this to indicate that headed *which*-relatives emerged directly from free *which*-relatives. More specifically, we assume that clause-final free *which*-relatives are the diachronic source of headed *which*-relatives (because clause-initial free relatives are not a likely candidate for reanalysis as postnominal headed relatives — see Truswell and Gisborne 2015). We will now investigate restrictiveness of headed *which*-relatives with and without NP complements against this background.

4.4 Nonreferential antecedents

As soon as headed *which*-relatives appear in the textual record, examples with non-referential antecedents are found. This means that there is no period during which the only headed relatives were nonrestrictive.¹⁴ Figure 5 shows this in two different

¹³Igor Yanovich (p.c.) points out that Figure 4 also admits an interpretation where some texts around 1200 are generated by grammars which categorically require an NP complement, and others are generated by grammars which categorically prohibit an NP complement. Because data is limited, we cannot discriminate between this interpretation and the one given in the main text.

¹⁴Although our focus in this article is on finding robust diagnostics of restrictiveness, we note also that clearly nonrestrictive examples are also found from the start. These include bare *which*-relatives, with no NP complement. In (i), the nonrestrictive nature is guaranteed by the fact that the relative modifies a proper name, of type *e*.

(i) *Judas Machabeus*, [**which** was goddes knyght]
 ‘Judas Maccabeus, who was God’s knight’

(late 14th century, PPCME2, cmctmeli-m3,235.C1.690)

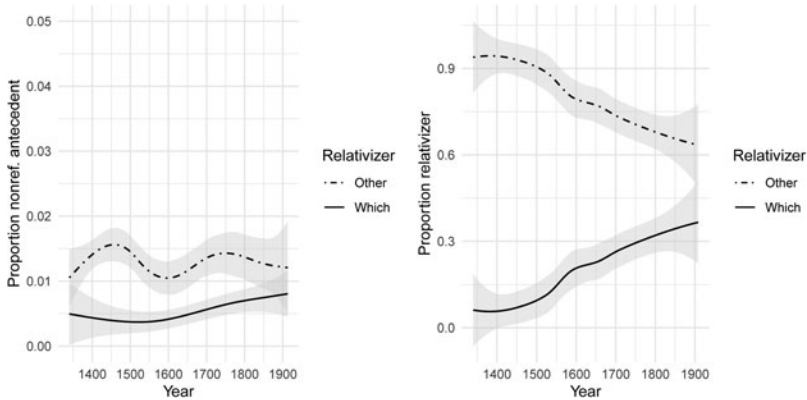


Figure 5: Proportion of *which*-relatives and other relatives that have a nonreferential antecedent.

ways: the left-hand plot shows how many *which*-relatives had a nonreferential antecedent, while the right-hand plot shows how many of the relatives modifying nonreferential antecedents were *which*-relatives.¹⁵ In each case, we see an upward trend across the period covered, but in each case, the regression line starts above zero.¹⁶

Text-by-text inspection of results confirm that even in the mid-late 14th century, every major text has a nonzero proportion of nonreferential antecedents for its headed *which*-relatives. Example (19) illustrates this for a selection of mid-late 14th-century texts, immediately after the emergence of headed *which*-relatives.

- (19) a. if we luf God in al oure hert, þar es *na thing* in us [**thurgh þe**
if we love God in all our heart there is no thing in us through the
whilk we serve to syn].
which we serve to sin
'If we love God in all our heart, there is nothing in us through which we serve to sin.'
(Mid-15th century copy of mid-14th century text, PPCME2, cmrollep-m24,110.794)
- b. for *fewe* ther ben [**the whiche** han this feruour to chastise her body so].
for few there are the which have this fervour to chastise their body so
'For there are few people who have this fervour to chastise their body in this way.'
(Mid-15th century, PPCME2, cmaelr4-m4,12.336)

¹⁵Nonreferential antecedents were operationalized as those with one of the determiners *each*, *every*, *few*, *little*, or *no*, along with orthographic variants determined manually by exhaustive search of the corpus and listed in the file WhRel.def in the Supplementary Materials, available at <http://doi.org/10.1017/cnj.2020.11>. We initially also included *any* and *all*, which are nonreferential in some of their uses, but these gave too many false positives, particularly with examples like *all the people* in the latter case.

¹⁶The upward trend is more pronounced in the right-hand graph simply because *which*-relatives increase in frequency throughout Middle and Early Modern English, as already shown in Figure 1.

- c. Y schal sle *ech fleisch* [in which is the spirit of lijf] vndir heuene
 I shall slay each flesh in which is the spirit of life under Heaven
 'I shall destroy each flesh [living thing] under Heaven in which is the spirit of life.'
 (Late 14th century, PPCME2, cmotest-m3,6,1G.224)

This falsifies the simplest form of the hypothesis that headed *which*-relatives inherit the semantic properties of free *which*-relatives: as noted in the introduction, there is a literature, beginning with Curme (1912), in which free *which*-relatives are taken to be closer to nonrestrictive relatives than to restrictive relatives, but at no point in the history of English did *which* only occur in free and nonrestrictive relatives, so the historical sequence of events has not been directly conditioned by this semantic overlap.¹⁷ However, in the following section we consider of the role of NP complements, which reveals a robustly nonrestrictive type of *which*-relative.

4.5 Headed *which*-relatives with and without NP

Without exception, no headed *which*-relatives with an overt NP complement take a nonreferential antecedent with *no*, *few*, *little*, *each*, or *every*. This absence is statistically highly unlikely to be a matter of chance. We can construct a simple estimate of the expected number of *which* NP-relatives with a nonreferential antecedent as follows: among all the corpus texts written since the *Ayenbite of Inwyt* in 1340, there are 223 examples of *which*-relatives with nonreferential antecedents. In the same texts, the frequency of NP complements of *which* in headed relatives is $1620 \div 18,318 \approx 0.09$. We therefore expect $223 \times 1620 \div 18,318 \approx 20$ *which* NP-relatives with nonreferential antecedents, as opposed to an observed value of 0. A binomial test (0 successes in 223 trials, with a hypothesized probability of success of $1620 \div 18,318$) returns $p < 10^{-8}$.

A more subtle estimate of the expected value, suggested by Igor Yanovich (p.c.), takes into account the fact that the use of *which* with nonreferential antecedents increases over this period, while the use of NP complements of *which* declines over the same period. To control for this, we repeated the same binomial test described above for each individual text (that is, if a text has n *which*-relatives with nonreferential antecedents, and if the proportion of all *which*-relatives in the text with an overt NP restrictor is p , we used binomial tests to obtain for each text the probability that 0 of the n *which*-relatives with nonreferential antecedents had an overt NP restrictor). We then took the product of these text-by-text probabilities, to obtain the probability of 0 observations across the whole dataset. According to this estimate, that probability $p = 0.002$.

For a final estimate, with somewhat different weaknesses from the previous one, we used loess smoothers with R's default parameter settings to estimate the frequency

¹⁷We cannot exclude the possibility that the late 13th and early 14th centuries were just such a period, coincidentally the period with fewest tokens of *which*-relatives. Strictly speaking, the considerations above show only that any such period was so short-lived as to be invisible in the textual record.

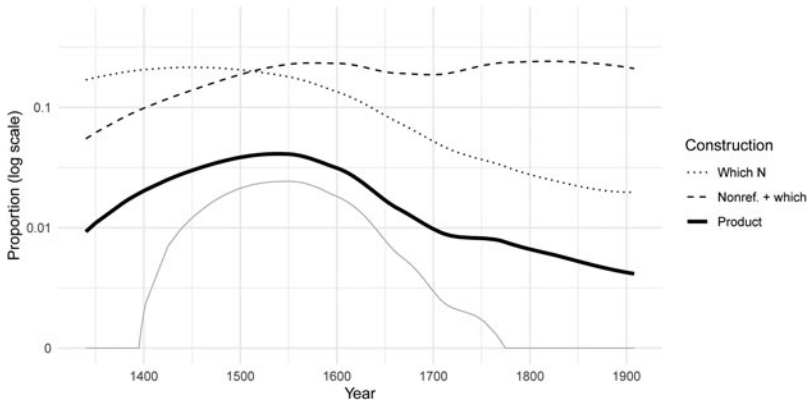


Figure 6: Expected frequency of *which* NP-relatives with nonreferential antecedents over time (thick black line), plus lower bound of 95% confidence interval (solid grey line), calculated as the product of loess smoothers tracking the frequency of *which* among all relative clauses modifying nonreferential DPs (dashed line), and the frequency of NP complements of *which* in headed relatives (dotted line). The y-axis has a logarithmic scale, except that the point marked '0' represents all values ≤ 0.001 .

of these two variables year-by-year (see the dashed and dotted lines in Figure 6), and then, for each text, estimated the expected number of *which* NP-relatives with nonreferential antecedents on the basis of these two values for the year of the text's composition (the thick black line in Figure 6). Summing these text-by-text estimates gives us a prediction of 22 such examples, almost unchanged from our first simple estimate. Although we do not have a precise *p*-value for 0 observations using this method, we used the 95% confidence interval on the product of the two loess smoothers (the solid grey line at the bottom of Figure 6) to derive a criterial value of 10 observations for $p < 0.05$. Accordingly, 0 observations is again very low probability.

We can then be reasonably sure that the absence of *which* NP-relatives with nonreferential antecedents is real, and surprising. This refines the picture from Section 4.4: although bare headed *which*-relatives are never categorically nonrestrictive, headed *which*-relatives with NP complements are always nonrestrictive, throughout their c.600-year existence.¹⁸

Given our confidence in this result, we can sharpen the notion of *referential antecedent* relevant to nonrestrictive *which* NP-relatives. In many respects, these relatives pattern just like Present-Day English nonrestrictive *which*-relatives (with no NP complement). For a start, classic donkey-anaphora configurations like (20) can be found, parallel to Sells' example (15a).

¹⁸As a reviewer notes, for those speakers of PDE whose grammar generates *which* NP-relatives, they remain categorically nonrestrictive even today. See Fabb (1990: 72) for discussion.

- (20) euery temporall man schuld paye *the xth parte of on yerly valu of hys londys and tene-mentis*, except lordis of the parlement. [**Whiche x part** amountith to the summe in euery shyre, cite and burgh as partyclerly herafter ensuete]: [...]

‘Every temporal man should pay the tenth part of the yearly value of his lands and tenements, except lords of the parliament, which tenth part amounts to the sum in every shire, city, and burgh, as particularly hereafter ensues ...’

(Late 15th century, PPCME2, cmreynes-m4,307.564)

This extends to modal and other subordination phenomena in the sense of Roberts (1987), where Sells’ (21) is structurally quite similar to Early Modern English (22).

- (21) Each boy might catch *a fish*, [**which** will struggle to get away].

(Sells 1985: 33)

- (22) everie Clothier within this Realme sholde sett *his seale of lead* unto everie Clothe and Kersey that shold be redy made and dressed to be put to sale, [**in whiche seale of lead** sholde be conteyned the true and juste content of the lenghe of everie of the same Clothes or Kerseyes]

‘Every clothier within this realm should put his seal of lead on every cloth and kersey that is ready made and dressed to be put on sale, in which seal of lead should be contained the true and just content of the length of each of the same cloths or kersseys.’

(1540–5, PPCEME, stat-1540-e1-p1,3,854.17)

Also broadly similar to Present-Day English, a quantified noun phrase does not license introduction of a plural discourse referent corresponding to the domain of quantification. That is, examples like (23) are infelicitous in Present-Day English and absent from the historical record.

- (23) #*Every book* was on the shelf, [**which** were arranged in alphabetical order].

However, unlike Present-Day English, the antecedent of a *which NP*-relative need not be a single accessible discourse referent. Examples like (24) are found, in which the antecedent of *which Townes* is the sum of *the Town of Rowcastell* and *the Town of Langton*. That is, the antecedent of a *which NP*-relative can correspond to the sum of multiple accessible discourse referents.

- (24) my broder Philipp Dacre with ccc. men which burnt and destroyed *the Town of Rowcastell* [...] and Sir Roger Fenwike with ccc. men burnt *the Town of Langton* [...] [**which Townes** er in the hert of the countre two myle beyond Jedworth opon the watere of Chevyot].

‘My brother Philip Dacre with 300 men, who burned and destroyed the town of Rowcastle ... and Sir Roger Fenwick with 300 men burned the town of Langton ... which towns are in the heart of the country two miles beyond Jedworth, on the water of Cheviot.’

(1513, PPCEME, dacre-e1-p2,1.1,94.6–8)

Corresponding configurations in Present-Day English are ungrammatical.

- (25) #*Coldstream* is in Scotland and *Cornhill* is in England, [**which** are on opposite sides of the Tweed].

We do not currently have a synchronic or diachronic account of this difference, but we suspect that the overt NP complement facilitates retrieval of this antecedent.¹⁹ A similar effect is found with Present-Day English demonstratives: *they* in (26a) is most naturally interpreted as referring to Philip, Roger, and their men, while (26b) shows that a demonstrative that explicitly mentions *towns* can refer to the presumably less topical Rowcastle and Langton.

- (26) a. Philip and his 300 men burned Rowcastle. Roger and his 300 men burned Langton. They are two miles beyond Jedworth.
 b. Philip and his 300 men burned Rowcastle. Roger and his 300 men burned Langton. These towns are two miles beyond Jedworth.

Finally, we note one example which apparently contradicts several of the above generalizations. Example (27) appears to make liberal use of coercion of the sort that is infelicitous in (25). The *wh*-phrase *the which holes* surely refers to the plurality of holes indirectly implied by the participle *holed*, even though the participle is in the scope of two universal quantifiers.

- (27) euery Spondel is holed on euery side, [through the which holes both Arteirs and veynes doo bring from the hart and the Lyuer both lyfe & nourishment]
 ‘Every vertebra is holed on every side, through the which holes both arteries and veins bring from the heart and the liver both life and nourishment.’
 (1548, PPCEME, vicary-e1-p2,74.266)

It is hard to interpret this single example. A corresponding structure (like (28)) would clearly be impossible in Present-Day English.²⁰

- (28) #Every vertebra is holed on every side, which [= the holes] are ...

This may mean that our conclusion about absence of coerced group antecedents in Middle and Early Modern English is inaccurate, but it may equally indicate that (27) is an outlier.

5. DISCUSSION AND CONCLUSION

Our main empirical result can be summarized as follows. *Which NP*-relatives are always nonrestrictive. Bare headed *which*-relatives can always be restrictive or

¹⁹In support of this, a reviewer suggests that (25) is acceptable for those Present-Day English speakers who allow *which NP*-relatives, if *which* is replaced by *which towns*. There is clearly a lot of idiolectal variation in this area, which we have not investigated in any depth.

²⁰This may be explicable in terms of the Formal Link Condition (Heim 1990, Elbourne 2001). Discourse anaphors typically require an overtly introduced discourse referent, and resist bridging of the sort apparently required in (28). See the contrast in (i), for instance.

- (i) a. Someone who has a *guitar* should bring *it*.
 b. #Some *guitarist* should bring *it*.

If *wh*-phrases in nonrestrictive relatives are discourse anaphors, as assumed throughout this article, the ungrammaticality of (28) is of a piece with the ungrammaticality of (ib).

nonrestrictive, although they come to be found more in restrictive relatives over time. Having clear, formally grounded criteria for identifying restrictive relatives has allowed us to find unambiguously restrictive *which*-relatives, even among the earliest headed *which*-relatives, and thereby falsify our (2015) claim that the early headed *which*-relatives are nonrestrictive.

There are two remaining questions, which we discuss briefly here as an invitation to further research. The first is why we don't find a period during which all headed *which*-relatives, even bare ones, are nonrestrictive. After all, the logic of reanalysis would lead us to expect such a stage. As noted above, there are clearly identifiable contexts for reanalysis of free relatives as nonrestrictive relatives, and also for reanalysis of nonrestrictive relatives as restrictive, but no context that we can see which would allow direct reanalysis of free relatives as restrictive. The natural diachronic pathway would then appear to be from free, to nonrestrictive, to restrictive relatives. We don't have an answer for this, but hope that it is related to other differences in bare and nonbare interrogative–indefinites, such as those discussed by Šimík (2018) and Belyaev and Haug (2018).

The second question is why *which NP*-relatives are so stably nonrestrictive, when so many other aspects of the grammar of English relativization are in flux. The question can be sharpened by considering the approach to E-type anaphora in Elbourne (2001). Elbourne assumes with Postal (1966) that pronouns are just intransitive determiners, so *the* differs from *she* or *it* in transitivity and inflectional marking. On Elbourne's analysis, E-type pronouns require a covert copy of the antecedent NP as a complement of the pronoun. That is, (29a) has an LF representation like (29b), where strikethrough represents elided material.

- (29) a. If a man owns a donkey, he beats it.
 b. [[always_{s₁} if a man(_{s₁}) owns a donkey(_{s₁})]_{s₂} he ~~man~~(_{s₄}) beats(_{s₂}) it ~~donkey~~(_{s₄}).
 (Elbourne 2001: 250)

Because nonrestrictive relativizers are a species of E-type anaphor, we should expect the same to hold of them. Moreover, anaphora has been analysed as an extreme form of deaccenting. We might then expect the distinction between nonrestrictive bare *which* and nonrestrictive *which NP* to reduce to the distinction between ellipsis and deaccenting.

In this way, Elbourne's analysis of E-type anaphora grows naturally into an account of why *which NP* relatives *can* be interpreted nonrestrictively. Something more needs to be said about why they *must* be interpreted like this.

We offer the following conjecture. The absence of restrictive *which NP*-relatives reflects the distribution of repetition, or redundancy, in discourse. Languages exhibit redundancy in abundance, as is well known, but at the same time, redundancy frequently leads to degradation or illformedness. An example of well-formed redundancy is verbal agreement, where the verb redundantly recapitulates information also encoded in its arguments. An example of ill-formed redundancy is (30), in which the property $\lambda x.bike'(x)$ is predicated of the individual in question twice.

- (30) #*a bike* [**which** (really) is (indeed) a bike]

In PDE, nonrestrictive relatives are more tolerant than restrictive relatives of such redundancy: although (31), out of context, is somewhat weird, it is clearly more acceptable than (30).

(31) *this bike*, [**which** (really) is (indeed) a bike]

Perhaps this reflects the fact that nonrestrictive relatives, unlike restrictive relatives, express an independent proposition. It is quite normal for a content noun to be repeated across independent sentences, for instance.

(32) Yesterday I bought a bike. This bike has twelve gears.

So our conjecture is that the problem with a restrictive relative like (33) is related to the problem with a restrictive relative like (30), in that both reflect a prohibition on certain types of redundancy within restrictive relatives.

(33) *a bike* [**which bike** has ten gears]

However, despite the stability of the association of *which NP* with nonrestrictive interpretations in English, Cinque (2011) surveys what he calls *doubly headed relatives*, with NP heads internal and external to the relative, in a range of languages. His data appears to contain both restrictive and nonrestrictive examples, although it is not clear whether he uses the same criteria we have used in his article. Cinque's survey would then appear to suggest that the association of *which NP* with nonrestrictiveness, however stable in the history of English, is still a parochial fact about English, and an explanation in the general terms just given may then not be appropriate. Alternatively, it may turn out that Cinque's operationalization of the notion of 'nonrestrictive' differs from ours. We are still some way, then, from a complete understanding of the basis of contrast between *which NP* and bare *which* in English, and its typological context.

CORPORA

PPCEME: Penn–Helsinki Parsed Corpus of Early Modern English.

PPCME2: Penn–Helsinki Parsed Corpus of Middle English, 2nd edition.

PCMEP: Parsed Corpus of Middle English Poetry.

PLAEME: Parsed Linguistic Atlas of Early Middle English.

YCOE: York–Toronto–Helsinki Parsed Corpus of Old English Prose.

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit <http://doi.org/10.1017/cnj.2020.11>.

REFERENCES

Allen, Cynthia. 1977. *Topics in diachronic English syntax*. Doctoral dissertation, University of Massachusetts, Amherst.

- Allen, Cynthia. 1980. Movement and deletion in Old English. *Linguistic Inquiry* 11(2): 261–323.
- Belyaev, Oleg, and Dag Haug. 2018. The genesis of *wh*-correlatives. Unpublished ms., Moscow State University and University of Oslo.
- Cecchetto, Carlo, and Caterina Donati. 2015. *(Re)labeling*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 2011. On double-headed relative clauses. *Revista de Estudos Linguísticos da Universidade do Porto* 6: 67–91.
- Curme, George. 1912. A history of the English relative constructions. *The Journal of English and Germanic Philology* 11(1): 10–29, 11(2): 180–204, 11(3): 355–380.
- Dayal, Veneeta. 1997. Free relatives and ever: Identity and free choice readings. In *SALT VII*, ed. Aaron Lawson, 99–116.
- Elbourne, Paul. 2001. E-type anaphora as NP-deletion. *Natural Language Semantics* 9(3): 241–288.
- Evans, Gareth. 1980. Pronouns. *Linguistic Inquiry* 11(2): 337–362.
- Fabb, Nigel. 1990. The difference between English restrictive and nonrestrictive relative clauses. *Journal of Linguistics* 26(1): 57–78.
- von Fintel, Kai. 2000. *Whatever*. In *SALT X*, ed. Brendan Jackson and Tanya Matthews, 27–39.
- Gisborne, Nikolas, and Robert Truswell. 2017a. On the stubborn refusal of English grammar to generate multiple correlatives. Paper presented at the annual meeting of the LAGB, University of Kent.
- Gisborne, Nikolas, and Robert Truswell. 2017b. Where do relative specifiers come from? In *Micro-change and macro-change in diachronic syntax*, ed. Éric Mathieu and Robert Truswell, 25–42. Oxford: Oxford University Press.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Doctoral dissertation, University of Massachusetts, Amherst.
- Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13(2): 137–177.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Hopper, Paul, and Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press, 2nd ed.
- Jackendoff, Ray. 1977. *X̄ syntax: A study of phrase structure*. Cambridge, MA: MIT Press.
- Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In *Quantification in natural languages*, ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara Partee, 451–486. Dordrecht: Kluwer.
- Johnsen, Olaf. 1913. On some uses of the indefinite relatives in Old English and the origin of the definite relatives. *Anglia* 37: 281–302.
- Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal methods in the study of language*, ed. Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, 277–322. Amsterdam: Mathematisch Centrum.
- Keenan, Edward, and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1): 63–99.
- Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2004. Penn-Helsinki Parsed Corpus of Early Modern English. <http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-1/>.
- Kroch, Anthony, and Ann Taylor. 2000. Penn-Helsinki Parsed Corpus of Middle English (2nd edition). <http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-2/>.
- Laing, Margaret. 2013. A Linguistic Atlas of Early Middle English, 1150–1325. Version 3.2, <http://www.lel.ed.ac.uk/ihd/laeme2/laeme2.html>.

- Postal, Paul. 1966. On so-called “pronouns” in English. In *Report on the seventeenth annual round table meeting on linguistics and language studies*, ed. F. Dinneen, 177–206. Washington, DC: Georgetown University Press.
- Roberts, Craige. 1987. *Modal subordination, anaphora, and distributivity*. Doctoral dissertation, University of Massachusetts, Amherst.
- Romaine, Suzanne. 1982. *Socio-historical linguistics: Its status and methodology*. Cambridge: Cambridge University Press.
- Sells, Peter. 1985. Restrictive and non-restrictive modification. CSLI report CSLI-85-28, Stanford University.
- Taylor, Ann, Anthony Warner, Susan Pintzuk, and Frank Beths. 2003. The York–Toronto–Helsinki Parsed Corpus of Old English prose (YCOE). University of York. <http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm>.
- Traugott, Elizabeth, and Richard Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.
- Truswell, Robert. 2016. Matching relatives in Middle English. Paper presented at the workshop on movement, University College London.
- Truswell, Robert, Rhona Alcorn, James Donaldson, and Joel Wallenberg. 2019. A Parsed Linguistic Atlas of Early Middle English. In *Historical dialectology in the digital age*, ed. Rhona Alcorn, Joanna Kopaczyk, Bettelou Los, and Benjamin Molineaux, 19–38. Edinburgh: Edinburgh University Press.
- Truswell, Robert, and Nikolas Gisborne. 2015. Quantificational variability and the genesis of English headed *wh*-relatives. In *Proceedings of Sinn und Bedeutung 19*, ed. Eva Csipak and Hedde Zeijlstra.
- de Vries, Mark. 2002. *The syntax of relativization*. Doctoral dissertation, Universiteit van Amsterdam.
- de Vries, Mark. 2006. The syntax of appositive relativization: On specifying coordination, false free relatives, and promotion. *Linguistic Inquiry* 37(2): 229–270.
- Šimík, Radek. 2018. Propositions and properties in the semantics of *wh*-clauses: Arguments from typology. Paper presented at the syntax–semantics colloquium, University of Potsdam.
- Zimmermann, Richard. 2015. *The Parsed Corpus of Middle English Poetry*. University of Geneva. <http://pcmep.net/>.