

'Ther been thinges thre, the whiche thynges troublen al this erthe'

The discourse-pragmatics of 'demonstrative which'

Alexander T. Bergs

This paper describes and analyses the origin, development, structure, and use of *which* as a demonstrative element. This largely qualitative study shows that demonstrative *which* had its heyday in the late Middle English and early Modern English periods. The subsequent reduction of both type and token frequency is seen in the context of both structural and sociolinguistic factors, e.g., the doubly-filled COMP filter and standardisation. From a functional point of view, its use is related to information structuring and the organisation of discourse through pragmatic markers. As a discourse marker, it refers back to a previously mentioned element and, at the same time, marks the end of one topic and provides a clear starting point for the next.

1. Introduction

Asked about the role of *which* in present-day English (PdE), most people would intuitively say that it functions as an interrogative pronoun, as in (1), and as a relative marker, as in (2):

(1) Which sauna was too hot for you?

(2) Yesterday I went to a traditional Finnish sauna, which is right off Eriksgratan.

However, we also find occurrences such as in (3):

(3) The appellate judges have several options. They can uphold Patel's decision, in which case Napster would be ordered to block all major-label music from being traded through its service.

Clearly, *which* in (3) is not interrogative. It does not even appear to be a proper relative marker, substituting a *missing* NP. It occupies the same syntactic position as the interrogative form, as pre-modifier or determiner to an overt head, but at the same time it also seems to function as some sort of relativiser in that it links the head to some previously mentioned discourse. The *Oxford English Dictionary* (OED) simply calls this an "ordinary relative adjective" (s.v. 'which'), the recent *Cambridge Grammar*

of the *English Language* (henceforth *CEL*) (Huddleston & Pullum 2002: 1043f.) refers to it as "determinative" rather than pronominal, and the *Shorter Oxford English Dictionary* comments (s.v. 'which'): "Rel. adj. 6 Introducing a clause and qualifying a noun referring to and esp. summing up the details of the antecedent in the preceding clause or sentence." The perhaps most detailed comment comes from Otto Jespersen:

Which may be used as an adjunct. This is never found in restrictive relative clauses, but only in a peculiar kind of appositive clause; very often the clause is at some distance from the antecedent, and some substantive is repeated so as to avoid any doubt as to what word is to be taken as the antecedent. This is only literary, but at best it is a rather clumsy device and is avoided by many writers.
(Jespersen 1927: §6.5.1, p. 126ff.; emphasis added)

This paper offers a mostly qualitative description and analysis of the origin, development, structure, and function of this "peculiar kind of appositive clause," the *which* plus noun construction (WPNC). In the first part, the origin and development of the construction will be outlined, followed by a detailed description of the structural background. The third part looks at the functions which the WPNC can fulfil in actual discourse. It will be suggested that the WPNC finds its main uses in information packaging (hearer-based economy and speaker-based rhetorical strategies, e.g. the presentation of information as given) and in organising discourse (as hearer guidance).

2. Origin and structural development of the WPNC

Which as a relative marker seems to have originated in the Early Middle English period (see Fischer et al. 2000: 92). As its precise origin and development are still somewhat unclear (e.g. did it develop out of generalising constructions such as *swa hwilc swa* or out of interrogative *hwilc*? – see Fischer et al. 2000: 92f.), these issues cannot and need not be discussed at this point. Suffice it to say that in Old English (OE), *hwilc* could be used both in generalising relative constructions and as an adjective or interrogative pronoun, meaning 'which, what sort of', as in (4)–(6):

(4) ond þa þider urnon swa hwilc swa þonne gearo wearp, ond radost.
and then thither ran whoever then ready was, and quickest.
"and so thither ran whoever was ready and quickest then"

(5) Gedenc hwilc witu us ða becomon for ðisse worulde
Consider, which punishments us then became for this world
"Consider which punishments we then had for this world"

(Alfred, State of Learning in *Preface to the Pastoral Care*; Whitelock, II, 25–26)

(6) *Hwylc þearf is ðe husles?*

Which need is you Eucharist?

“What need do you have for the Eucharist?”

(Bede’s Account of the Poet Caedmon; Whitelock X, 107)

What this amounts to is that in OE there were indeed the structural prerequisites for the Middle English (ME) developments of both proper relative pronouns and the “peculiar construction” of *which* plus overt head noun (WPNC). This development, in all likelihood, took place during the (early) Middle English period (the first WPNC quotation in the *OED* comes from 1175) and even perhaps under the influence of Latin.¹ One structural possibility (and rhetorical strategy) in Latin was to repeat the antecedent in the relative clause itself (see Menge 2001; Allen et al. 2002: §307):

(7) *diem dicunt, qua die ad ripam Rhodani omnes conveniant.*

a day determined, on which day on the River Rhone all should come together.
“They determined a day, on which day they should all come together on the River Rhone”

(Caesar, *De bello Gallico*: 1.6)

Alternatively, the antecedent could be taken up again through a near synonym or hyperonym:

(8) *Quod fere fit, quibus quisque in locis miles inveteraverit,*

As it usually happens, in which places troops will have spent a long time

ut multum earum regionum consuetudine moveatur.

so that often of the regions custom he is influenced

“as it usually happens that when troops have spent a long time in any district they are greatly influenced by the methods of the country”

(Caesar, *De bello civili*: 1.44)

A similar phenomenon in Latin is the ‘Relativer Satzanschluss’:

(9) *Auximum proficiscitur: quod oppidum Attius ... tenebat.*

To Auximum proceeds; which city Attius held.

“To Auximum he proceeds, which city Attius held”

(Caesar, *De bello civili*: 1.12)

Note, however, that *direct* syntactic borrowing seems to be rather unlikely in this case. Example (10) is taken from the Wyclif’s bible translation.

(10) a woman ... *which wommanis douȝtir hadde an unclene spirit, entride*

a woman which woman’s daughter had an unclean spirit, entered

“A woman, whose daughter had an unclean spirit, entered”

(Wyclif, Early Version, Mk: vii 25)

Example (10) obviously illustrates the construction which is at issue here. The exemplar in the Vulgate version, however, does not use this construction, but rather employs a simple genitive structure (11).

- (11) mulier ... filia cuius habebat spiritum immundum
 a woman ... daughter whose had spirit unclean
 "a woman, whose daughter had an unclean spirit"
 (Vulgate, Mk: vii 25)

Thus, it seems rather unlikely that the WPNC is a *direct* and simple borrowing from Latin. Nevertheless, an imitation of Latin style may have been an important factor in the rise (and demise) of this construction (also cp. Blatt 1957; Poutain 1999; Kohnen 2001).

By the late Middle and Early Modern English period, possibly under the influence of Latin, the *which* + nominal head construction was very popular and productive: "The use of *which* + Noun (that is, the referential rather than the pronominal relative) came to be almost a mannerism in the fifteenth century" (Traugott 1972: 157). In practice, *which* could combine with just about any sort of head noun, either repeating a previously mentioned antecedent (12)–(16) or summarising it (17)–(18):

- (12) But what schal I seye of delyces of body, of *whiche delices* the desirynges ben ful of angyssh
 (Chaucer, *Boethius*: Book III, Prosa 7, 1)
- (13) ther been thynges thre, the *whiche thynges* troublen al this erthe
 (Chaucer, *The Wife of Bath's Prologue*: 362)
- (14) Hir clothes weren makid of right delye thredes and subtil crafte of perdurable matere; *the wyche clothes* sche hadde wouen with hir owene handes.
 (Chaucer, *Boethius*: Book I, Prosa I, 20–23)
- (15) There is a yong man, a mercer in Chape, *the which* a Michaelmas purpose to sett up a shoppe of his owne, the which mercer makes great labor to my lady and to Jeffrey Dawne for my sister Isabell to marry with her;
 (Plumpton 8, 1464: p. 11)

- (16) Alle þe seyd lerned men telle me trewely þer is nother perill ne doute in þe takyng down of þe instrument and þe bille to no creature, *which instrument and bille* I send yow a-geyn by þe berare of this
 (Paston 1426: 03, 01)
- (17) accord is take attwyn þe seid parties that John Clopton, some and heir of þe seid William Clopton be þe grace of God, shall wedde Elizabeth, the daughter of þe seid Anneys. *For which marraige* the seid Anneys, &c., shall paye to þe seid John Clopton ccccth marc.
 (Paston 1454, Indenture of Marriage)
- (18) A prest of Norfolk þat spak with yow in Jull or August last passed told me þat he yede with yow to þe cardinales hous Tricaricensis to espie if any swych processe were sued agayn me as þe seyd lettres speciffen, and þat ye told þe same prest at alle tymes þer was þan no swich processe sued ne had; *þe which relaxation* I trust and beleue bettre þan þe seyd lettres.
 (Paston 1425: 11, 05)

Note that the antecedent may be either directly adjacent to the relative clause (12)–(13) or quite distant (14)–(15). (16) illustrates that even more complex antecedents (in this case a coordinated NP) may be taken up again and repeated verbatim. As re-

gards the summarising structures, (17) shows how some process or verbal situation is paraphrased ('shall wedde' > 'for which marage'), while (18) exemplifies that even quite complex states may be summarised or topicalised. In this case, the summarising head noun 'relation' can be interpreted almost meta-linguistically: the reported act of reporting is this 'relation', and the speaker trusts in this particular relation than in the previously mentioned letter. This kind of usage can still be found in the early Modern English period. (19) shows the repetition of the antecedent (even a proper name), while (20)–(21) show summarising or paraphrasing WPNCs:

(19) The eldest of the three wrestled with Charles the Dukes wrestler, which Charles in a moment threw him
(Shakespeare, *As You Like It*: I.2)

(20) I tell thee, Syracusan, twenty years
Have I been patron to Antipholus,

During which time he ne'er saw Syracusa:
I see thy age and dangers make thee dote. (Shakespeare, *Comedy of Errors*: V.1)

(21) And the offender's life lies in the mercy
Of the duke only, 'gainst all other voice.
In which predicament, I say, thou stand'st;
(Shakespeare, *Merchant of Venice*: IV.1)

Over time, two parallel processes may be observed, however. On the one hand, we can see a drop in overall token frequency, i.e. the WPNC in general becomes much rarer after the early Modern English period. On the other hand, we also see a sharp decline in type frequency, i.e. the number of possible head nouns is drastically limited in PDE (see Figures 1 and 2 below and Tables 1 and 2 in the Appendix).

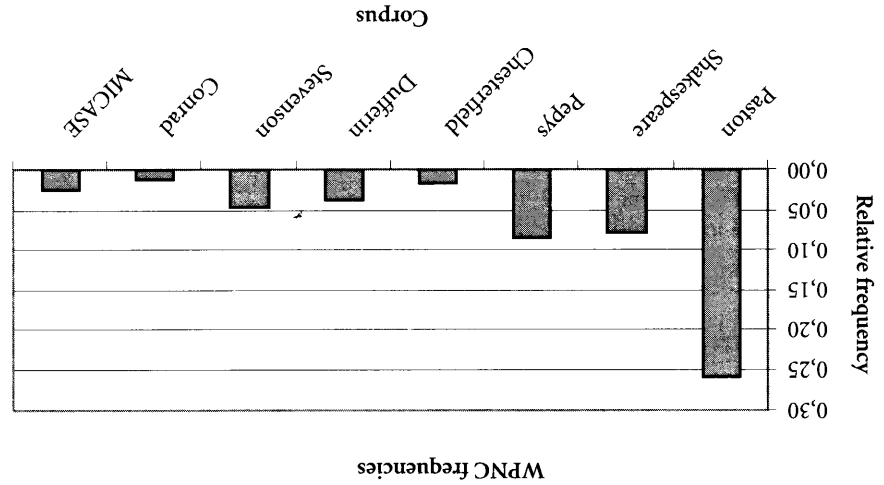


Figure 1. WPNC frequencies (see Table 1 in the data Appendix)

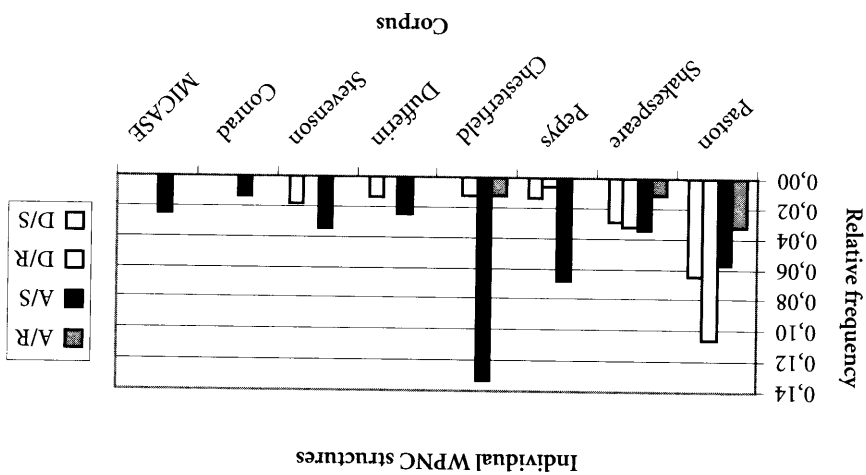


Figure 2. Individual WPNC structures (see Table 2 in the data Appendix)

The late Middle English Paston Letters (ed. by Davis 1971) with roughly 245,000 words in total contain about sixty-four occurrences of WPNCs (i.e. .261 per thousand words), including six with the head noun *time* and four with the head noun *letter*. Shakespeare has more than seventy occurrences in about 884,647 words (Stockwell & Minkova 2003), or .079 per thousand words, and more than fifty different head nouns (including *time* – ten occurrences – and *place* – three occurrences). *The Diary of Samuel Pepys* (1659–1669), with roughly 370,000 words, contains 32 (or .086 per thousand) WPNCs with more than 25 different head nouns. Surprisingly, about a hundred years later, in the letters by Lord Chesterfield to his son, written between 1748 and 1771, we can find 13 WPNCs in 83,000 words (.157 per thousand); four of these occur with *reason*, two with *time*. And we still find quite a number of occurrences of WPNCs between 1600 and 1900 (all of the following examples are quoted from Jespersen 1927: §6.5.1, p. 126ff.):

- (22) all the dangers of his solitary way, *which dangers*, tho he feared them more before, yet he saw them more clearly now
 (John Bunyan, *Pilgrim's Progress*: 1.4, 1675)
- (23) he takes any person of quality up to the said scaffold, *which person* pulls off his shoes
 (Jonathan Swift, *Works*: 1.264, 1735)
- (24) I happened upon a name, a family of *which name* lived at Hadley
 (Daniel Defoe, *Moll Flanders*: Ch. 48, 1722)
- (25) He rode twenty miles to see her picture in the house of a stranger, *which stranger* politely insisted on his acceptance of it
 (William Cowper, *Letters*: 1.212, 1912)
- (26) the application of the moral (*which said application* she was old enough to have made herself)
 (Charles Lamb, *Rosamund Gray*: Ch. 2, 1798)

- (27) Young woman with a wedding-ring and a baby *which baby* she carried about her when serving at the table (Arnold Bennett, *The Pretty Lady*: 309, 1918)
- (28) I continued about a year; *all which time* our neighbours did take me to be a very godly man (John Bunyan, *Grace Abounding to the Chief of Sinners*: I:31, 1666)

- (29) a fortnight; *during which time* the Emperor gave orders to have a bed prepared for me (Jonathan Swift, *Gulliver's Travels*: I:2, 1726)
- (30) By we I mean the bonne, the cook, the portress, and myself, *all which person-ages* were now gathered in the small chamber (Charlotte Brontë, *Villette*: Ch. 10, 1853)

- (31) He had hired a vessel to convey him to Constantinople, *for which city* he should sail in a few hours (Mary Shelley, *Frankenstein*: Ch. 14, 1818)

Evidently, (22)–(27) are examples of rather archaic repetitive WPNCs, while (28)–(31) illustrate the perhaps more modern pattern of summarising or paraphrasing head nouns.

Today, however, the construction is hardly ever acknowledged, or at best 'peculiar'; The *Letters from High Latitudes* by Lord Dufferin (c. 1856) only contain three WPNCs in c. 80,000 words (.0375 per thousand); in Robert Louis Stevenson's *Vailima Letters* (1890–1894) with roughly 85,000 words, there are merely four WPNCs (.047 per thousand words), albeit with four different nouns (*to which society, in which case, for all which mercies, to which class of exploit*). Joseph Conrad's *Notes of Life and Letters* (1921) have only one occurrence (*in which case*) in c. 80,000 words. This seems to suggest that dramatic changes must have occurred in the eighteenth and early nineteenth century. In the native speaker part of the *Michigan Corpus of Academic Spoken English* (MICASE) we find 37 (.0248 per thousand) occurrences of WPNCs, albeit exclusively in the collocations *in which case* (28), *at which point* (7), and *at which time* (2). A quick search in some of the major corpora of PdE shows similar results. For the probably most frequent WPNC in PdE, *in which case*, there were only between five and ten occurrences in each the ICE-GB, the LOB, the *Brown Corpus* (c. 1 million words each) and the LLC (c. 440,000 words). There were 895 occurrences in the BNC, a 100 million word corpus of spoken and written English. The LLC contained slightly more occurrences than the other corpora, which might be due to the fact that it is, just like the MICASE, a corpus which exclusively contains spoken English – but this question will have to be explored in future research. The CGEL confirms the general impression and argues that the repetition of the antecedent in PdE, as in (32), is actually "quite rare and formal, verging on the archaic" (1044).

- (32) They refuse to support the UN's expenses of maintaining the UN Emergency Force in the Middle East as buffer between Egypt and Israel, and the UN troops in the Congo, *which expenses* are not covered by the regular budget. (CGEL: 1044, ex. [261])

Thus, in present-day English, only *at/by/during* *which time*, *in which case*, *at which point* seem perfectly all right, albeit not too frequent in general:

- (33) Take Highway 101 to Palo Alto, and exit on University Avenue West. Take University across El Camino Real, *at which point* it becomes Palm Drive, and you will be on Stanford's campus.
- (34) The Rubidium clock should be ready for qualification by the mid 2002, *by which time* an engineering model of the Hydrogen maser should also be available.
- (35) the party challenging the decision was not present or represented when the decision was pronounced, *in which case* the time-limit shall run from the date on which the challenging party is notified of the oral decision.

It should be noted at this point, however, that, while dramatic changes in type and token frequency can be observed, and it might be speculated about concomitant register/genre differentiation, no significant changes in the function of the WPNC over time have been observed (see Section 4 below). It appears, nevertheless, that over time certain structural sub-types were more subject to change than others. The repetitive long distance WPNC, as in (23) above, appears to be almost impossible or at least archaic today, while certain summarising adjacent WPNCs (e.g. *in which case*) can still be found, albeit infrequently in present-day corpora. Structural differences in WPNCs will therefore be the topic of the following section.

3. Structure(s) of the WPNC

As has already been mentioned above, two different factors feature in the use of WPNCs: Distance of the antecedent and type of antecedent. On the one hand, the antecedent may be adjacent to the WPNC, or it may be further away (this is reflected in Lenk's distinction between local and global discourse markers, see below). On the other hand, the antecedent may either be repeated verbatim or the head noun in the WPNC may summarise or paraphrase the antecedent. These two axes are visualised in the following matrix (Table 3).

As has also been pointed out above, earlier English, particularly the late Middle English and Early Modern English periods, used the WPNC excessively, so that all four types (a-d) in Table 3 can be found. Over time, a gradual loss of possibilities seems to have occurred, so that today, while all forms are still in principle possible, example b, local/adjacent and summarising (e.g. *in which case*, *at which point* etc.), seems to be the most frequent form – see Table 2. This is of course easy to account for. In terms of iconicity it would seem that conceptual proximity feeds linguistic proximity and vice versa (cf. Haiman 1983, 1985). However, Hawkins (2001) has shown convincingly that demands placed on working memory by different numbers of syntactic and semantic relations in the processing of these relations is equally or even more important. In

Table 3. Constraint matrix of WPNCs

Antecedent	REPEATED	SUMMARIZED
------------	----------	------------

ADJACENT / LOCAL	(a)	(b)
DISTANT / GLOBAL	(c)	(d)

- (a) Were it not that I have bad dreames – *Which dreames* indeed are ambition (Shakespeare)
- (b) The inhabitants of the more temperate regions would at the same time travel southward, unless they were stopped by barriers, *in which case* they would perish. (Darwin)
- (c) he takes any person of quality up to the said scaffold, *which person* pulls off his shoes (Swift)
- (d) And the offender's life lies in the mercy
Of the duke only, 'gainst all other voice.
In which predicament, I say, thou stand'st; (Shakespeare, MOV)

the present case this means, without going into technical details, that with large(r) distances in the sentence – or even in the discourse! – both hearers and speakers would find it increasingly difficult to create and parse the anaphoric relationships, especially if the referring expression is only summarising the entities previously mentioned. Thus, while type (c) may be useful in ending a digression, establishing common ground again and continuing the conversation, type (d) increases the danger of misunderstandings and puts great cognitive strains on both speakers and hearers (note that the very high frequency of (d) in Shakespeare might be accounted for by referring to poetic, rhetoric, and aesthetic principles and functions).

Apart from these semi-functional factors, we need to take some formal-structural reasons for the developments into account, too. From a purely formal-structural point of view it has been claimed that present-day (standard) English is subject to a (generalised) doubly-filled COMP filter: "No projection may have both an overt specifier and an overt head at the end of the derivation" (Koopman & Szabolcsi 2000:40). This filter rules out structures such as *which that, for that, etc.*, and it could also cover the WPNC, which may be derived as follows:

- (a) [she [carried around [a (big) baby]]]
 - (b) [she [carried around [which (big) baby]]]
 - (c) [[which (big) baby]_i] [she [carried around [t_i]]]
 - (d) [[which \emptyset]_i] [she [carried around [t_i]]]
- Example (a) is the underlying declarative structure, example (b) represents what can be seen in partly transformed intonation questions, while (c) exemplified the WPNC as it was used in (27) above. Normally, however, the head noun or any other complementiser is deleted in this process, so that 'proper' modern structures can be generated:

This constraint developed during the Early Modern English period, which accounts for the gradual loss of this construction. Common constructions in PE such as *in which case* at first sight seem to violate the Doubly-filled COMP filter. However, it

might be argued that they constitute unanalysed ('fossilised') constructions – lexical expressions – which need not be derived but are simply inserted as a whole (cf. Wray 2001). Note, however, that this still does not explain occurrences such as (36):

- (36) William Allen White spent virtually his entire life as publisher and editor of the Emporio (Kansas) Gazette, from which *unlikely spot* he radiated an enormous influence on both journalism and politics.
 (H. L. Mencken, *Diary*, editors note, quoted in McCawley 1998: 445)
- The WPNC, from which *unlikely spot*, obviously cannot constitute an unanalysed lexical expression. How this can be unified with an active Doubly-filled COMP filter must remain a question unanswered for the time being. Suffice it to say that McCawley commented on this phenomenon thus:

Nonrestrictive clauses allow relative expressions of the form *which N* because a NP of the form [Det N] can be used anaphorically with an antecedent in a separate S... If formation of the non-restrictive clause converts as much as possible of the coreferential constituent into a corresponding Wh-word, this will yield such expressions as *which unlikely spot* in such cases. By contrast, in the structure underlying restrictive relative clauses, the relative pronoun corresponds to a pronoun that makes up the whole NP, and thus no noun is available to remain in the relative expression.
 (McCawley 1998: 453)

Coming back to constraints on the WPNC and the history of English: Just as the Doubly-filled COMP filter appeared, English also seems to have developed some functional surface rules and principles under the pressure of standardisation ideology (cf. Milroy & Milroy 1999; Stein & Tteken 1994; Cheshire & Stein 1997). One of these principles is "No double surface realization" (Stein 1997: 38). It may be argued that the WPNC is a somehow redundant double surface realization (just like *which that*, Stein 1997: 39) and that it is also reminiscent, to a certain extent and certainly counterfactually (see above), of overt orality ("Avoidance of perceived oral provenience"; Stein 1997: 39). This perceived conceptual orality is further underlined by a third factor: the WPNC is functionally related to prototypical discourse-pragmatic markers (e.g., *so, well, like*), as will be shown in the next section. The ideology of written standard languages seems to forbid the use of most, if not all, discourse markers in written language, partly because these are associated with the organisation of spoken discourse (avoidance of orality), but also because discourse-pragmatic markers can be considered "essentially interactive" (Stubbs 1983: 70, cited in Brinton 1996: 31). As interactive devices they belong to the emotive, epistemic aspect of communication and are thus (to be) avoided in educated, polite discourse ("Emotional meanings are dispreferred"; Stein 1997: 40). These are some of the factors and principles which systematically root out constructions such as *which that* or *which baby* in the ideologically motivated strive for elegant, written language – despite the fact that these structures first occurred in the model language of elegant discourse, Latin!