

The choice of relativizers in Early Modern English: evidence from the Helsinki Corpus

Emma Lezcano

UNIVERSIDADE DA CORUÑA

As is well known, relative clauses in English are the site of a clear case of syntactic variation since it is possible to choose among three different options: a *WH-word*, *that* or *zero* when a relative clause is going to be produced. A number of scholars in the recent and not so recent past have been considerably attracted by this variation in the choice of relativizers. I have in mind, in particular, the studies of Ryden (1966), Romaine (1982), Dekeyser (1984) and Rissanen (1984), who explored this aspect from a diachronic perspective or, more recently, the synchronic studies of Guy and Bayley (1995) and Fox and Thompson (1990). In fact, the variational approach has gained a prominent position in sociolinguistic studies since it gives a clearer picture of the syntactic development of a language.

The present paper will be concerned with relative pronoun choice in relative clauses in a particular period of the English language, namely, EModE. The crucial difference between previous periods of the English Language and Early Modern English is not the number of relativizers - which has not changed since the 15th century - but the system that governs its distribution, which is of great interest because in many respects it differs markedly from present English usage. The aim of this paper is, therefore, to provide a numerical account of relativizers in the the 16th century and first half of the 17th century. I will show figures of their distribution according to different parameters of variation and when possible I will explore explanations for the evidence found. This study is part of an ongoing research project in which I examine six different types of texts comprising 150, 000 words from the whole EModE period of the English language. As this is just a preliminary approach to the data, it seems prudent at present to show but some provisional results, which can certainly give us some hints of the state of relativizer choice by that time, but never definite and final conclusions.

Before taking up the above mentioned account, a few words seem in order concerning the constraints that appear to have significant effects on the choice of relativizers in PE. According to Quirk *et al* (1985) these constraints are:

- (i) The relation of the relative clause to its antecedent: restrictive and non restrictive
- (ii) The gender type of the antecedent: personal or nonpersonal
- (iii) And the function of the relativizer as subject, object, etc.

These three variables will be considered in my study. Other factors such as the adjacency of the antecedent and the relativized element, the function of the matrix NP and the complexity of the antecedent NP will also be explored. I will try to see if PE constraints on the choice of relativizers do work at all in Early Modern, the period when the consolidation of most structures has taken place.

As a source of data, a computerized corpus, *The Helsinki Corpus of English Texts*, has been used, as this may be representative of the formal and informal written language of the period. The EModE section of the corpus is divided into three subperiods. I have examined relative clauses

from two of these subperiods as illustrated in Table 1, which lists the total number of words investigated in each subperiod as well as the number of relative constructions found in them.

TABLE 1

	WORDS	RELATIVE CLAUSES	%
I. 1500-1570	12077	179	1.4 %
II. 1570-1640	10872	149	1.3 %
TOTALS	22949	328	

The total word count of period one was somewhat higher than that of period two. As can be seen in this table, there is quite an important number of occurrences, steadily distributed in the two periods. In view of these data we may presume that in Early Modern English relative constructions played a far more important role than in Present day English. In order to relate the relative clauses found in our corpus and style, a number of variables have been taken into account when selecting the material, namely, the type and register of writing, so as to obtain data representative of formal and informal settings.

Table 2 provides the breakdown of the types of texts that have been studied and the number of relative clauses in them. Private letters are informal and may reflect some of the characteristics of spoken language. On the contrary, educational treatises are formal, and bear no relationship to the spoken language.

TABLE 2

	<i>Private Letters</i>		<i>Educational Treatises</i>	
	No.	%	No.	%
Period I.	74	51,39 %	105	57,07 %
Period II.	70	48,61 %	79	42,93 %
TOTALS	144	43,9 %	184	56,1 %

As Table 2 shows, a larger percentage of relative clauses has been found in educational treatises, especially in the first period of our corpus. It would not be illegitimate to suggest that this fact may be largely due to the strong impact of Latin, especially if we take into account that the educational treatises under consideration are very formal and contain quotations and references from Latin and Greek authors. In Private letters, however, syntactic devices other than relative clauses seem to be preferred.

Table 3 shows an overall survey of relativizers in Early Modern English. This table reveals that *WH*-relative clauses are in expansion at this time: we can observe a sizeable increase in *WH*-relatives in the second period, entirely counterbalanced by a decline of the *that* strategy, much more popular in the first period, at least in view of our data. It is also during the second part of the Early Modern English period that the increase in the use of the pronoun deletion strategy takes place: I found five occurrences of zero relativizer and all five cases belong to period 2. My results coincide on the whole with those obtained by Ryden for the same period and with Dekeyser's survey of relativizers in a corpus of the 17th century.

TABLE 3

	THAT		WH-		ZERO	
	No.	%	No.	%	No.	%
Period I.	74	77,89 %	105	46,05 %	0	0,0 %
Period II.	21	22,11 %	123	53,95 %	5	100 %
TOTALS	95	28,96 %	228	69,51 %	5	1,52 %

If we break down the above mentioned figures according to the textual criterion, the results obtained will be those in Tables 4 and 5. We can see that *that* is much more frequent in private letters than in educational treatises. This is not surprising since it is well known that *that* is less common in formal texts than in colloquial ones. In the second period, though, there is a significant rise in the use of *WH* forms in private letters, as is proved by the fact that 51 out of the total 71 *wh*-relative clauses in private letters correspond to the second period. According to Rissanen (1984: 420) “the growing popularity of *wh*-forms is best accounted for through the load of functions given to *that*, which increased the risk of ambiguity”. The expansion of *who*, first recorded in the 15th century but which gained importance in the course of the 16th and 17th centuries may have also influenced these figures.

TABLE 4

	THAT		WH-		ZERO	
	No.	%	No.	%	No.	%
Private Letters I.	54	$\frac{72,97}{\%}$ 79,41 %	20	$\frac{27,03}{\%}$ 28,17 %	0	$\frac{0,0}{\%}$ 0,0 %
Private Letters II.	14	$\frac{20}{\%}$ 20,59 %	51	$\frac{72,86}{\%}$ 71,83 %	5	$\frac{7,14}{\%}$ 100 %
TOTALS	68	47,22 %	71	49,31 %	5	3,47 %

1: % EACH PERIOD — 2: % BOTH PERIODS

TABLE 5

	THAT		WH-		ZERO	
	No.	%	No.	%	No.	%
Educational Treatises I.	20	$\frac{19,05}{\%}$ 74,07 %	85	$\frac{80,95}{\%}$ 54,14 %	0	$\frac{0,0}{\%}$ 0,0 %
Educational Treatises II.	7	$\frac{8,86}{\%}$ 25,93 %	72	$\frac{91,14}{\%}$ 45,86 %	0	$\frac{0,0}{\%}$ 0,0 %
TOTALS	27	96,43 %	157	3,57 %	0	0,0 %

1: % EACH PERIOD — 2: % BOTH PERIODS

In the following tables (Tables 6 and 7) the different *WH*-relativizers are set out. This will give us a clearer picture of which *WH*-pronouns were on the decrease and which ones were favoured at the time. As may be noted, I have disregarded in my study *WH*-forms such as *where* or its compounds *wherein*, *whereof*, etc., but of course they have been included for the statistical count.

TABLE 6

	WHICH		THE WHICH		WHICHE		WHO		WHOM		WHOSE		WHOE	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Period I.	29	25, 89 %	6	100 %	29	100 %	7	24, 14 %	14	77, 78 %	6	66, 67 %	0	0, 0 %
Period II.	83	74, 11 %	0	0, 0 %	0	0, 0 %	22	75, 86 %	4	22, 22 %	3	33, 33 %	2	100 %
TOTALS	112	54, 63 %	6	2, 93 %	29	14, 15 %	29	14, 15 %	18	8, 78 %	9	4, 39 %	2	0, 98 %

TABLE 7

	WHICH		THE WHICH		WHICHE		WHO		WHOM		WHOSE		WHOE	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Private Letters I, II	54	48, 21 %	6	100 %	0	0, 0 %	9	31, 03 %	2	11, 11 %	6	66, 67 %	2	100 %
Educational Treatises I, II.	58	51, 79 %	0	0, 0 %	29	100 %	20	68, 97 %	16	88, 89 %	3	33, 33 %	0	0, 0 %
TOTALS	112	54, 63 %	6	2, 93 %	29	14, 15 %	29	14, 15 %	18	8, 78 %	9	4, 39 %	2	0, 98 %

The above mentioned quantitative evidence pertains to restrictive and non-restrictive clauses alike. Since in PE the three strategies occur side by side only in restrictive clauses, it is interesting to split up frequencies according to clause type, which is what I did in table 8. Traugott (1972: 66) observes that the contrast between restrictive and non restrictive relative clauses must have existed from the earliest times since the distinction is in essence that of fundamental semantic relationships. But there is no difference in surface structure since the Present English constraint which virtually confines *that* to restrictive clauses does not apply in this period. Punctuation is not a reliable criterion, either. In fact, I have found two examples of NR relative clauses with no punctuation mark at all and 4 restrictive clauses with a comma. I have therefore relied exclusively on semantic and pragmatic criteria. Example (1) shows a non-punctuated NR clause and example (2) illustrates the case of a Restrictive Clause with a comma:

(1) Thy loving husband who loves the more then his owne life. (<B CEPRIV2>|QE2_XX_CORP_KNYVETT: sample 1)

(2) almost in halfe the time, which the other will require. (<B CEEDUC2A>|QE2_EX_EDUC_BRINSLEY: sample 2)

In Table 8 the overall distribution of restrictive and non restrictive clauses is presented:

TABLE 8

	THAT		WH-		ZERO		TOTALS	
	No.	%	No.	%	No.	%	No.	%
RESTRICTIVE	84	88,42 %	73	32,16 %	5	100 %	162	49,54 %
NON RESTRICTIVE	11	11,58 %	154	67,84 %	0	0,0 %	165	50,46 %

Although most of the occurrences with *that* are used restrictively, it can be said that the regularizing tendency of limiting the use of *that* to restrictive clauses was not established yet: eleven cases of *that* used non-restrictively were found as example (3) illustrate:

(3) I wold desire you to mark wel my letter, that I sent you by Mr. Oughtred; (<B CEPRIV1>|QE1_XX_CORP_RPLUMPT: sample 2)

Table 9 presents the distribution of restrictive and non-restrictive relative clauses according to text-type in absolute terms.

TABLE 9

	Restrictive	Non-Restrictive
Private Letters I.	48	26
Private Letters II.	25	45
Educational Treatises I.	34	71
Educational Treatises II.	51	28

As can be seen, private letters and educational treatises of the first period serve to account for the normal tendency of using more restrictive clauses in informal types of texts and of favouring non restrictiveness in loose, descriptive and more formal writing. But on the second period the figures are reversed and surprisingly enough, there are more examples of non restrictive relative clauses in the informal kind of text than in the formal one. I cannot arrive at a tenable conclusion at this stage, but this is obviously an aspect which deserves further investigation and which will be contrasted and tackled in our larger project.

Before moving on to the next variable we should consider the use of *who* in restrictive and non-restrictive clauses. *Who* (still an innovation in the early 16th century) was first introduced in a non restrictive context. One century later, restrictive *who* was already established. The same trend can be deduced from our data. Our evidence shown in Table 10 confirms that all relative clauses with *who* in the first subperiod of the Early Modern Period were used non restrictively. However, in the second subperiod, we have 9 instances of *who* used in a restrictive environment, as example 4 in your handout illustrates:

(4) and to be imprinted in the memorie of every one who is desirous to get the best learning: (<B CEEDUC2A>|QE2_EX_EDUC_BRINSLEY: sample 2)

TABLE 10

WHO	Restrictive		Non-Restrictive		Totals
	No.	%	No.	%	
Period I.	0	0,0 %	7	100 %	7
Period II.	9	40,91 %	13	59,09 %	22

The following table, Table 11, shows the number of clauses which occur with each type of relativizer with human or non human antecedents.

TABLE 11

	WHICH		WHICHE		WHO		THAT		ZERO	
	No.	%	No.	%	No.	%	No.	%	No.	%
HUMAN	10	8, 93 %	5	17, 24 %	29	100 %	27	28, 42 %	0	0, 0 %
NON-HUMAN	102	91, 07 %	24	82, 76 %	0	0, 0 %	68	71, 58 %	5	100 %
TOTALS	112	41, 48 %	29	10, 74 %	29	10, 74 %	95	35, 19 %	5	1, 85 %

As can be seen, the animacy of the antecedent has no determining effect in the choice of relativizer since, unlike Present-day English, there are instances of *which* and *whiche* used after human heads. It is also worthy of mention that all cases with *zero* relativizer are used with a non-human antecedent, but the figures are so low that it would be risky to make strong generalisations out of these results. As regards *whose*, not present in the table, there appears not to be a constraint on *whose* in terms of Human or non Human antecedents although non human *whose* is not frequent in this period: I did not record any examples at all. It seems, though, that the present-day use of personal relativizers (such as *who*) with personal antecedents is being established and all 29 occurrences of *who* have a personal head. At this point it is interesting to note that when the antecedent is a proper name (there are 16 instances in our corpus) the pronouns favoured are *who*, *whom* or *whose*. I found, though, two examples of *which* used with a proper name as antecedent, something impossible in PE. (5) below illustrates one of those cases:

(5) as dyd the emperour Nero, whiche all a longe somers day wolde sit in the Theatre, (<B CEEDUC1A>|QE1_IS/EX_EDUC_ELYOT: sample 1)

My data also support Ryden's claim that *who* is always used when the antecedent is a word denoting the Deity, especially in closing phrases in letters as illustrated under example (6):

(6) as Jesu knowes, who preserve you in health. (<B CEPRIV1>|QE1_XX-CORP_WPLUMPT: sample 2)

Table 12 can be referred to for the breakdown of figures according to the different types of text in the two periods.

TABLE 12

	WHICH		WHICHE		WHO		THAT		ZERO	
	H ⁺	H ⁻	H ⁺	H ⁻	H ⁺	H ⁻	H ⁺	H ⁻	H ⁺	H ⁻
Private Letters I.	4	12	0	0	1	0	16	38	0	0
Private Letters II.	0	38	0	0	8	0	6	8	0	5
Educational Treatises I.	4	9	5	24	6	0	5	15	0	0
Educational Treatises II.	2	43	0	0	14	0	0	7	0	0

With regard to the syntactic function of relativizers, the general survey of data is shown in Table 13.

TABLE 13

AH	SU	OB	OBL	DET	SC
WHICH WHICHE THE WHICH	68	54	10	14	1
THAT	50	34	11	0	0
ZERO	0	3	2	0	0
WHO	29	0	0	0	0

In the same way as Romaine (1982) and Dekeyser (1984), I found quantitative evidence that the so-called Accessibility Hierarchy proposed by Keenan and Comrie (1977) worked nicely for Early Modern English. Their theory convincingly demonstrates that subjects are more accessible to relativization than DO's, DO's more than indirect objects, indirect objects more than objects of preposition, Genitives and object of comparison. As far as essentials go, the distribution shown in Table 13, agrees with Keenan's accessibility hierarchy for PE. The number of clauses relativizing subjects is higher than the number of clauses relativizing objects, prepositional complements and determiners. I have found no examples of object of comparison, at the bottom of the list, but I recorded one instance of a subject complement. Table 14 shows a more detailed distribution of functions in the different periods and text-types. That Accessibility Hierarchy works is also true for each text-type independently. Also note that relativizer *who* is only used in subject position.

TABLE 14

	WHICH WHICHE THE WHICH				THAT				WHO			
	SU	OB	OBL	DET	SU	OB	OBL	DET	SU	OB	OBL	DET
Private Letters I.	5	6	3	2	25	22	7	0	1	0	0	0
Private Letters II.	13	16	5	4	10	3	1	0	8	0	0	0
Educational Treatises I.	32	7	1	5	11	7	2	0	6	0	0	0
Educational Treatises II.	18	25	1	3	4	2	1	0	14	0	0	0
TOTALS	68	54	10	14	50	34	11	0	29	0	0	0

The final table, Table 15, presents the results obtained after analysing the function of the matrix NP in each relative clause. The distribution of figures is not very significant as to the possible influence that this factor may have on the choice of relativizers. A larger corpus will probably yield more interesting and relevant results.

TABLE 15

Matrix NP	WHO	WHICH	WHICHE	WHOM	WHOSE	THAT	ZERO	WHERE
SU	5	12	10	6	3	23	3	2
OB	9	28	9	4	1	26	1	11

OP	12	38	6	6	5	34	1	13
----	----	----	---	---	---	----	---	----

There are other factors that may affect the choice of the relative marker that I have not included in the tables. These are: the distance between the relative clause and its antecedent and the complexity of the head NP. As for the former, my data revealed that 27 of the relative clauses under consideration had some kind of intervening material between the antecedent and the relativizer as happens in example (7):

(7) an auncient and sad matrone, attendynge on hym in his chambre, whiche shall nat haue any yonge woman in her company. (<B CEEDUC1A>|QE1_IS-/EX_EDUC_ELYOT: sample 1)

The relative marker preferred in these instances is *which* and its variants *whiche* and *the which* (11 cases). The text type where this relative clause is more common is Educational Treatises 1, which may prove that the more formal the text, the more distance between the head and relativizer we find.

With regard to NP complexity, it is remarkable that there are a few cases of antecedents with extended modification in the form of relative clauses as in example 8, where the head of the second relative clause is a NP which is itself postmodified by a relative clause.

(8) I thanke you for your letter which you sent me fromTuddington: which gaue me satisfaction of your being well, (<B CEPRIV2>|QE2_XX_CORP_HARLEY: sample 2)

I found 10 cases similar to the one in the example and all of them where NR relative clauses. Again, *WH-* forms are favoured here, probably because of their greater “carrying power”.

As a matter of fact, according to Rissanen (1984: 420) the basic factor influencing the choice of relative marker at the early stages of development was the tightness of the link between the head and the postmodifying clause: the tighter the link, the less risk of ambiguity there was and consequently the less need to use the newer and more emphatic *WH-*forms. This may well serve to prove the fact that *WH-* forms are preferred when the distance between the antecedent head noun and the relativizer is considerable and that conversely *that* is far more common than *which* when the antecedent is a pronoun (57 instances of pronoun antecedents were found in our corpus and 29 had *that* as a marker, whereas only five had *which*). It seems that the relative clause is more closely linked with a pronoun antecedent, possibly because of the vague semantic content of the pronouns. Obviously, the link seems to be particularly loose when the antecedent is a whole clause. In these cases, as example (9) seems to prove, *which* is also preferred. I found 25 instances in the corpus and all of them have *which* or *whiche* as markers.

(9) to deliuer the children of Israhell out of captiuitie, which he coulde nat haue done, if he had nat bene of suche pacience and charitie. (<B CEE-DUC1A>|QE1_IS-/EX_EDUC_ELYOT: sample 2)

Taking everything into account I would suggest by way of conclusion that there are no clear constraints on the choice of relativizers in this period. Obviously, the amount of variation tolerated in this period is considerably less than a couple of centuries earlier, yet it is much important than the variation available in Present day English, even though some of the syntactic changes already foreshadow the Present English grammar of relativizers.

This being a preliminary study, many questions remain unanswered and many assumptions are probably based on inadequate evidence. It is possible that the analysis of larger samples might modify our data and show a more regular picture for the distribution. Closer textual study, together with further inspection of data both backwards and forwards in time from the period studied for this paper, would certainly give a deeper insight into the development of this syntactic construction.

REFERENCES

- Dekeyser, X. 1984: Relativisers in Early Modern English: A Dynamic Quantitative Study. Fisiak, J. ed. 1984: *Historical Syntax*. Berlin, Mouton.
- Fox, B. A. & Thompson, S. A. 1990: A Discourse Explanation of the Grammar of Relative Clauses in English Conversation. *Language* 66/ 2: 297-316.
- Guy, G. R. & Bayley, R. 1995: On the choice of relative pronouns in English. *American Speech* 70/2: 148-162.
- Keenan, E. L. & Comrie, B. 1977: Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8: 63-99.
- Kytö, M. 1991: *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and List of Source Texts*. Helsinki, Helsinki University Press.
- Moessner, L. 1992: Relative Constructions and Functional Amalgamation in Early Modern English. In Rissanen, M., Ihalainen, O., Nevalainen, T. & Taavitsainen, I. eds. 1992: *History of Englishness. New Methods and Interpretations in Historical Linguistics*. Berlin, Mouton.
- Quirk, R. et al. 1985: *A Comprehensive Grammar of the English Language*. London, Longman.
- Rissanen, M. 1984: The Choice of Relative Pronouns in 17th century American English. In Fisiak, J. ed. 1984: *Historical Syntax*. Berlin, Mouton.
- Romaine, S. 1982: *Socio-historical Linguistics. Its Status and Methodology*. Cambridge, CUP.
- Ryden, M. 1966: *Relative Constructions in Early 16th Century English*. Uppsala, Almqvist & Wiksell.
- Traugott, E. C. 1972: *A History of English Syntax: A Transformational Approach to the History of English Sentence Structure*. New York, Holt, Rinehart and Winston, Inc.

* * *

