SOCIO-HISTORICAL ASPECTS OF RELATIVIZATION IN LATE 16TH CENTURY ENGLISH: CA. 1550—1600*

Xavier Dekeyser

University of Antwerp/UFSIA and Catholic University of Leuven

Mia Ingels

Catholic University of Leuven

1. *Introduction*

1.1. In the field of socio-historical linguistics Romaine, *Language in Society* (1980) and *Folia Linguistica Historica* (1981), convincingly demonstrated that Keenan and Comrie's *Case Hierarchy* together with the distribution of the relative markers THAT, WH- (WHO and WHICH) and Contact Clauses, i.e. clauses with a zero relative, correlate with register and also have a diachronic dimension. To substantiate these claims she processed a stylistically stratified corpus of Scots English written between 1530 and 1550. At the end of the article in "Language in Society", page 96, Romaine calls her study "an example of how one might tackle a particular problem within what I have called a *socio-historical linguistic theory*, i.e. an approach which embeds function in an historical context", which is in fact the gist of her book on *Socio-historical Linguistics* (1982).

In a way, our paper is roughly in the same vein. However, it is more extensive than Romaine's data, as it is based on a sample of over 1500 Relative

---

Clauses drawn from an Early Modern English corpus (ca. 1550—1600), while at the same time it has a wider coverage as more para meters will be implemented so as to verify our hypotheses with greater certainty. Moreover, the present paper rigidly distinguishes written and written-to-be-spoken English, which has not been done before, and is also more diachronically oriented. We hope that this investigation carries the work begun by Romaine somewhat further.

1.2. The sample of relative clauses has been drawn from 20 different units of ca. 6000 running words each, nearly all of them published in the second half of the 16th century: 8 of these units belong to the register of Informative Prose or IF (essays dealing with literary theory, philosophy, etc.), 12 are dramatic texts, representing the written-to-be-spoken register. The latter set is further subdivided into 2 substrata: Tragedy or T (6 units) and Comedy or C (again 6 units). From each work the first pages, ranging from 20 to 25, were taken until the required number of running words per unit was reached.

Here follows a stratified and chronological survey of the material we have analysed:

INFORMATIVE PROSE

1551 More Th.	*Utopia* (R. Robinson's translation)
1570 Ascham R.	*Of Imitation*
1579 Lodge Th.	*Defence of Poetry*
1583 Sidney Ph.	*Apology for Poetry*
1586 Webbe W.	*A Discourse of English Poetry*
1591 Harrington Th.	*A Brief Apology for Poetry*
1605 Bacon F.	*The Advancement of Learning* (1st book)
	The Advancement of Learning (2nd book)

DRAMA: TRAGEDIES

1561 Norton Th. and Sackville Th.	*Gorboduc*
1582 Kyd Th.	*The Spanish Tragedy*
1587 Hughes Th.	*The Misfortunes of Arthur*
1590 Marlowe C.	*Tamburlaine the Great*
1594 Marlowe C.	*Dido, Queene of Carthage*
? Anon	*The Lamentable Tragedy of M. Arden of Feversham in Kent* (a late 16th century play sometimes ascribed to Shakespeare)

DRAMA: COMEDIES

1563 Bale J.	*King John*
1584 Lyly J.	*Campaspe*
1594 Greene R.	*The Honorable Historie of Frier Bacon and Frier Bongay*
1594 Shakespeare W.	*A Midsummer Night's Dream*
1595 Peele G.	*The Old Wives' Tale*
1600 Dekker Th.	*The Shomaker's Holiday*

1.3. Before we can look at relativization as a measure of syntactic complexity and at its socio-linguistic correlates, we should account for the tri-partite segmentation of the sampled material, which is intended to represent three different registers. However, the very use of the notion of register lands us in difficulties: what is register, and how can it be applied to Early Modern English texts?

Leech (1969 : 9) defines register as usage according to situation or circumstances, these being affected by three parameters: the *medium* of communication, especially speech versus writing, the *social relations* between the participants, and the *role* of the communication. On the basis of these parameters we can safely assign what we have labelled IP to the register of written scientific English, and as it addresses itself to a scholarly audience, it most probably has the characteristics of a formal register. The matter is more complicated for the dramatic samples, which we would refer to as written-to-be-spoken. However, this straightforward classification raises more questions than we can answer: to what extent do these dramatic texts mirror contemporary speech, and which level of speech is reflected? And what are the differences, if any, between "tragedy" and "comedy" in general, and between the individual plays in particular, and even between the participants within each play? Clearly, these are questions that cannot be satisfactorily answered here.

The only course, then, we can take is to start from the assumption that Comedy probably comes closest to informal spoken English and is therefore the least complex, while IP is its stylistic antipode and Tragedy ranks in between as an example of formal or elevated spoken language. This rough assessment also rests on the assumption that syntactic complexity correlates with register.

In order to verify these claims we shall attempt to measure syntactic complexity in the selected samples, taking relative clause formation as a touchstone: with a view to this we shall make use of the following variables: Keenan and Comrie's *Case Hierarchy*, the various relativization strategies, the distribution of restrictive and non-restrictive clauses, frequency of relativization, and, finally, the average length of the relative clause; the overall reliability of these parameters will be further enhanced by our data on complementizer *that* deletion, generally regarded as a good indicator of syntactic complexity.

2. *Case Hierarchy and register*

In a pioneering article which appeared in *Linguistic Inquiry* (1977) Keenan and Comrie hypothesized that the possibilities of relativization are determined by the function that NPs fill in the relative clause, subjects being more easily relativizable than objects, etc. In terms of performance this implies that the Case Hierarchy is translatable into language-specific frequencies; or as Keenan (1975: 139) puts it: "The frequency with which people relativize in discourse conforms to the CH, subjects being the most frequent, then direct objects, etc." In addition Keenan (141) argued that this frequency scale is register-sensitive: "Authors who are reliably judged to use syntactically simple sentences will present a greater proportion of RCs near the high end of the CH than authors independently judged to use syntactically complex sentences." These two predictions were verified with data from *Animal Farm*, *Sun* and *Mirror* (simple sources) and Woolf and Strawson (complex sources).

Table I below shows the frequencies with which NPs in different slots are relativized in 16th century English (Ingels 1985 : 65); the data in the second row pertain to the first half of the 17th century and have been borrowed from Dekeyser 1984a.

Table I: Frequency of relativization in different syntactic positions

	Subject	Object	Adverbials	Oblique	Genitive	Comparison	Totals
1550–1600	939 53.5%	375 21.4%	216 12.3%	113 6.5%	111 6.3%	0 —	1754
1600–1649	1251 52%	479 20%	331 14%	235 10%	119 5%	1 0.04%	2416

Oblique stands for prepositional NPs; *Adverbials* includes the temporal and locative adverbs *when* and *where* as well as compound adverbials like *wherein*, etc. In Keenan (1975 : 140) the adverbial category occupies a place between *Oblique* and *Genitive*; its position immediately after *Object* in the present material has to be ascribed to the numerous instances of compound adverbials, which correspond to prepositional NPs in Present-day English.

The data of the two samples run remarkably parallel. In Table II we have grouped the frequencies lower than *Subject*; the small value for X^2 proves the differences between the two periods to be entirely due to random sampling.

Table II: Frequency of relativization in subject and non-subject positions

	Subject	Non-subject	Totals
1550–1600	939 (53.5%)	815 (46.5%)	1754
1600–1649	1251 (51.8%)	1165 (48.2%)	2416
Totals	2190	1980 (47,5%)	4170

X^2 = ca. 1.2 (not significant)

If Keenan's second prediction is correct, we should arrive at different CHs depending on the registers involved. Let us first look at the data in general (Table III):

Table III: Case Hierarchy and register (Ingels 1985: 69 and 74)

	Subject	Object	Adverbials	Oblique	Genitive	Totals
IP	544 (56.2%)	190 (19.6%)	119 (12.3%)	85 (8.8%)	29 (3%)	967
T	234 (47.2%)	120 (24.2%)	66 (13.3%)	17 (3.4%)	59 (11.9%)	496
C	139 (56.7%)	55 (22.4%)	17 (6.9%)	11 (4.5%)	23 (9.4%)	245

Table IV presents the same data in a dichotomic configuration:

Table IV:

	Subject	Non-subject	Totals
IP	544 (56.2%)	423 (43.8%)	967
T	234 (47.2%)	262 (52.8%)	496
C	139 (56.7%)	106 (43.3%)	245

The difference between T and C in Table IV is significant at the .01 level X^2 = ca. 6), while the frequencies for IP and C are almost identical. The Case Hierarchy ranks C as the simpler segment, T as the more complex one, which comes up to our expectations. This order is also confirmed by the frequencies of relativization in the subject position for each of the 6 sample units: C ranges from 72% down to 44%, T from 51% to 42% (Ingels 1985 : 103). Interestingly, there is only one C sample (Lyly) with a frequency of 44%, which comes close to the more complex T units with 43% and 44%; all the other Cs are simpler tnan any T unit.

These data prove the Case Hierarchy to be a sensitive measure of complexity, at least as far as Tragedy and Comedy are concerned. But what about the Informative Prose segment? One would expect it to be at least as complex as Tragedy, which is manifestly not the case in our material. On closer inspection it appears that the data in Table III are biassed by the surprisingly high values for *Genitive* in both drama samples, which are not in agreement with Keenan's Modern English data and with his second prediction; neither do they agree with the data collected by Roggero (1967), who recorded no instances at all of genitival relativization in a 100 page sample from two modern plays. However, there is a similar, though smaller, deviation in the 17th century material (Dekeyser 1984a), which shows the comparatively high frequency for genitives to be a stylistic feature inherent to the written-to-be-spoken register in Early Modern English: in the Comedy sample Genitive ranks between Object and Adverbial, in Tragedy between Adverbial and Oblique. If we discard

Table V: Incidence of WH/TH correlated with syntactic position (Ingels 1985: 81 - 82)

Va: COMEDY — Increasing complexity →

Functions	%	THAT	%	%	WH	%	Totals
Subject	(85.7)	96	(70.6)	(41.7)	40	(29.4)	136
Object	(13.4)	15	(37.5)	(26)	25	(62.5)	40
Oblique	(0.9)	1	(11.1)	(8.3)	8	(88.9)	8
Genitive	—	—	—	(24)	23	(100)	23
TOTALS		112	(53.9)		96	(46.1)	208

Vb: TRAGEDY — Increasing complexity →

Functions	%	THAT	%	%	WH	%	Totals
Subject	(83.1)	167	(73.6)	(31.6)	60	(26.4)	227
Object	(16.9)	34	(38.6)	(28.4)	54	(61.4)	88
Oblique	—	—	—	(8.9)	17	(100)	17
Genitive	—	—	—	(31.1)	59	(100)	59
TOTALS		201	(51.5)		190	(48.6)	391

Vc: Inf. Pr.: — Increasing complexity →

Functions	%	THAT	%	%	WH	%	Totals
Subject	(83.3)	210	(39.1)	(58.2)	327	(60.9)	537
Object	(13.9)	35	(20.8)	(23.7)	133	(79.2)	168
Oblique	(2.8)	7	(8.75)	(13)	73	(91.25)	80
Genitive	—	—	—	(5.2)	29	(100)	29
TOTALS		252	(31%)		562	(69)	814

Notes:

(1) Adverbials have been left out in the above Tables as we are interested in the distribution of THAT as compared with WHO(M)/WHOSE and WHICH.

(2) Percentages at the right of the absolute values pertain to the right-hand totals and have to be read horizontally, as indicated by the small arrows: they indicate the frequency of WH/TH in a particular slot. Percentages at the left are related to the totals at the bottom and consequently have to be read vertically: they show the frequency of either TH or WH in the various slots.

Genitive and confine ourselves to the other "complex" slots, Informative Prose and Tragedy emerge as more or less equally complex, with resp. 40.7% and 40.8%, while Comedy now ranks as the least complex register, with a frequency of 33.8%, which, as we shall demonstrate later, is roughly in agreement with the evidence provided by the other parameters that we are going to use.

3. *Register, Case Hierarchy and relativization strategies*

3.1. WH and THAT

It is known that WH is a relativization strategy that entered the English language in the 11th or 12th century; there is also evidence from various sources that it first appeared in the more difficult slots of the Case Hierarchy and then worked its way down the scale; see Romaine (1980 and 1981) and Dekeyser (1983 and 1984b). This aspect of the history of relative clause formation, to which we shall return at the end of our paper, is almost ideally documented in our 16th century data as represented in Tables Va, Vb and Vc on page 30.

We find WH to be particularly frequent in the complex functions, where diachronically it originated, while the province of TH is largely confined to the easy subject slot, as indicated by the vertical percentages. Following Romaine (1982 : 151) we hypothesize that TH is a "less abstract or simpler" strategy than WH, the spread of which in the course of ME is often put down to influence from the prestigious Romance acrolects, first French and then also Latin.

Yet there is a striking discrepancy between Drama and Informative Prose, as evidenced in the horizontal relative frequencies: in the former the ascendance of WH over TH ranges from Object down to Genitive, while in the latter it begins from the subject slot, which proves WH to be more widely spread here. This is also convincingly borne out by the overall distribution of WH/TH in each sample, as demonstrated in Table VI, which is derived from V and contrasts written and written-to-be-spoken English.

Table VI: Relativization correlated with register

	WH	TH	Totals
Informative Prose	562 (69%)	252 (31%)	814
Comedy and Tragedy	286 (47.7%)	313 (52.3%)	599
Totals	848 (60%)	565 (40%)	1413

$X^2 =$ ca. 65

WH is significantly more frequent in Informative Prose than in the two drama segments together, while the reverse holds for TH (a value for X^2 as high as ca. 65 is utterly significant). With $X^2 = 0.30$ for Tragedy and Comedy in Tables

Va and Vb above, we have to accept that the incidence of the two relativiza-tion strategies is very much the same in the written-to-be-spoken strata, which runs counter to what we expected to be the case.

3.2. Contact clauses or Ø

Not only do WH and TH correlate with register, a correlation also exists be-tween Ø and register (Romaine: 1980 and 1981, Dekeyser: 1984b). In Table VII the slight difference between T and C is statistically not at all significant; X^2 for the proportion between drama and IP is ca.10.8, which shows Ø to be con-siderably less frequent in the latter stylistic stratum.

Table VII: Contact clauses and register

	WH/TH	Ø	Totals
IP	814 (96.5%)	30 (3.5%)	844
T	391 (92.7%)	31 (7.3%)	422
C	208 (92.5%)	17 (7.5%)	225
Totals	1413 (94.8%)	78 (5.2%)	1491

4. Register and +R/—R clauses

Given the fact that non-restrictive clauses only provide additional information about the antecedent NP, and that less complex language, such as spontaneous speech, tends to present this in a paratactic or non-embedded structure, we expect —R to be a mark of the more complex registers. Our data prove this to be the case for IP and Drama, while the difference between C and T is once again a marginal one.

Table VIII: Register correlated with +R/—R (Ingels 1985: 95)

	+R	—R	Totals
IP	468 (55.5%)	374 (44.5%)	842
T	258 (62.2%)	156 (37.8%)	414
C	144 (63.2%)	84 (36.8%)	228
Totals	870 (58.6%)	614 (41.4%)	1484

$X^2=7.3$ for Drama and IP; C and T do not significantly differ.

5. Frequency of relativization and length of the relative clause

It appears from the foregoing that the Case Hierarchy should be handled with caution when 16th century texts are involved. The other variables imple-mented so far, viz. WH/TH, Ø and +R/—R clearly differentiate what we have

referred to as Informative Prose and Drama as regards syntactic complexity. They also seem to suggest that T is slightly more complex than C. We now turn to two other parameters which will lend further support to these hypotheses: frequency of relativization and length of the relative clause.

5.1. Register and frequency of relativization.

It has been suggested that parataxis is a feature of "primitive" languages and that it "is the simpler construction of the two (our note: *parataxis* and *hypotaxis*) because of its presence in the early stages of many languages" (Romaine 1984: 446). If this assumption is correct, one would expect more relative clauses in IP than in Drama, and presumably also more in T than in C, which is exactly what our 16th century data reveal. We have calculated an *index of relativization* by dividing the total number of relative clauses of a given substratum by the corresponding total of words, divided by 1000; the bracket-ed frequencies for IP include sentential relatives:

Sample	Number of Rel. Cl.	Numer of running words	Number of Rel. Cl. per 1000 words
IP	967(1013)	48,000 (8 units × 6,000)	20(21)
T	496	36,000 (6 units × 6,000)	14
C	245	36,000 (6 units × 6,000)	7

We can infer from this that T contains twice as many relative clauses as C, and IP (inclusive of sentential relatives) even three times, which allows us to set up the following tripartite stylistic scale:

SIMPLE   C ———————— T ———————— IP   COMPLEX
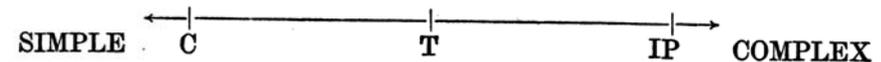
Interestingly, deletion of complementizer **THAT** correlates with syntactic complexity in exactly the same way as the index of relativization does, but then in the inverse direction; deletion increases as we move from IP over T to the colloquial C register (Ingels 1985: 91):

Informative Prose: 7.5%
Tragedy: 25.5%
Comedy: 39%

5.2.
A last measure of syntactic complexity that we want to apply here is clausal length. Quirk (1968: 107—108) has demonstrated that in the London corpus of educated spoken English WHICH, THAT and Ø correlate with length of the relative clause. Given the fact that in our sample relativization

strategies are to some extent determined by register, we may expect length to be significant here as well. With a view to this we made a random selection of 5 clauses per unit, which yields a total of 100 items:

	Number of words	Number of clauses	Average length
Informative Prose	491	40	12.3
Tragedy	371	30	12.4
Comedy	213	30	7.1

Once again this count shows Comedy to be less complex than both Tragedy and Informative Prose, which match each other in terms of clausal length.

6. *An index of syntactic complexity and register*

Both frequency of relativization and complementizer deletion allow us to set up a tripartite scale of increasing complexity as shown above. Three other parameters, WH/TH, Ø and +R/—R unmistakably rank IP and Drama respectively as more and less complex. Given the non-significance of the differences between T and C, we should be wary of jumping to conclusions here. However, the slight leaning of T towards the more complex register, which these parameters seem to suggest and which is also confirmed by the length index, is remarkably corroborated by the index of relativization and complementizer deletion in Table IX. Case Hierarchy proves C and T to be two stylistic extremes, but contrary to expectation ranks IP with the former rather than with the latter.

In the table below we have assigned points reflecting the ranking of the registers for each of the parameters involved; as stated before, there is only a bipartite division for the first four variables and clausal length.

Table IX: Indices of syntactic complexity

	CH	WH/TH	Ø	+R/—R	Freq.Rel	Cl. length	(COMP-del.)
SIMPLE	C/IP (1)	C/T(1)	C/T(1)	C/T(1)	C(1)	C(1)	C(1)
INTERMEDIATE					T(2)		T(2)
COMPLEX	T(2)	IP(2)	IP(2)	IP(2)	IP(3)	T/IP(2)	IP(3)

Exclusive of COMP-deletion, this Table yields the following overall values: Comedy: 6, Tragedy: 9 and Informative Prose: 12.

If we divide these figures by the total number of variables, viz. 6, we obtain three equidistant indices of syntactic complexity; the bracketed indices are obtained if we include COMP-deletion:

SIMPLE	Comedy	1 (1)
↓	Tragedy	1.5 (1.6)
COMPLEX	Inf. Prose	2 (2.1)

As argued by Romaine (1981: 87), such indices constitute more accurate measures of stylistic levels than any other parameter on its own. Comedy, which we assumed to represent colloquial spoken English, and Informative Prose are each other's antipodes on all scores; the solemn and more or less elevated written-to-be-spoken style of Tragedy typically ranks in between. Given the stratification provided by these composite indices and the discrepancy between T and C as regards frequency of relativization and complementizer deletion, the non-significant differences for WH/TH, Ø, +R/—R and clausal length are probably also indicative of stylistic differentiation. Case Hierarchy, apparently a sensitive parameter for Comedy and Tragedy, assigns a place to Informative Prose which as yet we cannot accommodate to the overall index of syntatic complexity; stylistically more refined analyses will hopefully remove this "anomaly".

7. *The diachronic dimension*

Clearly, the analysis of relative clauses as pursued here has a socio-linguistic dimension; but what are the historical implications, if any? Some of the parameters that we have used, notably WH versus TH and the Contact Clause, also have a diachronic dimension.

7.1. WH and TH

In the relevant literature we can find that WH derives either from Old English indefinites with an implied antecedent or from interrogatives in certain types of indirect questions, or perhaps from both; see e. g. Mustanoja (1960: 191ff.) or Traugott (1972: 153—154). Whichever their origin is, there is evidence that WH originated in prepositional slots, because, unlike þe and þat, these new relatives did not require preposition stranding; the earliest Middle English examples can probably be found in the *Peterborough Chronicle*, for which we refer to Dekeyser (1983). Linguists are also agreed that the spread of these new relatives was to some extent promoted by the corresponding models from the French and Latin acrolects.

On the basis of this general evidence we would make three claims as to the spread and distribution of WH in the course of Middle English, which will be documented by the extensive data we have at our disposal and which are presented in Table X below:

Claim 1: the distribution of WH is conditioned by the Case Hierarchy, there being more occurrences in the difficult slots than in the easy ones.

Claim 2: the distribution of WH is determined by the period in which a parti-

cular text was written: a Late Middle English corpus of material contains more instances than an Early Middle English one.

Claim 3: finally, the distribution of WH is determined by register.

In other words, the occurrence of WH in Middle English is governed by three parameters: Case Hierarchy, chronology and register. Let us now look at the diachronic data of Table X:

Table X: The Middle English disribution of WH

Homilies 12th and 13th centuries (based on data from Van den Eynden 1984: 167)

	SUBJECT	OBJECT	OBLIQUE	GENITIVE	Totals
þe	58(52.3%)	26(51%)	2(25%)	—	86(50.6%)
þat	52(46.8%)	25(49%)	3(37.5%)	—	80(47%)
WH	1(0.9%)	—	3(37.5%)	—	4(2.4%)
Totals	111(100%)	51(100%)	8(100%)	—	170(100%)

Hidgins' Polychronicon translated by Trevisa (1387) (based on Martens 1986: 16)

	SUBJECT	OBJECT	OBLIQUE	GENITIVE	Totals
þat	252(99.6%)	29(96.7%)	10(35.7%)	—	291(93.6%)
WH	1(0.4%)	1(3.3%)	18(64.3%)	—	20(6.4)
Totals	253(100%)	30(100%)	28(100%)	—	311(100%)

Middle English Sermons (MES) written between 1378 and 1417 (Van den Eynden 1984: 169)

	SUBJECT	OBJECT	OBLIQUE	GENITIVE	Totals
þat	95(91.3%)	42(93.3%)	3(14.3%)	—	140(80.5%)
WH	9(8.7)%	3(6.7%)	18(85.7%)	4(100%)	34(19.5%)
Totals	104(100%)	45(100%)	21(100%)	4(100%)	174(100%)

Hidgins's Polychronicon in a mid-15th century translation (Martens 1986: 31)

	SUBJECT	OBJECT	OBLIQUE	GENITIVE	Totals
that	23(12.8%)	6(11.1%)	—	—	29(9.8%)
WH	157(87.2%)	48(88.9%)	57(100%)	6(100%)	268(90.2%)
Totals	180(100%)	54(100%)	57(100%)	6(100%)	297(100%)

Note: The percentages have to be read vertically and indicate the distribution of WH versus THAT (or þe) in a particular slot; horizontal percentages have not been calculated as we are not interested in Case Hierarchies per pronoun this time.

Most strikingly, this Table reveals the increasing incidence of WH in the course of Middle English, together with its spread from difficult to easy slots. The 12th and 13th century *Homilies* clearly constitute an example of Early Middle English, with Old English þe still frequently used and WH only in an incipient stage, while the material from Trevisa (1387) and MES (1378—1417)

show a moderate use of WH mainly in the difficult slots, though MES already has a fair number of WH-subjects. Impressionistically we would assume the 15th century translation of the *Polychronicon* to be more sophisticated than Trevisa's, so the unusually high incidence of WH here has to be put down both to chronology and register. It should be added that the results arrived at by Caldwell (1974) for Early Scots (ca. 1375—ca. 1500) point in the same direction and interestingly corroborate our findings. As evidenced in Tables V and VI, the ascendance of WH over TH in the latter half of the 16th century is significantly more advanced in Informative Prose than in the written-to-be-spoken registers of Tragedy and, above all, Comedy, which our data prove to be closer to an older and simpler relativization strategy.

7.2. Contact Clauses

"The colloquial nature of the asyndetic relative clause cannot... be disputed", Phillipps (1965: 324) writes, which accounts for the seemingly erratic distribution of Contact Clauses in most of the Middle English texts: some manuscripts, like Trevisa or the *Homilies* contain no or hardly any examples at all, while in others, such as the *Peterborough Chronicle* continuations or *Piers Plowman* the construction is well attested; see Dekeyser (1986). This diachronic dimension of Contact Clauses is also well reflected in our 16th century material, where Ø is significantly more frequent in Drama than in the Informative Prose segment; the same holds for our 17th century drama data (Dekeyser 1984a): here Contact Clauses account for 11% of all relative clauses, while Informative Prose only reaches 2.5%.

8. *Conclusion*

At the end of this paper it is tempting to contrast the Early Modern English data with those of the corpus of educated spoken English, given in Table XI below. Typically, the frequencies arrived at in Quirk's survey rank in between "our" Drama and Informative Prose, at least as far as WH and TH are concerned. Now, if we assume, as we have done before, that the dramatic sample more or less approximates the spoken language in the Early Modern English period, we can claim that the development of present-day educated spoken English is characterized by a considerable extension of WH coupled with an equally remarkable shrinking of TH. Romaine (1980: 235), however, points out that this only holds for educated (?English) English; in Modern Scots and American English THAT and Contact Clauses are still preferred to WHICH: "they are, Romaine remarks, closer to a more colloquial or vernacular norm".

Table XI: Relativization strategies in Early Modern and Modern English

	WH	TH	Ø	Totals
Informative Prose	562 (66.5%)	252 (30%)	30 (3.5%)	844 (100%)
Drama (T/C)	286 (44.2%)	313 (48.4%)	48 (7.4%)	647 (100%)
Quirk's Survey	695 (53.5%)	373 (29%)	228 (17.5%)	1296 (100%)
Totals	1543	938	306	2787

$X^2 =$ ca. 190, which for 2 df is a highly significant value.

Another characteristic of present-day educated spoken English is that Contact Clauses are definitely on the increase in Modern spoken English, even if we allow for the fact that the incidence of ø in "drama" is artificially low because a written-to-be-spoken register is involved. Given the highly significant value for X^2 in the Table above, we can state with great certainty that present-day (educated) spoken English reveals two opposite developments: on the one hand it continues and even extends the tradition of earlier colloquial English in that it makes frequent use of Contact Clauses, on the other hand its widespread adoption of WH is obviously rooted in the more complex registers of written English. As amply evidenced in this paper and elsewhere, these developments are the outcome of changes that started centuries ago.

SOURCES

Bowers, F. (ed.). 1973. *The complete works of Christopher Marlowe.* Vol. I. Cambridge University Press.

Churton Collins, J. 1904. *Sir Thomas More's Utopia, edited, with introduction and notes, by J. Churton Collins.* Oxford at the Clarendon Press.

Comedies by William Shakespeare. Library of Classics. London and Glasgow: Collin's Clear-Type Press.

Cunliffe, J. W. (ed.). 1912. *Early English classical tragedies.* Oxford at the Clarendon Press.

Mc Ilwraith, A. K. (ed.). 1961. *Five Elizabethan tragedies.* London: Oxford University Press.

Mc Ilwraith, A. K. (ed.). 1962. *Five Elizabethan comedies.* London: Oxford University Press.

Mc Ilwraith, A. K. (ed.). 1965. *Elizabethan history plays.* London: Oxford University Press.

Smith, G. G. (ed.). 1959. *Elizabethan critical essays.* Volumes I and II. London: Oxford University Press.

Wright, W. A. 1868. *Bacon's Advancement of learning, edited by William Aldis Wright.* Oxford at the Clarendon Press.

REFERENCES

Caldwell, S. J. G. 1974. *The relative pronoun in Early Scots.* Mémoires de la Société Néophilologique de Helsinki, XLII. Helsinki: Société Néophilologique.

Cambridge Bibliography of English Literature (The). 1974. Vol. I: 600—1600.

Daems, F. and Goossens, L. (eds.). 1983. *Een spyeghel voor G. Jo. Steenbergen.* Leuven: Acco.

Dekeyser, X. 1983. "Relative markers in the Peterborough Chronicle: 1070—1154". In Daems, F. and Goossens, L. (eds.). 1983. 95—107.

Dekeyser, X. 1984a. *Socio-historical linguistics and relativization in 17th century English: ca. 1600—1649.* Unpublished paper.

Dekeyser, X. 1984b. "Relativizers in Early Modern English: a dynamic quantitative approach". In Fisiak, J. (ed.). 1984. 61—87.

Dekeyser, X. 1986. "English contact clauses revisited: a diachronic approach". *Folia Linguistica Historica 7/1*: 105—118.

van den Eynden, N. M. 1984. *The process of relativization in Middle English: a diachronic approach based on the analysis of 12th—late 14th century theological prose.* Unpublished dissertation, K. U. Leuven.

Fasold, R. W. and Shuy, R. W. (eds.). 1975. *Analysing variation in language.* Washington, D. C.: Georgetown University Press.

Fisiak, J. (ed.). 1984. *Historical syntax.* Trends in Linguistics, Studies and Monographs 23. Berlin—New York—Amsterdam: Mouton Publishers.

Ingels, M. B. 1985. *Socio-historical aspects of relativization in 16th century English (ca. 1550—1600).* Unpublished dissertation, K. U. Leuven.

Keenan, E. L. 1975. "Variation in universal grammar". In Fasold, R. W. and Shuy, R. W. (eds.). 1975. 136—148.

Keenan, E. L. and Comrie, B. 1977. "Noun phrase accessibility and universal grammar". *Linguistic Inquiry 8*: 63—99.

Leech, G. N. 1969. *A linguistic guide to English poetry.* English Language Series. London: Longman.

Martens, L. 1986. *The process of relativization and personal pronouns in Middle English: a diachronic approach based on the analysis of two translations from Hidgin's "Polychronicon" (ca. 1380—1440).* Unpublished dissertation, K. U. Leuven.

Muller, Ch. 1968. *Initiation à la statistique linguistique.* Paris: Librairie Larousse.

Mustanoja, T. F. 1960. *A Middle English syntax.* Part I: Parts of speech. Mémoires de la Société Néophilologique de Helsinki, XXIII. Helsinki: Société Néophilologique.

Phillipps, K. C. 1965. "Asyndetic relative clauses in Late Middle English". *English Studies 46*: 323—329.

Quirk, R. 1968. "Relative clauses in educated spoken English". In *Essays on the English language, Medieval and Modern.* Longmans' Linguistics Library. London: Longmans. 94—108.

Roggero, J. 1967. "*Whose* et *of which*: recherche en vue d'une application à l'enseignement". *Linguistique 61*: 405—416.

Romaine, S. 1980. "The relative clause marker in Scots English: diffusion, complexity and style as dimensions of syntactic change". *Language in Society 9*: 221—247.

Romaine, S. 1981. "Syntactic complexity, relativization and stylistic levels in Middle Scots", *Folia Linguistica Historica 2*: 71—97.

Romaine, S. 1982. *Socio-historical linguistics. Its status and methodology.* London: Cambridge University Press.

Romaine, S. 1984. "Towards a typology of relative clause formation strategies in Germanic". In Fisiak, J. (ed.). 1984. 437—470.

Traugott, E. C. 1972. *A history of English syntax: a transformational approach to the history of English sentence structure.* Transatlantic Series in Linguistics. New York: Holt, Rinehart and Winston, Inc.