DIACHRONIC CONSIDERATIONS ON ENGLISH RELATIVE CLAUSES: SEMANTIC AND STRUCTURAL EVOLUTION

Paloma Tejada Caller
Universidad Complutense

"Historical linguistics may be characterized as the art of making the best of bad data".
"Linguists are a marvellously clever bunch of scholars: there is really no limit to the imaginative, elegant and intellectually satisfying hypotheses they can dream up to account for observed linguistic behaviour".

The subject of this paper —relative clauses (RC)— constitutes an example of the diachronic approach I think most clarifying. In it, apart from the analyses offered by historical linguistics and the data taken from the texts themselves, synchronic linguistics, linguistic typology and data from different linguistic varieties are taken into account. With them all, therefore, I will try to weave a net of relationships, so as to explain the behaviour and evolution of RC's through the history of English, if such a concept -RC- can be said to exist from the very beginning. That is, I will establish a general frame for the subject, considering the syntax of the clause as part of a wider textual system, comparing different local and social varieties of language and assuming that Old English (OE) and contemporary British English may —and perhaps should— be considered two different languages: one semantic-oriented; the other, syntactic-oriented, in recent linguistic terms [1].

1. To study the evolution of RC's we should first of all try to define what a RC is. However, what we usually understand for RC is restricted to what this concept stands for in contemporary British English, and this turns clearly insufficient for the historical study of language. So, assuming that we are dealing with different languages in which we may find considerably different phenomena, the definition should be as close as possible to the universal idea of RC. The concept, as it has been defined up to now [2], is unclear, more and more so as we widen our field of study.

The initial problem to be faced is therefore the lack of a definition which covers all those structures we come across as we move from Old to Modern English

(and necessarily in the analysis, from standard to non-standard varieties). We set out accordingly from the problems and inconsistencies found in the traditional definition of the concept, in order to suggest a new one more suitable to our goals.

1.1. The first point in the controversy refers to whether RC's are *subordinate*, i.e. dependent, structures or not. Most authors consider them to be dependent structures –the term "clause" itself confirms this [2], but many of them add certain adjectives to that definition or feel the need to modify it slightly, and so RC's become "special subordinates", "endocentric dependent clauses" (Winter 1982), or "structures with different degrees of subordination" (Jacobson 1977).

We will not go further into the subject or give any solution to it. It seems quite clear that RC's are not independent constructions, but just a look at example (1) will give us an idea of how complex the subject may be. In this sentence the reference of the pronoun *he* varies, and therefore, the structure is close to what we consider to be an independent clause.

(1) I told John, who told his brother, and he told his wife

The consideration of RC's as such often involves two further questions, the one referring to the presence or absence of a finite-verb or of a verbal form at all within the structure; the other, trying to specify which the elements that introduce true RC's are, and by which they may be defined. For the sake of brevity we will have to restrict our pages to constructions including finite verb forms, disregarding therefore structures such as (2) through to (8).

(2) The language used by McDonald proved unintelligible
(3) The passengers leaving at 10.30 found their flight had been cancelled
(4) The great evils to be fought are those of low salaries and unemployment
(5) The painter good at landscapes can easily earn his living here
(6) The linguist from New York left without a word
(7) I was wakened by a barking dog
(8) The newly-arrived population includes five million African people

As for the second of the above-mentioned subjects, i.e. which elements can appear in the introductory slot for a clause to be defined as "relative", there is a wide disagreement as to whether adverbial RC's, the so-called "independent RC's", cleft and pseudo-cleft structures ought to be considered RC's or not. We have examples of them all in (9) through (12):

(9) The house where you lived is being repainted
(10) What you said is wrong
(11) It was John who did it
(12) Who did it was John

The controversy is great. Examples (9) and (13) might be clear enough; not so much (14) and (15), often called conjunctive clauses:

(13) That was the time when she went to Amsterdam
(14) This is where I was born
(15) That was when she went

At the other end we have (16), somewhere in between pure adverbial and RC's:

(16) He felt at home the moment I left the room

Moreover, clauses such as those in (17) and (18) constitute an even further step in the controversy:

(17) That's what work I've done
(18) I've rung up what friends I have

Reflections multiply, even if we disregard the never-ending discussion on the nature of clefts and pseudo-clefts. We will leave the question temporarily unconcluded and move forward.

1.2. A second big focus of controversy in the definition refers to the classification of RC's according to the concept of *restrictiveness*. Once more the problem spreads in all directions. Is the difference between restrictive and non-restrictive clauses a formal difference? Is it semantic? Is it both? Are there any tendencies for the occurrence of either restrictive or non-restrictive clauses in different languages or linguistic varieties? Is the concept of restrictiveness a gradient or are we dealing with a twofold division, with a dichotomy? Does the classification of RC's depend on the semantic peculiarities and organisation of the lexicon of a given language?

The existence of two different types of clauses —even of two different types of RC's— is almost generally accepted, whatever the not always successful terms coined to name them. That there exists a semantic and functional difference between restrictive and non-restrictive clauses no one can deny. However, the syntactic difference so clearly observed in contemporary standard British English is not universal by any means. Even worse, according to Comrie (1981), the pragmatic and semantic difference is conveyed by a formal distinction only occasionally across languages. Therefore, contemporary British English turns out to be a rare species in this respect, the contrast being expressed in surface structure. In most languages the difference between restrictive and non-restrictive clauses must be looked for in the context or corresponding sequence. The difference is not sentential, but textual and this must be taken into account for our purposes.

On the other hand, anyone can prove how non-restrictive clauses tend to be avoided in the spoken language, in more pragmatic linguistic varieties and, we may add, in semantic-oriented languages. In all these cases coordinated or juxtaposed structures are preferred, as can be observed in examples (19) to (21). Sentences under b/ represent what the usual spoken discourse would be:

(19) a) My niece, whom you met yesterday, lives in London
     b) You remember my niece yesterday? She lives in London
(20) a) My house, which has just been redecorated, looks nice
     b) My house has just been redecorated and looks nice
(21) a) I've broken it, which is a nuisance
     b) What a nuisance! I've broken it

Up to now RC's were classified into restrictive and non-restrictive, in a bipolar division of what constituted a clear dichotomy. However, depending on the language or variety chosen, with changing frequency we come across examples really difficult to allocate. To illustrate this, some examples in Spanish ([22]-(24)] have been selected (Ojea López 1987). They appear to be more suitable to our purposes and often closer to OE than the ones we could take from the current language, as shall be seen further on.

(22) Continué con los muchachos, con los que se sentían con fuerzas
(23) Los muchachos, los que se sentían con fuerzas, continuaron
(24) Es un Madrid que no conozco

The idea of restrictivity being a gradient and not a bipolar idea, then, begins to emerge from cases like the ones quoted above, which proves to be most fruitful for the understanding of the OE system and the further evolution of the structures considered. For our later definition and analysis of RC's it should also be remarked that this idea of restrictivity being a gradient has much to do with the lexico-semantic network of a given language. Lexical and referential univocity varies in space and time and, we could even say, from text to text, if the speaker wants it so.

1.3. A third subject in our analysis of RC's refers to the so-called *"strategies of relativization"*. Among these the most common and usually the only one reckoned for Present-Day English (PdE) is that of pronominalization. In the current language there are said to be three series of pronouns: *wh-*; *that*; and *zero*, each of them constrained by clear and well-known rules which prevent their free distribution. We should perhaps mention that the acceptability of the zero-pronoun in subject position grows not only in oral registers, but also in certain structures such as cleft and existential sentences [see (25) and (26)], and it is also widely accepted that it is the semantic or perceptual ambiguity conveyed by sentences such as (27) and (28) that

often justifies the insertion of the corresponding relative particle (Traugott 1972, Bolinger 1968).

> (25) There's a man at the door wants to see you
> (26) It was my father bought it
> (27) I saw the man voted
> (28) The fellow over there just voted is an FBI agent

In recent years, however, there has been a strong reconsideration of the subject and the theory that *that* is not a pronoun at all, but a subordinating conjunction or an invariable complementizer is gaining ground (Stahlke 1976, Grimshaw 1985, Dekeyser 1984). *That* is only a particle to signal the dependent nature of the following clause. The formal correspondence with the antecedent is therefore non-existing; there is no indication about the function or morphological features of the head of the clause, which must be accounted for through the verbal and extraverbal context. Under this interpretation PdE would have two relativization strategies, coming close with it to the usual behaviour of many other languages in the world.

Much less clearly heard is the opinion of those who claim a third relativization strategy in contemporary English (Stahlke 1976, among others), drawing their attention mainly towards non-standard varieties. This third strategy is the retention of an anaphorical personal or demonstrative pronoun within the RC, a mechanism considered unimportant or left aside by many and illustrated in (29). However, despite not being central to the system, this third strategy must be carefully taken into consideration, if we want to throw some light on earlier stages of the English language.

> (29) The girl that she was crying left

The parameter we are dealing with now —relativization strategies— may be said to be the one that has attracted most the attention of comparative linguists. To the strategies mentioned above, we should add one more found in languages not so close to our own, and which could be called "noun repetition". This strategy —illustrated in (30)— constitutes the other end in the line of syntactization:

> (30) The girl the girl was crying left

Most interesting for our purposes, however, is not the fact that there exist several relativization strategies in different languages, but their potential co-existence and, above all, distribution. The rules constraining this distribution will be much clearer if we introduce the concept of Accessibility Hierarchy (AH).

This term and its corresponding concept was introduced by Keenan and Comrie in 1972 [3] and has proved really useful and illuminating for our goals. The

basic idea is that not everything can be relativized, i.e. that there are functions for which this transformation is not possible. And this must be understood not in an absolute sense but relatively. Different languages will be able to relativize different number of functions. So the AH —reproduced under (31)— must be interpreted as follows: the possibility of relativizing a given function in a language A involves the possibility of doing it with those functions placed to its left. Contemporary English covers the whole range of possibilities —as is illustrated in examples (32) to (37)—, but this is not so in all languages. There are languages that can only use this structure up to the indirect object, for example. The head would then presumably occur as subject, direct object and indirect object within the RC, but not as prepositional object.

(31) *AH: subject/DO/ IO/ PrepO/ Poss/ Obj. of comparative particle*
(32) The man who was coming today...
(33) The man you mentioned...
(34) The man to whom you gave the book...
(35) The man about whom you spoke...
(36) The man whose car is red...
(37) The man than whom no one is taller...

This AH must, therefore, be considered an index of syntactic complexity, since it clearly involves that it becomes increasingly difficult to make up RC's as we move to the right in the hierarchy; but it also constitutes an index of semantic and pragmatic complexity, since the hierarchy itself stems from an increasing difficulty of psychological perception. From this it must be concluded that this increasing complexity affects not only the number of occurrences registered for different functions; it must also determine the syntactic resources used to relativize a given funtion. The more "difficult" the function, the more redundant and explicit the syntactic structure will be. Therefore, far from being arbitrary, the distribution of the different relativization strategies in a given language is mainly functional. Even in PdE, for example, a more explicit technique than the ususal ones must be used for the least accessible positions, as in the extreme and perhaps doubtful case of (38), where a pronoun is included anaphorically. However, the fact that this strategy is only marginal in standard written English does not necessarily imply that this is so in all languages and varieties. Pronoun-retention may even be the only technique available, or a central one, if —as it usually happens— the hierarchy for that language or variety is shorter.

(38) This is the road that I know where it leads

A deeper study into this essential concept would prevent us from coming to the end, or even the core, of this paper. To sum up, we must keep in mind that languages may enjoy different relativization strategies which include a larger or smaller

number of pragmatic indicators. The distribution of these formal resources depends on the language or linguistic variety considered and on the function performed by the head within the RC. The codification of this function reflects an increasing degree of perceptual complexity, as will be proved when we deal with dialectal and early English.

2. At this stage and through the information outlined up to now, a suitable frame for the analysis of RC's can be set up. As we have seen, it is difficult to give a formal definition of RC's. The best point of departure might be, therefore, to consider them within the subject of textual reference.

It is well known that the existence of correferential mechanisms constitute a pragmatic option, ranging from lexical repetition or renominalization to the absence of any formal sign in surface structure. That is, the choice of a particular mechanism depends on the need to identify or re-identify a NP, together with the need of referential redundancy. In other words, the option of a structure being more or less explicit is constrained by co-textual and contextual circumstances, such as thematic continuity, easiness with which information is recovered, communicative importance of an element, etc.

On identifying RC's with correferential structures we are putting forth the link between RC's and other independent constructions. At the same time we are remarking on the existing relationship between the above-mentioned scale of correferential explicitness and the different ways there exist to codify RC's. This relationship may be analysed interlinguistically —there are languages with a more explicit codification than others— and intralinguistically; in a given language or variety there may be different ways of codifying RC's (according to the scale of perceptual complexity referred to when talking about AH). The more pragmatic a language is, the more obvious this scale of codification will be.

Taking this into account, we must argue at an intermediate stage that relativization is but textual correference codified through one or more syntactic mechanisms. The distribution of these strategies —if there are more than one— depends, at least partly, on pragmatic and communicative factors.

At the same time, we must not forget that formal subordination is not a universal phenomenon and that there are languages where it is not syntactic inter-sentential relationships that are signalled, but referential continuity. This means, therefore, that it may not always be possible to distinguish between subordination and coordination, i.e. RC's from mere juxtaposed correference. The current syntactic categorization of PdE darkens the existing semantic and functional categories. However, in OE, as in Greek and other ancient languages, the process had not begun yet; semantic and functional categories were still preserved.

Coming back to the idea of textual correference, we have just mentioned that in the relationship holding between RC's and other correferential formulae, relatives are not only substituting structures, or structures found in alternating distribution

with those formulae; they are at times formally identical elements diachronically related to them. In our analysis attention should be drawn to the relationship observed between relative structures and other textual and communicative constructions, namely theme/rheme (T/R) structure, definite articles and cleft periods.

There have been many to notice the relationship between RC's and topic or sentential theme (Dahl 1974, Comrie 1981, Sankoff & Brown 1976, Danes 1976). To mention just some examples, cross-linguistically the unmarked position of topic and the only function universally relativizable coincide; in some languages the noun to be relativized can only be topic of the sentence; very often the markers themselves for both topic and RC are the same; RC's have been said to be a strategy to express T/R in cases where there is a high degree of fusion or cohesion, etc.

With regard to the relationship obtaining between relatives and definite articles, it is most concluding that in many typologically distant languages there is a formal coincidence between articles and relativizers: German, Greek, OE, Tok Pisin, etc. From a functional point of view, the article is used to signal an element as [+ given]; i.e. it performs an identifying, backward-looking function, so that the listener is able to identify, or re— identify, the reference as unique. RC's bracket a given material with this very identifying function within discourse. The idea of "definiteness" is, therefore, essential to relativization. Relative structures are not so often modifiers as a special kind of textual determiners.

Winter (1982:57) offers an interesting example. In newspaper reporting, after a headline *Informer lied to the Court* there come others meant to act as a guide to the reader, such as *Informer who lied*. This means that the reader identifies a given NP not by name, but by the only activity he knows about him. A RC -in Winter's words- "picks up a preceding or already known clause and uses its lexical uniqueness to the next mention of the clause participant."

And it is this identifying function itself that explains the semantic and functional relationship between RC's and cleft structures and the formal coincidence of the elements used in either group. In a cleft structure the identifying function is clearly maintained; to that a contrastive component is added; a deviation from what is expected, be it about the participants or about the actions involved:

(39) It was the policeman who lied

To sum up, we could define RC's as *textual formulae of codified correference with a central identifying function*. And within this frame we can understand Comrie's assumption that restrictive RC's are more central to the notion of RC than non-restrictive ones. The former are essential to identify the referent, whereas the latter constitute a means to introduce new information on an already well-known —identified— referent.

3. The next axis of our discussion is focused on relative structures found in English linguistic varieties different from the current British written standard. The interest here is not dialectal as such, obviously, but diachronic and follows some not so recent trends of historical study for which the interrelationship between one and the other subject is essential [4]. First of all, the standard language is not necessarily the variety that most clearly reflects the evolution of a certain structure. So, the study of other linguistic varieties helps to reconstruct the process followed with a higher degree of reliability; and secondly, with the study of other linguistic varieties we widen the typological linguistic range, which in itself may contribute to clarify certain phenomena.

We will point out what get relativized in some less well-known varieties of English and how it is achieved. However, the aim here will not be an exhaustive analysis, but only the establishing of a referential frame of possibilities existing today, similar to the ones that were central in earlier stages of the language. It must be born in mind that in dialectal studies morpho-syntactic information is constituted by disperse, anecdotic and far from systematic data.

One of the most common phenomena found in informal British and American English is the overgeneralized use of the particle *that* as relativizer, and the consequent almost non-existence of the remaining pronouns. *That* is used for anything, even to express the function of possession, as we have in (40) and (41). (The American, where *leg* is not the subject of the clause, constitutes a step further down in the scale).

> (40) The man that's leg is broken...(British English)
> (41) The man that's leg the wolf broke...(American English)

This overgeneralized use of *that*, whatever the function and nature of the antecedent, is also found in proverbs and fixed expressions, as in:

> (42) Handsome is that handsome does
> (43) He that fights and runs away may live to fight another day

In all these cases *that* acts as an invariable particle, as a mere indicator of nexus or dependency between elements. This fact leads easily to constructions often called "freer"; in other words, to constructions including the retention of an anaphoric pronoun, such as those in:

> (44) The man that John talked to him was a liar
> (45) The man that I found his book came in last night
> (46) A man that I'm younger than him got the appointment

In standard British English some of these structures would require a more complex syntax; some others as (47) could not be codified as proper RC's:

(47) Some guy that John knows the girl who married him has been elected to
the Board.

This relativization strategy -the retention of an anaphoric pronoun- is used in
all non-standard English varieties as well as in the language spoken in Scotland and
Ireland; it is also found in pidgins (Tok Pisin) and children speech (48). According to
Lass (1987), sentences such as (49) are typical all around the UK, with the exception
of some southern counties.

(48) There's two fellows that their dads are millionaires
(49) The man that his brother drowned left the place soon

Much more important than their geographical limits, however, is the fact that
these structures are frequent mainly in genitive; when it is the possessor, a
pragmatically and syntactically complex position, that is relativized. As will be seen
later on, most formal inconsistencies and variation are found in these less accessible
positions. This proves, therefore, that these last positions in the scale are the ones
that require a higher number of explicit signs to smoothe over the processing of the
message.

To illustrate the above-mentioned relationship between deictics and
relativizers, two cases have been selected. Firstly, that of Tok Pisin, where *ia* (here)
performs both functions (50). Here, the relative/ deictic particle marks the limits of
the RC almost bracket-like. The occurrence of the anaphoric pronoun *em* is not
compulsory, but constrained by pragmatic factors, according to Sankoff & Brown
(1976):

(50) Meri ia em i yangpela meri draipela meri ia em harim stop
(The girl who was young and big was listening) [*em*: 3rd p.sg.pron]

Secondly, in varieties closer to us, the use of "relative pronouns" as
determiners of a given NP is not uncommon. It often happens in non-restrictive
clauses, and the structure is so much more popular as the message gets more difficult
to process. See examples (51) and (52), close to juxtaposition itself.

(51) We arrived at noon, by which time the demonstration was over
(52) It was necessary that she should go and start a new life with that old
friend of hers, for which purpose....

Another relevant point in the discussion is the use of zero-relative in varieties
different from the standard. It should be included in the scale of formal explicitness
referred to above, which is dependent ultimately on pragmatic and communicative
criteria.

There is no doubt about the acceptability in informal language of structures

with zero-relative in subject function [(53)-(55)]. As we have already pointed out, these constructions were also admitted in more standard registers when included in existential or cleft sentences.

(53) Any man says that is a liar
(54) The man did it was a friend of mine
(55) Anybody does that should be locked up

Authors like Bolinger (1968) and Traugott (1972) maintain that zero—relative is used in English whenever there is no risk of ambiguity, again a communicative reason. However, it is the subject that originates a higher number of confusing situations (56), and this could explain the compulsory use of an explicit pronoun in the standard variety. Under the same parameters, we must understand the rule that prevents the use of zero-relative in non-restrictive clauses, since these structures usually demand a more explicit marking.

(56) I saw the man was coming [a) I saw the man who...; b) I saw that...]

At the other end of the continuum, we find the case of Black English, where the tendency to use zero-relative is so widely spread that a different formal sign —the so-called subject agreement— has proved necessary to mark main clauses:

(57) The man came to dinner was really hungry (RC)
(58) The man he came to dinner and left (MC)

A last feature to be taken into account in this analysis of English varieties is the obvious instability of the system and the fluctuations found in the distribution of pronouns, something which in standard is perfectly established. In non-standard varieties we come across examples of *what* or *what's* instead of the above-mentioned *that* and *that's* to indicate possession (59); or the use of *what* and *which* in structures similar to the one with an anaphoric pronoun (60), (61). Jespersen (1909) even quotes examples of *what* used with personal antecedents from West Somerset (62). As for *which*, in Black English Vernacular and in some dialects of New York City, it is assuming the role of *that* and acts as an invariable particle. See (63), with the absence of the standard *at* after *leaves*:

(59) The man what's leg is broken...
(60) The man what his brother is a doctor...
(61) The woman which her name I can't remember....
(62) It's his mother wot has come
(63) The time which the bus leaves is 9.30

Finally, the scarce occurrence of non-restrictive clauses in more pragmatic

varieties must be recalled. As we saw in examples (19)-(21), coordinated structures are then preferred.

To sum up this section, one must conclude that: a/ the distribution of relative pronouns and strategies is highly variable and non-stable; b/ the role of pragmatic and communicative factors reveals itself as essential in the selection and distribution of relativization techniques and in the distribution of these and other correferential structures that do not involve subordination. This may be observed in the frequency of occurrence of zero-relative; in the abundant use of pronoun-retention; and in the way that least accessible positions in the hierarchy are treated. It has been also proved that c/ the connection between correferential deictic pronouns and relatives is clearer in non-standard varieties; and finally, d/ that in these varieties the use of just a relative marker or complementizer is much more widespread.

The fact that all these features so briefly outlined here characterize earlier stages of the English language brings into view the need to consider them in a historical analysis of RC's.

4. In this last section our attention will be devoted to the diachronic process undergone by RC's in English, setting off from OE.

4.1. There are many problems entailed in the analysis of any OE structure: the information we get is far from complete or ideal and so are the explanations offered for the interpretation of the data. These rely most of the times on modern categories and concepts which do not quite fit the earliest stages of the language. In any case, we will try to collect all data available [5] and filter them through the parameters established above.

In OE, as is well known, several strategies of relativization are found:

* *Pronominalization*: many of the so-called RC's are apparently introduced by the pronoun *se, seo, ðæt*, which is also the main deictic element. That is why at times it is impossible to decide whether we are facing a RC or a simply juxtaposed correferential structure. Very often examples such as (64) and (65) have been classified on punctuation criteria, wholly unreliable, since punctuation has been rhetorical and not grammatical up to modern times.

> (64) An ðeodwitta wæs on Britta tidum, Gildas hatte, *se* awrat one heora misdædum
> (65) ...e mon hæt æt Hædum; *se* stent...

* *Complementizer / relativizer* ðe. In OE this particle occurs frequently and steadily from the end of the 10th c. onwards. It is the most widely used technique, that indicates merely an existing and close relationship between the NP and the clause.

* A combination of the first two: *pronoun* (with anaphoric reference) + *ðe*, sometimes referred to as "emphatic relative":

(66) Datia ða ðe iu wæron Gotan...

* *Pronoun retention* in clauses introduced by ðe. It is often used in poetry for rhythmical reasons; also emphatically, or to avoid ambiguity on the referent. It is, on the other hand, the only way of relativizing the function of possessor:

(67) se mon ðe ic *hine* seah...
(68) Eadig bið se wer ðe *his* tohoðan bıð to Drihtne
(69) Dol bið se ðe *him* his dryhten ne ondrædeð

* Finally, *zero-relative* (70). There is an open disagreement as to how frequent it may have been; one must not forget that the texts available belong most of them to formal varieties:

(70) Frencisce abbot Turolde wæs gehaten

* There were certainly other techniques -*as*, *swylc*-, but being of minor importance, they will be disregarded here.

Traditionally no explanation is given to this cold collection of data, no general principle governing the distribution of all these techniques, clearly other than the one in PdE. However, considering what has been stated in previous sections, it seems reasonable to defend a global underlying principle —mainly pragmatic and communicative—, similar to the one found in non-standard varieties, and powerful enough to explain RC's from OE all through the 18th c. According to this, the proliferation of (explicit) marks may be said to depend on how difficult it is to recover certain information; on how complex it is to relativize a given syntactic function; on how much new information the speaker wants to add. Much more specifically, on how tight the link between antecedent and RC is. This concept of *tightness of link* between head and RC proves essential to understand the distribution of relativization strategies in English.

Tightness is a gradient determined by two components: a) stronger or weaker dependency between head and clause, and b) structural and semantic complexity. In particular, tightness may be said to depend on several factors:

1) on the identifying role of the RC. Both elements (head and clause) will be so much more closely linked, much more interdependent, when the RC gives the referent its lexical uniqueness. We must bear in mind that the concept of restrictiveness is a gradient in itself and that between sheer identification and new information being added on a head, there are afterthoughts, need of re-identification, physical distance between elements, desire for emphasis or contrast, etc.

2) most probably, on the meaning of the antecedent. Proper nouns are not the same as common nouns, regarding univocity of reference and, therefore, identification. But, as stated above, univocity of reference is something that varies

according to the linguistic system involved, the cultural setting and even the text itself.

3) on the difficulty to perceive the syntactic function involved in the process. Subjects can, therefore, be supposed to be much more closely linked to their clause than prepositional objects. Complexity is much lower and the relationship can easily be inferred.

With all these parameters in mind and without disregarding other possible sequential constraints, the following general principle can be established: *the tighter the link, the smaller the number of signals* that are necessary in surface structure; and vice versa: *the looser the link, the more explicit the structure must be*.

This can be proved considering that in OE δ*e*/ zero-relative are used in tight contexts; *se, se δe*, as well as pronoun retention, in looser ones. And what is more, it seems that extreme tightness or extreme dependency tends to be expressed —transparently— through formal subordination, the final position of the verb being perhaps a temporary sign of it. On the other hand, if the link between a NP and the following clause is not so tight, a structure similar to that of the main clause is preferred in OE, with the verb in medial position. However, in the gradient suggested, the potential element combinations are almost endless. Let's have a look at examples:

(71) Wat se δe cunnaδ

(72) ...sum δara monna δ*e* in δam here weorδuste wæron

(73) ...δas land, δ*a* synd hatene ærest Blecinga

(74) ...se betsta *hwælhunt*δa, δ*a* beδo eahta and feowertiges elna lang.

(75) Datia, δ*a* δ*e* iu wæron Gotan...

(76) ...δe fotspure δe wæs undernæden his fote δ*æt* wæs eall of read golde...

(and)...brohton dune δæt næcce δe δær wæs behid, *hit* wæs eall of gold and of selfre

(77) ...on δam mere δ*e* Truso stande *in sta*δ*e*

To get a whole picture of the OE system of relativization, however, we should have in mind not only these scattered examples, but some numerical evidence. Most of the structures analyzed for this study present a high degree of restrictiveness, i.e. perform an identifying function; and, out of these, some 80% corresponds to relativization of a human subject or of a direct object, the two left-most positions in the AH. There is no formalization for the genitive other than pronoun retention, and the prepositional object usually requires the retention of the adverb (often interpreted as postposition) in a final slot.

On the other hand, as is well known, intersentential dependency is not fully categorized in OE. Therefore, very often we come across structures similar, if not identical, to the ones referred to above, which have been traditionally classified as adverbial or appositional nominal clauses, following contemporary criteria. These

structures illustrated in examples (78) to (83) are also introduced by the complementizer *e* and include a deictic antecedent with cataphoric reference, closey linked to the following defining clause. This fact suggests that the term RC itself might even be inadequate when analyzing OE structures, and should not, therefore, be used:

(78) ...iii ger ðæs ðe hie on Linemuan comon hiðer...
(79) ...ða *hwile* ðe hie hira corn geryðon...
(80) ...ðo ð*one* dæg ðe hi hine forbærnæd...
(81) ...ðam middanwintre ðe se cyng forðferde...
(82) ...ða comon *forðy* onweg ðe ðara oðerre scipu asæton...
(83) ...ær ð*æm* ðe hit eall forhergod wære ond forbærned

4.2. Moving forward into Middle English (ME), a first stage of noticeable changes, already initiated in OE, results in a rather complex distribution of forms. In this period three relativization strategies can be said to exist: 1) complementizers: ðe, ðet/ðat; 2) zero-relative; and 3) retention of anaphoric pronouns. *That* will soon prove to be overladen with functions and —from examples similar to those in (84) and (85)— opens the way into the system for indefinite and interrogative pronouns.

(84) I didn't see whom he struck
(85) Who kills a man should die

From the 13th c. onwards relativization techniques are almost the same as in OE, the only important difference being that pronominalization, once consolidated, implies specific subordination. The picture in ME confirms the pragmatic use of relativization strategies ranging from the least to the most explicit: zero-relative is much more widespread than in OE and is used for leftmost functions in the scale; *that*, as complementizer, is still the most common of all techniques for unmarked contexts. The retention of anaphoric pronouns, finally is still reserved for syntactically more complex structures: genitives and prepositional objects.

Little by little, however, through most formal and written varieties, wh-pronouns come to be used. They will be gradually spread along a stylistic continuum up to their generalization at the end of Early Modern English (EME). Most important of all, though, is the specific rhythm with which they were introduced and their particular distribution within the ME system. *Which* is the first pronoun to find its way into the system. It is recorded from the end of the 13th c., but it appears almost exclusively in cases of oblique objects and genitives, i.e. in positions difficult to relativize, according to the AH, when it came to represent and explicit mark. The introduction of *whom* and *whose* is dated in the 15th c. and we will have to wait until the 16th c. to see *who* included in the system of English relatives. This pacing proves how relativization techniques and marking depends on communicative criteria. It is

difficult functions that are served first. *Who* enters the system very late because, being the subject the most accessible position in the scale, it was originally unnecessary to mark it with other specific elements. On the other hand, wh-pronouns come to appear mainly in clauses with a low degree of restrictiveness, i.e. in those marked contexts that require a more explicit signalling.

ME is also characterized by formal instability and inconsistencies. To give just some noticeable examples, *that* is still felt as general subordinator and is added almost to anything: *there that*, *which that*, etc. *What* is sporadically used instead of *that* (as we saw it happened in dialects). And combinations of wh-relatives, articles and nouns are frequent, in a clear effort to make anaphoric correference more explicit: *the which books that...* .

In short, we can say that in ME the language has not reached yet the stage of full syntactic codification of relative strategies. The process has only timidly begun, if at all, and the distribution of the different existing techniques is still constrained by communicative principles. Clauses of a high degree of restrictiveness constitute still the great majority of the structures used.

4.3. Our next stage, that of EME, can be defined as a a period of morphosyntactic upheaval. In general terms, two systems —one more pragmatic, the other more syntactic— coexisted in a not too stable and defined relationship. There are many well known intra- and extra-systemic factors which contributed to this situation. Within EME, however, two periods can be distinguished: a first one up to the 17th c., and another one closer to the 18th c.. This was characterized by prescriptivism and a purist struggle that would eventually decide on the system of relatives in contemporary English.

In the first stage the complementizer *that* is still the most common relativization technique, with some 90% of the occurrences. It gets weakly restricted to non-human antecedents, though its significant fall will occur around 1800. Zero-relative becomes more and more frequently used even in positions defined as "difficult to relativize". The retention of anaphoric pronouns is still maintained for cases of referential distance between antecedent and postmodifying clause, or to express possession and complex complementation, as in (86) where a second relative, and not a personal pronoun, appears anaphorically. *What* gets out of use as relative pronoun; the use of *which* is remarkably increased for mainly non-restrictive contexts, and whatever the nature of the antecedent. *Who*, in turn, appears timidly for cases in which the antecedent is a proper noun distant from the usually non-restrictive clause. It will be through the 18th c. that *who* and *which* acquire certain specialization according to the feature [+ human] of the antecedent.

(86) She that from whom we are all sea-swallowed

In fact, most of the rules and constraints affecting the system of relatives in contemporary English stem from two main debates carried through in the 18th c..

One concerns the use of *who* and *which* versus *that* and the distribution of the first two according to the feature [+ Human]; the other refers to the distribution of *who* and *whom*. It was not clear whether *whom* should be used as object or only as prepositional object. *That* and zero-relative, in turn, were considered colloquial barbarisms, and pronoun retention together with final prepositions vulgar and substandard speech. Relativization would have been restricted therefore to pronouns distributed according to the animate/ inanimate nature of the antecedent.

Obviously not all these constraints succeeded, not even in formal standard registers. The 19th c. shows the effects of these 18th c. discussions, and in the 20th c., as we have seen, it is the nature of the antecedent and the bipolar division of clauses into restrictive and non-restrictive what gives shape to relativization.

5. After the necessary theoretical discussion on RC's and the interpretation of their diachronic evolution just over-viewed, the following conclusions might be drawn:

1. What we now call RC results from the convergence of two originally different semantic and pragmatic functions, each of which is placed at one end of a gradient, which we define as tightness of link between a head and its corresponding clause. The link is so much tighter or looser, according mainly to: (a) the identifying nature of the clause, which constitutes a gradient in itself (from strict identification through the textual need of re-identification to the adding of new information —emphasis and contrast being included here—); (b) the nature of the antecedent, whether it is a pronoun, a noun or a whole clause; (c) the semantics of the antecedent; (d) the syntactic role of the antecedent within the clause. The identifying function is determined, in turn, by the parameters mentioned under (b) through to (d).

2. The basic meaning of RC's is apparently that of tightest link, these structures being therefore considered unmarked.

3. English up to modern times and current varieties different from British standard reflect the distinction between marked and unmarked structures through what we might call a rather transparent use of relativization techniques; i.e. the more marked a structure is, the more explicit signals will be used to point out the underlying complexity. The scale of syntactic complexity ranges from the absence of any open mark in surface structure, or the use of complementizers, to the retention of an anaphoric pronoun within the clause, or the combination of two elements.

4. Relatives have been constrained up to modern times by pragmatic and communicative factors, which has been proved through the following data: (a) the distribution of relativization techniques and the use of formal dependency on most unmarked OE structures; (b) the gradual spread of wh-pronouns from more to less marked structures; and (c) by the distribution of different coexisting resources in modern English, according to the formal/ informal character —or what is the same, the pragmatic or syntactic orientation— of the emerging varieties, etc.

5. Not all languages are equally syntactized; i.e. are able to relativize economically and through specialised elements the same number of syntactic functions (90% of the OE relativization being restricted to leftmost positions in the AH); nor can they all distinguish with the same degree of accuracy in surface structure between what we have called RC's and other related structures placed at either end of the continuum of tightness. I.e., in OE RC's cannot be properly separated on the one hand from the so-called adverbial and appositional nominal clauses; and, on the other, from independent correferential merely coordinated structures.

6. 18th c. prescriptivism and the desire to achieve and impose on English the logics of Latin (together no doubt with an ongoing tendency of the language itself) favour the substitution in standard of a probably weakened system, and the establishment of another, similar to the one we have today. In it gradients disappear and strict bipolar options decide on slightly different syntactic forms. Relativization techniques get reduced almost only to pro-nominalization.

7. A last marginal conclusion: instability and formal inconsistencies observed in ME and EME are identical to those found in dialects and non-standard varieties, as is the tendency to the overgeneralized use of complementizers whenever possible.

Notes

[1] This article is mainly based on a postgraduate seminar I taught in 1988 and 1989. The seminar was centred on the analysis of Old and Middle English selected prose texts. Data on Early Modern English rely on personal testing and on the bibliography available.

[2] See grammars in the bibliography or any other available.

[3] See also Keenan 1975.

[4] See Rissanen 1984, Dekeyser 1984, Romaine 1982 and 1983, and the work being carried out at the English Department of the University of Helsinki.

[5] A particularized bibliography for this section is considered unnecessary. I refer the reader to usual textbooks and general works on Old, Middle and Early Modern English syntax.

References

BOLINGER, D. (1968), *Aspects of Language*, N. York, Harcourt, Brace & World.

COLE, P. & SADOCK, J.M. (eds.) (1977), *Syntax and Semantics. Grammatical relations*, N. York, Academic Press.

COMRIE, B. (1981), *Language Universals and Linguistic Typology*, Oxford, Basil Blackwell.

DAHL, O. (1974), "Topic-Comment Structure Revisited", in Dahl, O. (ed.).

DAHL, O. (ed.) (1974), *Topic and Comment; Contextual Boundness and Focus*, Hamburg, Helmut Buske.

DANES, F. (1976), "Zur semantischen und thematischen Struktur des Kommunikats", in Danes F. & Viehweger D. (eds.).

DAVENPORT, M. et al (eds.) (1983), *Current Topics in English Historical Linguistics*, Odense, U.P.

DEKEYSER, X. (1984), *"Relativizers in Early Modern English: a Dynamic-quantitative Study"*, in Fisiak, J. (ed.).

FASOLD, T. & SHUY, R.W. (eds.) (1976), *Analyzing Variation in Language*, Washington, Georgetown, U.P.

FISIAK, J. (ed.) (1984), *Historical Syntax*, The Hague, Mouton.

GIVON, T. (1976), "Topic, Pronoun and Grammatical Agreement", in Li, C. (ed.).

GIVON, T. (1979), "From Discourse to Syntax: Grammar as a Processing Strategy", in Givón, T. (ed.).

GIVON, T. (1983), "Spoken English", in Givón, T. (ed.).

GIVON, T. (ed.) (1979), *Syntax and Semantics: Discourse and Syntax*, N. York, Academic Press.

GIVON, T. (ed.) (1983), *Topic Continuity in Discourse: a Quantitative Cross-Language Study*, Amsterdam, John Benjamins.

GRIMSHAW, J.B. (1985), "Evidence for Relativization by Deletion in Chaucerian Middle English", in *Papers in the History and Structure of English*, University of Massachussets.

HUDDLESTON, R. (1988), *English Grammar: an Outline*, C.U.P.

JACOBSON, B. (1977), "Adverbs, Prepositions and Conjunctions in English: a Study in Gradience", *Studia Linguistica, XXXI-I*.

JESPERSEN, O. (1909), *A Modern English Grammar*, Heidelberg.

KEENAN E.L. (1975), "Variation in Universal Grammar", in Fasold T. & Shuy, R.W. (eds.)

KEENAN, E.L. & COMRIE, B. (1972), "Noun Phrase Accessibility and Universal Grammar". Paper presented at the Linguistic Society of America meetings (Winter), and at the Linguistic Association Meetings (Hull, 1973).

KUNO, S. (1976), "Subject, Theme and the Speaker's Empathy. A Reexamination of Relativization Phenomena", in Li, C. (ed.).

LASS, R. (1987), *The Shape of English*, London, J.M. Dent & Sons.

LEECH, G. (1969), *Towards a Semantic Description of English*, London, Longman.

LEECH, G. (19), *A Communicative Grammar of English*.

LEHMAN, C. (1986), "On the Typology of Relative Clauses", *Linguistics*, 24-4.

LI, C. (ed.)(1976), *Subject and Topic*, N. York, Academic Press.

LINDE, C. (1979), "Focus of Attention and the Choice of Pronouns in Discourse", in Givón, T. (ed.).

McINTOSH, A. (1948), "The Relative Pronouns þe and þat in Early Middle English", *English and Germanic Studies*.

NORRICK, N.R. (1979), "The Lexicalization of Pragmatic Functions", *Linguistics*, 17 7/8.

OCHS, E. (1979), "Planned and Unplanned Discourse", in Givón, T. (ed.)

OJEA LOPEZ, M. (1987), "Relativas apositivas en español: ¿Tienen propiedades?", Paper presented at the Congreso de la Sociedad Española de Lingüística.

PALMER, H. (1976), *Grammar of Spoken English*, C.U.P.

QUIRK, R. et al. (1985), *A Comprehensive Grammar of the English Language*, London, Longman.

RISSANEN, M. (1984), "The Choice of Relative Pronouns in 17th c. American English", in Fisiak, J. (ed.).

ROMAINE, S. (1982), *Socio-historical Linguistics*, C.U.P.

ROMAINE, S. (1983), "Syntactic Change as Category Change by Re-analysis and Diffusion: Some Evidence from the History of English", in Davenport, M. et al. (eds.).

SANKOFF, G. & BROWN, P. (1976), "The Origin of Syntax in Discourse", *Language*, 52, nb.3.

SINCLAIR, J. (1972), *A Course in Spoken English*.

STAHLKE, H. (1976), "Which That", *Language*, 52, nb.3.

STOCKWELL, J. (1973), *The Major Syntactic Structures of English*, N. York, Reinhart & Winston.

TOMLIN, R. (1985), "Foreground-background Information and the Syntax of Subordination", *Text*, 5, 1-2.

TRAUGOTT, E.C. (1972), *The History of English Syntax*, N.York, Holt, Reinhart & Winston.

WINTER, E. (1982), *Towards a Textual Grammar of English*, London, George, Allen & Unwin.