RELATIVE CONNECTIVES IN PRESENT-DAY ENGLISH: A CORPUS-BASED ANALYSIS

FRANCISCO JAVIER DÍAZ PÉREZ
UNIVERSIDAD DE SANTIAGO DE COMPOSTELA

1. INTRODUCTION

The complexity entailed by the description of relative clauses in English has resulted in a great number of studies devoted to different aspects of this type of construction. One of the topics which has aroused most interest is the examination of the factors which condition the selection of the relative pronoun. Grammatical, stylistic, and psychological or cognitive aspects combine to determine to a greater or lesser extent the choice of the connective which introduces the relative clause. The purpose of this study is to shed some light on the subject of pronouns introducing relative clauses. My data are the 180 relative clauses contained in three files of the Lancaster Parsed Corpus of British English (an abridged version of the Lancaster Oslo/Berger Corpus. These three files are LPCA (8188 words), which contains 728 clauses from written press, LPCJ (7537 words), containing 450 clauses extracted from scientific English, and LPCK (11279 words), which consists of 1119 clauses from general literary texts. The whole corpus, therefore, amounts to 2297 clauses and 27004 words. The reason for using this corpus is that it constitutes a fairly representative sample of present-day written

Miscelánea: A Journal of English and American Studies 18 (1997): 69-88.

English. On the one hand, it is a homogeneous corpus, since all the clauses belong to a formal written type of language, but on the other hand, three different major styles—literary, journalistic and scientific—are represented, which allows a comparative study of the way the stylistic differences affect the choice of relative pronoun. This paper partly draws on Yamashita (1994), but its focus and objectives are different. The aim of Yamashita's paper is to analyse the language processing principles by means of which the processing effort involved by relative clauses can be minimized. In our paper, however, it is the factors which condition the selection of the relativizer that have been focused on. Some of these factors are determined by strategies which facilitate the production and understanding of relative clauses, but other aspects, such as for instance stylistic characteristics, have been found to be relevant. On the other hand, while the corpus used in Yamashita (1994) corresponds to the spoken language, ours has been extracted from written material. This fact may be responsible for some differences in the results obtained in both papers. In section 2 grammatical factors conditioning the selection of relative pronoun are dealt with; section 3 concerns itsef with cognitive strategies which affect the choice of relative connective, namely the closure strategy and the non-ambiguity strategy; section 4 is devoted to the analysis of the effect of stylistic characteristics over the distribution of relative connectives. Finally, section 5 summarizes the main conclusions reached in this study.

2. GRAMMATICAL FACTORS

The most frequent relative linkword in this corpus is *which*, with 69 occurrences (= 38.33%), followed by zero, which occurs 44 times (= 24.44%) and at a distance by *who* and *that*, with 21 occurrences each (= 11.76%). These data indicate that, as an alternative to *which* and above all to *who*, *that* is not chosen very often, which is probably a consequence of the fact that all the relative clauses of this corpus have been extracted from written material. The general data of the frequency of the relative linkwords in the corpus is presented in Table 1.

	WHO	WHICH	WHEN	WHERE	WHOM	WHOSE	THAT	ZERO	AS
N	21	69	10	9	3	2	21	44	1
%	11.76	38.33	5.55	5	1.67	1.11	11.76	24.44	0.55

Table 1. Frequency of the relative linkwords.

One problem which has to be faced in an investigation of this type, as Taglicht points out (1973: 329), is the separation of the restrictive from the nonrestrictive relative clauses, since there is no single criterion that can be considered satisfactory on its own. Phonological and punctuation criteria have proved insufficient to distinguish between both kinds of relative clauses. Thus, for instance, several examples have been found in this corpus in which, contrary to the prescriptive rule, the nonrestrictive clause is not preceded by a comma:

(1) *the mass of oxygen per year at this rate would be approximately 8x10 g/year which is a quite insignificant quantity.*

(2) *she was afraid and could only hope that the girl had gone to Erich who loved her, however hopelessly.*

The converse phenomenon, the insertion of a comma before a restrictive relative clause, is much more infrequent and the only example found in this corpus is due to the introduction of a prepositional phrase between the antecedent and the relative clause:

(3) *it was fate, in fact, that was making fools of all of us.*

The punctuation criterion, therefore, was relied on only in those cases in which either the grammatical structure or the context could not offer a better tool. The distribution of restrictive and non restrictive relative clauses is the following: there are 143 restrictive relative clauses, which amount to 79.44%, and 37 nonrestrictive relative clauses, which represent 20.56%.

As could be expected, nonrestrictive relative clauses are introduced by *wh-* pronouns or adverbs. There is not any nonrestrictive clause which is introduced by *that* or zero. As regards restrictive clauses, although there are more *wh-* clauses than *that* or zero clauses, the difference is not very significant. Seventy-seven relative clauses contain a *wh-* pronoun or adverb, and 65 have *that* or zero as relative connective. Table 2 shows the distribution of the relative pronouns in restrictive relative clauses.

The choice of relative pronoun in nonrestrictive relative clauses is much more limited, since, as has been said above, the use of *that* and zero in this type of clauses is extremely rare, and in this particular corpus no instance of these connectives has been found. The distribution of relative connectives in nonrestrictive relative clauses is presented in Table 3.

WHO		WHICH		WHERE		WHEN		WHOSE		WHOM		THAT		ZERO		AS	
N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
12	8.39	48	33.57	5	3.50	9	6.29	1	0.70	2	1.40	21	16.48	44	30.77	1	0.70
77 (53.85%)												65 (45.45%)				1 =0.7%	

Table 2. Distribution of relative linkwords in restrictive relative clauses

WHO		WHICH		WHERE		WHEN		WHOM		WHOSE	
N	%	N	%	N	%	N	%	N	%	N	%
9	24.32	21	56.76	4	10.81	1	2.70	1	2.70	1	2.70

Table 3. Distribution of relative linkwords in nonrestrictive relative clauses

SUBJECT										OBJECT							
WHO		WHICH		THAT		ZERO		AS		WHICH		THAT		WHOM		ZERO	
N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
21	25.61	44	53.66	16	19.51	----	0	1	1.22	9	19.57	2	4.35	1	2.17	34	73.91
82 (45.55%)										46 (25.55%)							

Table 4. Functions of relative pronouns and relative adverbs

PREPOSITIONAL COMPLEMENT										DET.	
WHO		WHICH		THAT		WHOM		ZERO		WHOSE	
N	%	N	%	N	%	N	%	N	%	N	%
----	0	16	72.73	1	4.55	2	9.09	3	13.64	2	100
22 (12.22%)										2 (1.11%)	

Table 4 continued.

ADVERBIAL									
WHERE		WHEN		THAT		ZERO		WHICH	
N	%	N	%	N	%	N	%	N	%
9	31.03	10	34.48	2	6.90	7	24.14	----	0
28 (15.55%)									

Table 4 continued.

Apart from the absence of *that* and zero, there are not many other divergent aspects in the two types of relative clauses with respect to the distribution of the relative connectives. If the percentages of *who* and *which* in nonrestrictive clauses are added up (81.08%) and compared with the percentages of *who, which, that* and zero in restrictive clauses (87.41%), the results are fairly equivalent. Assuming, as De Haan (1991) does, that zero and *that* are alternatives for *who* and *which*, the implications are that relative pronouns of basically the same sort are used in both types of relative clauses. There is a difference with relative adverbs, since *where* is more frequently used in nonrestrictive clauses and *when* in restrictive ones. However, due to the small sample of nonrestrictive clauses, these results cannot be considered very significant.

The function of each relative pronoun and relative adverb has been analyzed and the relation between syntactic functions and relative connectives is depicted in Table 4. All the functions fulfilled by the different relative connectives are illustrated in the following examples from our corpus:

SUBJECT:

(4) *Labour should not take any steps [which would appear to prop up an out-dated institution].*

DIRECT OBJECT:

(5) *he quoted from this letter [which Gaitskell had received].*

PREPOSITIONAL COMPLEMENT:

(6) *but governments should be free to negotiate and refuse proposals [with which they did not agree].*

DETERMINER:

(7) *I feel rather like a father [whose child is bleeding to death].*

ADVERBIAL:

(8) *they were a reminder of the time [when the districts had been little hamlets before they were swallowed up in London's vast sprawl].*

(9) *he was ashamed of the way [() she credited other people with her own memory].*

When the position occupied by the relative connective is that of subject, *wh-* pronouns tend to be selected against relative *that*. While *who* acts as

subject on 21 occasions, which amount to 25.61% of the subjects, and *which* on 44 occasions (= 53.66%), there are only 16 examples (= 19.51%) in which *that* is subject. An aspect which is worth mentioning is the fact that when the subject relative has a personal antecedent, it is only *who* that occurs as relative pronoun in our corpus. In none of the 16 examples in which relative *that* functions as subject the antecedent is personal. There is not any example in this corpus either in which the zero relative pronoun occupies the subject position.

The relative pronoun which most often functions as direct object is zero, with 34 examples, which represent 73.91% of the object relative pronouns. *Which* is direct object on 9 occasions (= 19.57%), *that* on 2 (= 4.35%), *whom* on one (= 2.17%), and *who* does not fulfil that function in this corpus. The total number of objective relative pronouns is 46. Of these 46 relative pronouns, only 3 have a human antecedent; two are zero relative pronouns and the other one is *whom*. With such a small sample, no final conclusions can be drawn with regard to the choice of relative pronoun acting as object when the antecedent is human.

SUBJECT		OBJECT		PREP. COMP.		ADVERBIAL		DETERMINER	
N	%	N	%	N	%	N	%	N	%
82	45.55	46	25.55	22	12.22	28	15.55	2	1.11

Table 5. Frequency of the syntactic functions fulfilled by the relativizers.

Among all the possible syntactic functions (see Table 5), subject is that which is most often fulfilled by the relative pronouns, followed by direct object, adverbial, complement of preposition and determiner. In 82 examples the relative pronoun is the subject of its clause, which means 45.55% of the total. The number of relative pronouns which function as object is 46 (= 25.55%). Twenty-two relative pronouns (= 12.22%) fulfil the function of prepositional complement, 28 (= 15.55%) that of adverbial, and 2 are determiners (= 1.11%). In this connection, it has been suggested that those clauses in which a subject has been relativized are easier to process than those in which a genitive, for instance, has been relativized. Thus, Keenan and Comrie introduced the theory of the Accessibility Hierarchy, according to which certain syntactic functions are more easily relativizable than others.

Pronoun		which	who	where	when	whose	whom	that	zero
	N	4	1	---	---	---	---	3	9
	%	23.53	5.88	0	0	0	0	17.65	52.94
Proper noun	N	---	7	3	---	---	---	---	1
	%	0	63.64	27.27	0	0	0	0	9.09
Count singular	N	33	9	3	6	1	1	5	21
	%	42.31	11.54	3.85	7.69	1.28	1.28	6.41	26.92
Count plural	N	20	2	3	2	1	1	8	6
	%	45.45	4.54	6.82	4.54	2.27	2.27	18.18	13.64
Non-count	N	11	2	---	2	---	1	4	7
	%	40.74	7.41	0	7.41	0	3.70	14.81	25.92

Table 6. Distribution of linkwords by realisation of noun phrase antecedent.

They have established a scale in which the last positions can be relativized if and only if the previous positions in the scale can too. This is, according to them, a universal principle which applies across languages:

> We argue that languages vary with respect to which NP positions can be relativized, and that the variation is not random. Rather, the relativizability of certain positions is dependent on that of others, and these dependencies are, we claim, universal. The Accessibility Hierarchy (AH) below expresses the relative accessibility to relativization of NP positions in simplex main clauses. Accessibility Hierarchy (AH) SU>DO>IO>OBL>GEN>O COMP (1977: 66).

If the occurrences of relevant examples in our corpus are counted, the results will be: Subject: 82 (45.55%), Object: 46 (25.55%), Oblique: 22 (12.22%) and Determiner: 2 (1.11%). The data of our corpus confirms, therefore, the Accessibility Hierarchy.[1] As Schmied indicates, this hierarchy is related to the psychological principle of 'right-branching', according to which "the easiest structure is when a subject relativizer is attached to the last matrix noun phrase" (1993: 89).

Another aspect which influences the choice of the relative linkword is the realization of the head of the noun phrase, that is to say, whether it is a pronoun, a proper noun, a singular count noun, a plural count noun or a non-count noun. Table 6 shows the distribution of the relativizer over the realisation of its antecedent noun phrase. *Which* scores high on singular count nouns, with 33 occurences and 42.31% of the relative pronouns which have a singular count noun antecedent. It is followed in this category by the zero relative pronoun, which appears in 21 examples (= 26.92%). *Who*, on the other hand, is used after more than half of all proper nouns (7 out of 11= 63.64%), followed by *where*, used in 3 examples (= 27.27%) and zero, which appears in what constitutes an exceptional example (= 9.09%):

> (10) *in the large airport bus she had a better view of the London () she had not seen for over two years.*

With respect to plural count nouns, the most frequent relative pronoun is *which* (20 cases and 45.45%) followed at a great distance by *that* (8 cases= 18.18%). With pronoun antecedents, the zero relative pronoun scores high, since it is used in 9 of the 17 examples (= 52.94%). As regards the category formed by those examples in which a non-count noun is the antecedent, *which* is the most frequent pronoun (11 cases= 40.74%), closely followed by zero (7 cases= 25.92%).

3. COGNITIVE STRATEGIES

The relative clause constitutes one of the possible constituents which can function as post-modifier of a noun phrase. Adverbs and non-finite clauses can also be noun phrase post-modifiers. However, relative clauses add information which neither adverbs nor non-finite clauses can provide. According to Quirk et al., "part of the explicitness of the relative clause lies in the specifying power of the relative pronoun" (1985: 1245). Yamashita (1994: 73) believes that the facts that the relative pronoun concords with its antecedent in gender and that it has a function within the relative clause on the one hand result in more explicit information, but, on the other hand, create difficulties in language processing, both from the perspective of production and of understanding. However, language has developed some linguistic principles which have helped to minimize the problem, among which we encounter some cognitive strategies. These cognitive strategies often influence the choice of the relative pronoun and they are also related to the function which the relative pronoun fulfils.

3.1. The Closure Strategy

One of the cognitive strategies which simplifies the production and understanding of relative clauses is the closure strategy, which Prideaux and Baker define in the following terms:

> CLOSURE. In processing a particular linguistic unit (phrase, clause, etc.), the language user (speaker or hearer) attempts to obtain closure on that unit as early as possible. (1986: 32)

The closure strategy predicts that non-interrupted structures, like example 11 below, are easier to process than interrupted ones (example 12).

Final: (11) *boiling the effluent with more than 5 ml. of acid resulted in gravimetic recoveries [which were too low to be tolerated].*

Medial: (12) *she found no difficulty in assuming for convenience that the attack [she had begun on Martin] hadn't happened.*

Therefore, the closure strategy conditions the position that the relative clause occupies with respect to the main clause, since it favours relative clauses in

final position as opposed to those which occupy medial position. In this corpus there are 155 relative clauses (= 85%) which occupy final position, and only 27 which are placed at medial position (= 15%). In this connection, Kuno argues that center-embedding reduces comprehensibility, and this "is related to the limitation of the human capacity of temporary memory" (1974: 120).

The prediction that non-interrupted sentences are more easily processable than interrupted ones implies that OS and OO sequences (in which the first letter indicates the function of the antecedent NP and the second one that of the relative pronoun) are easier to process than SS and SO types. To test this hypothesis, only those examples whose relative pronouns and antecedents function as subject, direct object or subject complement have been considered. The reason for including subject complements in the analysis is, as Yamashita points out (1994: 79), that object and complement take the same position, and in this analysis what is important, rather than the grammatical function, is the ordering or syntactic position. Examples of CS and CO in our corpus, included under the category of OS and OO are:

CS: (13) *stones in a soil profile are those things [that have failed to weather to form a soil].*

CO: (14) *there are sides to me [() you'll never know].*

Categories SC and OC have not been able to be included, since there is no example in our corpus in which the relative pronoun functions as subject complement. The hypothesis is confirmed by the data of the corpus (Table 7), since the number of OS sequences is 23, which added to the OO ones (29 ex.) make 52, whereas SS (12 ex.) and SO (5 ex.) types together constitute 17 cases. Examples of these four types are the following:

OS: (15) *you just wait till you meet the girl [who thinks you are a god].*

OO: (16) *he looked up into her eyes and saw the immeasurable depth of eternity [that God has put in there for man to lose himself].*

SS: (17) *A correspondent [who travelled yesterday within a few miles of Vang Vieng] was told by officers that this village was still held by the procommunists.*

SO: (18) *anything [which physical principles suggested might be relevant] was included, and the advice of experienced forecasters at London Airport was sought.*

SS		SO		OS		OO		TOTAL	
N	%	N	%	N	%	N	%	N	%
12	17.39	5	7.25	23	33.33	29	42.03	69	100

Table 7. Distribution of relative clauses by the types to which they belong.

Cofer defends the position that the deletion of the relative pronoun in medial relative clauses responds to a psychological principle, since it "aids production and comprehension, because it minimizes the interruption of the higher sentence, thus insuring continuity of form and meaning" (1975: 31). Therefore, the fact that in relative clauses which are embedded in the middle of another clause the zero relative pronoun is preferred over *wh-* relative pronouns and adverbs is also related to the closure strategy. This demonstrates that this cognitive strategy has an effect not only over the position of the relative clause and the function of both NP antecedent and relative pronoun, but also over the choice of the relative pronoun. The data obtained from medial relative clauses in our corpus are misleading at first sight, since there are 21 *wh-* relative pronouns and adverbs (12 *which*, 2 *when*, 1 *where*, 4 *who*, 1 *whom*, 1 *whose*), 7 zero relative pronouns, and 2 instances of *that*, but it must be noticed that some syntactic functions fulfilled by these pronouns, like subject, determiner of prepositional complement preceded by the preposition, cannot possibly be fulfilled by zero. If we consider relative pronouns which function as object, adjunct, and prepositional complement not preceded by the preposition, functions which all the relative pronouns can fulfil (Cf Quirk 1957: 107), the tendency to select zero against *wh-* pronouns in medial relative clauses is self-evident. As Table 8 shows, there are 7 zero relative pronouns,1 *which*, 1 *that*, 0 *who(m)*.

WHO(M)		WHICH		THAT		ZERO		TOTAL	
N	%	N	%	N	%	N	%	N	%
0	0	1	11.11	1	11.11	7	77.78	9	100

Table 8. Distribution of relative linkwords in medial clauses.

3.2. The Non-ambiguity Strategy

Non-ambiguity is another cognitive strategy which is at work to make the production and understanding of relative clauses easier.

> NON-AMBIGUITY. The language user assumes that the unit being processed is not ambiguous. (Prideaux and Baker 1986: 32)

The fact that the zero relative pronoun hardly ever functions as subject is a consequence of the non-ambiguity strategy. In other words, the fact that a relative pronoun functioning as subject cannot normally be omitted is due to ambiguity reasons, and not to its syntactic role. This can be proved by the fact that a relative pronoun which functions as direct object cannot be substituted by zero if this creates confusion. Thus, an example like ? *They are girls boys don't like to date* (Bolinger 1972: 12-13) is doubtfully acceptable, since it is a source of ambiguity, due to the tendency of nouns to function as premodifiers in English. It seems to suggest a compound *girls-boys* or a Saxon genitive *girls' boys*. According to this, the greater frequency with which the zero relative pronoun is used before a personal pronoun results from the fact that there is no combination noun-noun in that case which can lead to confusion. In this connection, when the relative connective is not a subject and it is followed by a subject personal pronoun, a marked tendency can be observed in this corpus to select zero as relative pronoun. As Table 9 displays, zero has been chosen on 37 occasions (= 68.52%), *which* appears in 8 examples (= 14.81%), *where* in 5 (9.26%), *whom* in 2 (3.70%) and *that* and *when* have been used once each (1.85%).

	WHICH	WHOM	WHERE	WHEN	THAT	ZERO	TOTAL
N	8	2	5	1	1	37	54
%	14.81	3.70	9.26	1.85	1.85	68.52	100

Table 9. Distribution of relative linkwords followed by a pronoun.

	WHICH	WHOM	WHERE	WHEN	THAT	ZERO	TOTAL
N	15	1	3	9	7	4	39
%	38.46	2.56	7.69	23.08	17.95	10.26	100

Table 10. Distribution of relative linkwords followed by a noun phrase.

However, when a noun phrase functioning as subject follows the relative pronoun (See Table 10), the tendency is the opposite one. The relative pronoun *which* is followed by a noun phrase subject in 15 examples (= 38.46%), *when* appears before a noun phrase on 9 occasions (= 23.08%), *where* on 3 (= 7.69%), and *whom* on 1 (= 2.56%). *Who* does not appear in this context, because it fulfils itself the function of subject in all its occurences. With respect to zero and relative *that*, the former is followed by a noun phrase subject in 4 examples (= 10.26%) and the latter in 7 (= 17.95%). Thus, while *wh-* relative pronouns and adverbs precede a noun phrase on 28 occasions (71.79%),*that* and zero do it on 11 occasions (28.21%).

Another consequence which derives from the non-ambiguity constraint is the influence which the length of the relative clause has on the selection of the relative pronoun or adverb (See Tables 11 and 12). In Cofer's opinion (1975: 31), the relation between length of clause and selection of relative pronoun is connected with a psychological principle, ease of sentence processing. In long relative clauses, the use of *who* or *that*, instead of zero, makes the structure and meaning of the clause easier to identify by the listener or reader. If the zero variant is selected, the relative clause structure may not be recognized until the place where the co-referential noun phrase used to be is reached, and by that moment, the beginning of the clause may have been forgotten. Short clauses (1-4 words, excluding the relative connective itself) tend to select zero relative pronoun, whose presence declines slightly in clauses of medium length, from 25 short clauses (= 55.55%) to 19 medium ones (= 42.22%). There is only one clause with more than ten words in

which zero occurs (= 2.22%). All in all, the average length of zero clauses is 4.59 words. This average goes up to 7 words in the case of *which* clauses. *Which* occurs mainly in clauses of medium length (35 ex.= 50.72%). On the other hand, the proportion of *which* clauses that are long (10 words or more) is much higher (15 ex.= 21.74%) than that of zero clauses, while, conversely, the percentage of *which* clauses that are short (19 ex.= 27.54%) is lower. As regards relative *that*, it occurs in the same proportion in short as in medium clauses (9 cases of each), while the occurrences of *that* in long clauses go down to 3. *That* clauses have an average length of 5.90 words. The average length of clauses with *who* is the highest, with 8.52 words, and the percentages of short, medium and long clauses with *who* are respectively 28.57% (6 occurences), 38.09% (8 occurences) and 33.33% (7 occurences). The analysis of these data reveals the length of the relative clause as an influencing factor in the selection of the relative pronoun.

	WHO		WHICH		THAT		ZERO	
	N	%	N	%	N	%	N	%
short	6	28.57	19	27.54	9	42.86	25	55.55
medium	8	38.09	35	50.72	9	42.86	19	42.22
long	7	33.33	15	21.74	3	14.28	1	2.22
Total	21	100	69	100	21	100	45	100

Table 11. Distribution of relative linkwords by the clause length.

	WHO	WHICH	THAT	ZERO
Average length	8.52 words	7 words	5.90 words	4.95 words

Table 12. Average lenth of the relative clauses according to the linkword.

The influence which the separation or non-separation of the relative pronoun from the antecedent has on the choice of the relative pronoun seems to be a consequence of the non-ambiguity strategy as well (Table 13). Although relative clauses normally follow their antecedents immediately, when this is not the case, *wh-* pronouns are preferred over *that* and above all over the zero relative pronoun. This fact reflects another cognitive strategy which helps to recognize the structure of the clause, since a relative clause which does not come immediately after its antecedent is more difficult to separate from the material that intervenes if there is no relative pronoun, and on some occa-

sions, it could produce ambiguity, as in *the man at the museum John admires is beginning to speak.*[2] (Cf. *the man at the museum whom John admires is beginning to speak.*). Of the 25 relative clauses in which the relative pronoun is not strictly adjacent to its antecedent, 10 are introduced by *which*, 4 by *who*, 3 by *when*, 3 by *that*, 3 by zero and 1 by *as*. The *wh-* series, with 17 occurences is therefore preferred over *that* and zero, which together total 6 occurrences.

	WHO	WHICH	WHEN	THAT	ZERO	AS	TOTAL
N	4	10	3	3	3	1	24
%	16.67	41.67	12.5	12.5	12.5	4.17	100

Table 13. Distribution of linkwords in relative clauses separated from their antecedent.

4. STYLISTIC FACTORS

As has been mentioned above, three different genres are represented in this corpus: Press, Fiction and Learned & scientific writings. The number of relative clauses which correspond to each of these genres are 51, 82 and 47 respectively. Each genre corresponds to a different style with its own defining characteristics. The aim of this section is to find out to what extent the stylistic differences condition the selection of the relativizer.

The distribution of the relative linkwords over the three genres is shown in Table 14. The percentage of zero relative pronouns in the fictional sample, which is equivalent to 41.46% of the relativizers (34 out of 82), is much higher than in the relative clauses extracted from scientific English, where only 4.25% of the relativizers are zero relative pronouns. The subcorpus which corresponds to a journalistic style occupies a mid-way position in this respect, with 8 zero relative pronouns of its 51 relativizers (= 15.69%). While in the Fiction group, zero is the most frequent relative pronoun, in the Press sample it is the third relativizer in frequency and in the science one it occupies the last position. The distribution of *which* is the opposite one. It is the prevailing relativizer in the scientific subcorpus, with 35 occurrences and 74.47%. The percentage in narrative texts goes down to 20.73% (17 occurrences), even though it is the second relativizer in frequency. The journalistic genre is again in the middle, since the percentage of *which*, with 17 instances, is 33.33%, making it the most used relative linkword. The differences with respect to the distribution of *that* are not so significant. The highest percentage corresponds to fictional texts (13.41%), followed by that of written texts taken from newspapers (11.76%); scientific writings occupy the last position (8.51%).

		WHO	WHICH	THAT	ZERO	OTHERS[3]	TOTAL
Fiction	N	12	17	11	34	8	82
	%	14.46	20.73	13.41	41.46	9.76	100
Press	N	9	17	6	8	11	51
	%	17.65	33.33	11.76	15.69	21.57	100
Science	N	0	35	4	2	6	47
	%	0	74.47	8.51	4.25	12.76	100

Table 14. Distribution of the relative linkwords across the genres.

If the distributions of the *wh-* series and the *that* and zero series are examined across the three genres (Table 15), it will be found out that the differences are quite outstanding. In the Fiction genre the *that*-zero series, with 54.88% overcomes the *wh-* series, with 45.12%. In the other two genres, on the other hand, *that* and zero are surpassed by the *wh-* pronouns and adverbs. However, there are still differences between these two genres, since, while in the scientific style the *that* and zero series (12.77%) is very widely overcome by the *wh-* series (87.23%), in the journalistic style the percentages which correspond to both series are slightly more equilibrated: 70.59% of the relative linkwords belong to the *wh-* series and 27.45% corresponds to *that* and zero.

	WH-SERIES		THAT & ZERO		TOTAL	
	N	%	N	%	N	%
Fiction	37	45.12	45	54.88	82	100
Press[4]	36	70.59	14	27.45	51	100
Science	41	87.23	6	12.77	47	100

Table 15. Distribution of the *wh-* and *that*-and-zero series across the genres.

A possible explanation for these divergences in the distribution of relativizers across the different genres is, as De Haan points out, that in fiction texts "very often the dialogue parts are aimed at representing a more collo-

quial style" (1991: 63), which explains the high proportion of zero relative pronouns and the predominance of the *that* and zero series over the *wh-* series. The scientific style, on the contrary, is very formal, and this fact is reflected in the overwhelming supremacy of *wh-* relative linkwords and the very few instances of zero relative pronouns. The journalistic style is also formal, but not as rigidly so as the scientific one. This results in a prevalence of the *wh-* series, but not as absolute as in the case of scientific English.

5. CONCLUSIONS

The main conclusions which can be drawn from this analysis can be summarized as follows:

1) The syntactic function which the relative pronoun fulfils clearly influences the selection of this relative pronoun. Thus, the zero relative pronoun never acts as subject, a function which is mainly realised by *which* when the antecedent is non-personal and by *who* when it is personal. When the relative pronoun occupies the object position, zero is the most frequent pronoun.

2) The most often relativized function is that of subject, which is followed by object, adverb, complement of preposition and determiner. These data confirm the Accessibility Hierarchy proposed by Keenan and Comrie. This hierarchy is nevertheless even more strongly confirmed by the data in Yamashita (1994). This may be due to the fact that the need to lessen the effort imposed on addresser and addressee to process relative clauses is more pressing in spoken than in written language.

3) The realization of the noun phrase antecedent also conditions the choice of the relativizer. Proper nouns tend to be followed by the relativizer *who*, and pronouns by the zero relative pronoun. Singular and plural count nouns, on the other hand, seem to prefer the relative pronoun *which*.

4) Relative clauses which occupy a final position are preferred over those which are embedded in the middle of a superordinate clause. When the relative clause is in medial position and the relative pronoun functions as object, adverbial or prepositional complement not preceded by the preposition, zero tends to be selected (7 zero, 1 *which*, 1 *that*). These two facts are related to the closure strategy, according to which OS and OO types in relative clauses are more frequent than SS and SO types, which is confirmed by the results of this corpus.

5) There is a relation between relative pronoun and length of the relative clause. The shortest clauses are those introduced by the zero relative pronoun, with an average length of 4.59 words, followed by *that* clauses (5.90 words)

and *which* clauses (7 words); the longest are clauses introduced by *who* (8.52 words).

6) Relative clauses strictly adjacent to their antecedent are preferred over those which are separated from it. But when the relative clause does not immediately follow its antecedent, *wh-* relative pronouns occurred in a higher percentage than zero or *that,* 17 occurences of *wh-* pronouns and adverbs and 6 occurences of *that* and zero.

7) In those cases in which the subject of the relative clause is a personal pronoun, zero, with 37 occasions (= 68%), is preferred over any other relative connective.

8) Conclusions 5, 6 and 7 are related to a cognitive strategy which makes easier the production and understanding of relative clauses, the non-ambiguity strategy.

9) Different styles are related to the choice of the relative linkword as well. While in the fictional style zero relative is the prevailing relative pronoun, in the scientific style, which is in the opposite pole, zero is very rarely used, and the prevalence of the *wh-* series is overwhelming.

NOTES

1. As Yamashita states, IO (indirect object) in English corresponds with OBL (complement of preposition). Therefore IO is included in OBL.

2.. Example taken from Cofer 1975: 31.

3. Under the category *OTHERS* the following relativizers are included: *whom*, *where*, *when*, *whose*, and, just in the Press sample, the conjunction *as* used as relative linkword.

4. In the Press subcorpus there is an example of the conjuction *as* used as relativizer which is not included in the table.

WORKS CITED

BOLINGER, D. 1972. *That's That*. The Hague: Mouton.
COFER, T. M. 1975. "Performance Constraints on Relative Pronoun Deletion."*Linguistics* 157: 13-32.
GARSIDE, R. G. N. LEECH and T. VÁRASI. *Manual of Information for the Lancaster Parsed Corpus*. Lancaster: U of Lancaster, n.d.
HAAN, P. J. M. de. 1991. "On the Exploration of Corpus Data by Means of Problem-Oriented Tagging: Postmodifying Clauses in the English Noun Phrase." In Johanson and Stenström 1991: 51-65.
JOHANSSON, S., G. N. LEECH and H. GOODLUCK, comp. 1978. *Lancaster-Oslo/Bergen Corpus of British English*. Oslo: U of Oslo.
JOHANSSON, S and A. B. STENSTRÖM, eds. 1991. *English Computer Corpora*. Berlin and New York: Mouton de Gruyter.
KEENAN, E. L. and B. COMRIE. 1977. "Noun Phrase Accessibility and Universal Grammar."*Linguistic Inquiry* 8.1: 63-99.
KUNO, S. 1974. "The Position of Relative Clauses and Conjunctions." *Linguistic Inquiry* 5: 117-136.
PRIDEAUX, G. D. and W. J. BAKER. 1986.*Strategies and Structures: The Processing of Relative Clauses*. Amsterdam: John Benjamins.
QUIRK, R. 1957. "Relative Clauses in Educated Spoken English." *English Studies* 38: 97-109.
QUIRK, R. et al. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
SCHMIED, J. 1993. "Qualitative and Quantitative Research Approaches to English Relative Constructions." In Souter and Atwell 1993: 85-96.
SOUTER, C. and E. ATWELL, eds. 1993. *Corpus-Based Computational Linguistics*. Amsterdam: Rodopi.
TAGLICHT, J. 1973. "The Choice of Relative Pronouns in Written English." *Scripta Hierosolymitana* 25: 327-336.
YAMASHITA, J. 1994. "An Analysis of Relative Clauses in the Lancaster/IBM Spoken English Corpus." *English Studies* 75.1: 73-84.

EL PROCESO DE CAPTACIÓN RACIONAL Y SU EXPRESIÓN LÉXICA EN INGLÉS Y ESPAÑOL

NATIVIDAD FERNÁNDEZ DE BOBADILLA LARA
UNIVERSIDAD DE GRANADA

1. INTRODUCCIÓN

El objeto del presente estudio es tanto delimitar la parcela semántica en la que se ubica el proceso de "captación racional" como ofrecer un estudio contrastivo de las unidades léxicas verbales que la conforman en inglés y en español mediante un análisis detallado de las mismas. Este análisis incluye información de tipo semántico y sintáctico y se va a llevar a cabo a nivel intra e interlingüístico, estableciendo las principales semejanzas y oposiciones a ambos niveles. Para ello seguimos las bases del método lexemático-funcional propuesto por L. Martín Mingorance (1984, 1989, 1990), modelo de análisis léxico y descriptivo que integra la Lexemática de Coseriu y Geckeler y la Gramática Funcional (GF) de Dik. Dicho modelo presenta una gran validez en el estudio del léxico—y, por tanto, resulta muy apropiado para nuestro propósito—porque, como afirma su propio autor (1990), consiste en la integración armónica de dos modelos de análisis complementarios: la Lexemática, modelo estructural esencialmente paradigmático de análisis léxico, y la GF, modelo gramatical basado en el lexicón, que proporciona un desarrollo altamente estructurado de las relaciones sintagmáticas de las unidades léxicas.

De la *Lexemática* tomamos los conceptos de estructura de *campo léxico*—o "paradigma léxico que se origina por la distribución de un continuo de contenido léxico en diferentes unidades, dadas en la lengua como palabras, que están recíprocamente en oposición inmediata mediante rasgos distintivos de contenido simples" (Coseriu, 1967: 294); articulación en *dimensiones* y

Miscelánea: A Journal of English and American Studies 18 (1997): 89-114.