

Is that a fact? A corpus study of the syntax and semantics of *the fact that*

Solveig Granath, Karlstad University

1. Introduction

Even though corpora have been used in language studies for the past few decades, up until now very few grammars of English have based their descriptions of the language on the wealth of information that can be gathered from them. There are indications that this is beginning to change. For example, Biber et al (1999) use a corpus-based approach in their handbook, the *Longman Grammar of Spoken and Written English*. It can be expected that this new approach will bring about both a revision and a refinement of descriptions of structure and patterns of usage in present-day English. The advantages of using corpora as the basis of grammatical descriptions should be obvious to anyone interested in grammar: not only is it possible to provide examples of actual use instead of invented sentences, but in addition, information on frequency of use, style, and the meaning of different features can be included.

The use of *the fact that* (i.e. *the fact* followed by an appositional *that*-clause) is a case in point. Learner grammars tend to stress that *the fact* is used to link a preposition to a *that*-clause, since prepositions may not govern *that*-clauses in English (Svartvik & Sager 1996:331, Hasselgård et al. 1998:349). Sometimes it is added that the use of *fact* presupposes that the *that*-clause expresses a fact. According to this description, *fact* can thus be regarded both as a function word and as a content word. As a reflection of this, it is interesting to note that the first two entries of *fact* in *Collins Cobuild English Dictionary* (1995) refer to the syntactic use rather than the meaning of the word. Thus the first entry says that “you use *the fact that* after some verbs or prepositions ... to link the verb or preposition with a clause”, and in the second entry the reader is informed that “you use *the fact that* instead of a simple *that*-clause either for emphasis or because the clause is the subject of your sentence”. Only in the fifth entry is a semantic definition of the word given: “when you refer to something as a *fact*... you mean that you think it is true or correct”. In Quirk et al (1985:1001-02) *the fact* is even referred to as a marginal subordinator, which stresses its syntactic function over and above its semantic meaning.

Whereas grammars and dictionaries thus demonstrate the need for *the fact that* in certain syntactic patterns, usage handbooks often denounce the use of *the fact* as “deadwood” (Perrin 1942: 515, 161): “*The fact that* is very often a circumlocution for which *that* alone would do as well: *He was quite conscious [of the fact] that his visitor had some other reason for coming.*” Even quite modern usage guides will warn users that *the fact that* “is often inserted quite superfluously into a sentence”, and exhort the user to avoid the phrase, because it is “ugly and pretentious” (Kahn and Ilson 1985:233). It is noteworthy that Quirk et al. (1985:659, 1002) also refer to prepositional phrases ending in *the fact that* as “stylistically clumsy”.

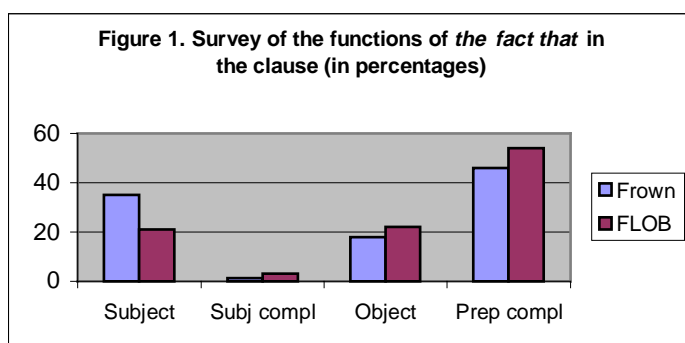
Both grammars and handbooks thus tend to regard *the fact* primarily as a function word, which fulfils little semantic function in the clause. This is not supported by Mair (1988), who draws two major conclusions from his corpus study of spoken and written English: first, “the function of *the fact* is primarily semantic: it indicates that the speaker takes the following clause to refer to a fact,” and second, “*the fact that* is not a mere variant of the conjunction *that* but a genuinely suppletive form which substitutes for *that* in contexts where the latter is ruled out” (Mair 1988:70). The present study is an attempt to shed light on some syntactic and semantic features of *the fact that*. The paper will begin with a presentation of the frequency with which *the fact that* appears in different functions in the clause. Next, a closer analysis will be made of the use of the phrase in its functions as subject, subject complement, prepositional complement, and object of a transitive verb. One feature that is rarely noted is that the phrase in itself is not syntactically frozen, and examples of this will be given in section 7 below. Finally, the question whether *fact* is semantically meaningful or just a function word empty of meaning will be discussed (section 8).

The results presented in the quantitative part (section 2) are based on two million-word corpora, FLOB and Frown, the Freiburg update of LOB and Brown with 1991 and 1992 used as sampling years. A pilot study of LOB and Brown, which contain texts that antedate FLOB and Frown by 30 years, showed results that were approximately the same. This makes it reasonable to assume that the figures are fairly representative, at least of these two varieties of English, but possibly also of English in general. The three sections that discuss *the fact that* in subject position, as subject complement, and as the complement of a preposition are likewise based on FLOB and Frown, whereas the three sections that deal with *the fact that* in object position and the syntax and semantics of the phrase have culled

examples not only from these two corpora, but also from a number of British and American newspaper and news broadcast corpora. The majority of examples are from written texts, but in the later sections, a few examples of spoken English are included.

2. Survey of the functions of *the fact that* in the clause

The two corpora on which the first part of this study is based, the Frown corpus of American English and FLOB which is a similar corpus of British texts, are rather small by comparison with some of the corpora in use today, but since they contain a variety of texts representing many different text categories, it can be assumed that they reflect general tendencies in the language fairly well. The total number of *the fact that* was somewhat higher in the British corpus (121 in FLOB as compared to 87 in Frown), but when it comes to syntactic function of *the fact that*, the results are similar (figure 1). Thus, in close to half of the sentences in both corpora, *the fact that* is the complement of a preposition, something that is reflected in grammatical descriptions of this feature, since we saw above that this is the context in which *the fact* is most frequently mentioned in grammars today.



The second most common function of *the fact that* is that of subject; however, almost equally frequent is *the fact* as the head of an object clause, something that is all but ignored in grammatical descriptions. Syntactic accounts generally classify *that*-clauses as nominal clauses, which means that their function is similar to NPs (e.g. Quirk et al. 1985:1048). It is therefore curious that a subset of the transitive verbs that occur with *the fact* + a *that*-clause will take a *that*-clause only as the complement of a head noun (see below, section 6). The least common grammatical function is that of subject complement. It should be kept in mind that the figures are based on written corpora, and that they might be somewhat different in a corpus of spoken English.

3. *The fact that* as subject

Clauses in subject position are often felt to be awkward in English, since they violate the principle of end-weight, which requires that “the part of the sentence following the verb should be as long as, and preferably longer than, the part that precedes the verb” (Quirk et al. 1985:1040). This is normally resolved by extraposition of the clause, which means that the clause is moved to the end of the sentence and *it* takes its place as the subject, so that (1b) is generally preferred to (1a):

- (1) a. That the cable networks and the advertisers that found them will continue to encourage the trend is far from certain.
 b. It is far from certain that the cable networks and the advertisers that found them will continue to encourage the trend.

If the *that*-clause is an appositional clause, however, extraposition is blocked, since *it* cannot be used in subject position when the extraposed constituent is an NP (2):

- (2) * It is a real advantage the fact that Harry doesn't have a cart of his own.

There are certain conditions under which a subject *the fact that* can move to clause-final position. This happens for instance when the predicate phrase consists of a clause-initial adjective (often in the comparative) plus the copula, as in (3a). The participle of a verb may also occur in clause-initial position, in which case auxiliary *be* follows the rest of the verb phrase (3b):

- (3) a. More worrying, I fear, is the fact that most people will have terrible trouble just understanding what Condren is trying to argue. (Frown J58 144)
 b. Complicating the issues of social class is the fact that in the United States there is a large overlap between lower-middle-class, working-class, and lower-class membership (Frown J49 175)

As can be seen from table 1, clause-final subjects are generally longer than clause-initial subjects, but it is still striking that when *the fact that* is in its normal subject position, it often “outweighs” the predicate part of the clause, as in (4), where the subject consists of 24 words and the verb phrase of only 3 (see also example (16a,b) below).

- (4) The fact that she is still somewhat tentative in the role and that her command of English is rather less secure than her arabesques, are minor blemishes. (FLOB C11 125)

Table 1. Survey of position and length of subject *the fact that* in Frown and FLOB

	Frown	FLOB
Clause-initial subject		
No. of sentences	19	21
Length	7 – 34 words	6 – 24 words
Average length	15 words	13 words
Clause-final subject		
No. of sentences	11	3
Length	17 – 41 words	14 – 35 words
Average length	31 words	21 words

It is often the case, however, that these clauses have long and information-loaded predicates in addition to the long subjects, as witness (5), which has a subject part consisting of 34 words, and a verb phrase containing as many as 37 words.

- (5) The fact that this is still going on, in both this and related semidetached systems (the term comes from the fact that only one of the stars is in contact with the critical surface), means that accretion flows onto the companion have played a role in the orbital dynamics and that this has fed back into the stellar evolution through the alteration of the mass and boundary conditions on the stars. (Frown J07 64)

The explanation for the reversal of weight in sentences such as (4) most likely has to do with information structure, so that even if the subject in a sentence like (4) contains several items of new information, they are presented as “facts”, i.e. as something given that cannot be questioned by the addressee. It is interesting to note that an initial subject *that*-clause will also be understood factively, i.e. presupposed to be true, even when it is not preceded by *the fact*. This is shown in (6), where only (6c) presupposes that Smith had in fact arrived (see Kiparsky & Kiparsky 1970:167f, Granath 1997:36f).

- (6) a. The UPI reported that Smith had arrived.
 b. It was reported by the UPI that Smith had arrived.
 c. That Smith had arrived was reported by the UPI.

In conclusion, when *the fact that* functions as the subject of the clause, the principle of end-weight no longer applies, so that we may find long and heavy elements with a high information load at the beginning of a sentence, contrary to the normal place of such elements in English. This is confirmed by the results quoted in Mair (1988:65), who found that there is a tendency to have *the fact that* rather than only *that* in English subject clauses that have not undergone extraposition. It seems that *the fact* may serve as a structural marker of a long constituent, and that it will signal to the listener/reader to expect the verb phrase to appear late in the clause. As such, it serves a communicative purpose, and omission of it would enhance the complexity of the sentence and make it more difficult to comprehend, as is shown in (7a) and (7b).

- (7) a. **That** patients still have to wait 21 weeks to get into hospital for treatment and 12 weeks for day surgery demonstrates yet again that the health service has not got enough money to treat patients in a reasonable time.
 b. **The fact that** patients still have to wait 21 weeks to get into hospital for treatment and 12 weeks for day surgery demonstrates yet again that the health service has not got enough money to treat patients in a reasonable time." (FLOB A14 192)

4. *The fact that* as subject complement

The least common function of *the fact that* is that of subject complement. Only one instance was found in Frown, compared to four in FLOB. Not only is *the fact that* rather rare in this function, it also seems that semantically, *the fact* serves only to reinforce the statement somewhat, and that the message is communicated almost as forcefully if it is omitted. The reader is welcome to test this on the following two examples from FLOB (my slashes).

- (8) a. “the most obvious encumbrance on this picture is /**the fact/ that** it is woefully late” (FLOB G49 140)
b. Another component of her success was /**the fact/ that** she rated well under both the RORC and American CCA rules of the time. (FLOB E18 23)

Considering the low number of sentences in which *the fact that* was used in this function, it seems likely that this construction is less frequent than *that*-clauses (without an introductory *fact*) as subject complement.

5. *The fact that* as the complement of a preposition

Standard grammatical descriptions of English stress that prepositions cannot govern *that*-clauses in English (see e.g. Svartvik & Sager 1996:331, Hasselgård et al. 1998:349, Quirk et al. 1985:658f).¹ This means that some other syntactic device is needed to link clauses to prepositional phrases. One such device is “to use an appositive construction with a ‘general’ noun such as *fact*” (Quirk et al. 1985:659). Both in grammars and in English usage guides, this usage is referred to as “clumsy” and “superfluous”, and the language user is told to avoid it if possible (for examples of how this can be done, see (9) and (11a) below). The advice from handbooks is usually simply to leave out the sequence preposition + *fact*, or in the case of the two prepositions *despite* and *due to*, to replace them with a subordinator such as *although* or *because*. However, evidence from the corpora shows that this is rarely an option for the user: in most cases where this structure is used, no other alternative exists.

In Frown, 40 out of 87 sentences with *the fact that* were of the structure preposition + *the fact*. The number was even higher in FLOB, so that this structure was found in 65 out of 121 sentences. This means that this is by far the most common use of *the fact that*. In each corpus, there were only a handful of cases where the writer would have had the choice of leaving out the preposition + *the fact*. This was for instance the case in the four sentences where the preposition was the complement of an adjective (*content with*, *grateful for*, *oblivious of/to*, *responsible for*) and it was possible with two of the four nouns postmodified by preposition + *the fact that* in Frown and FLOB: *recognition of*, *concern about*. However, only in two cases would it have been possible to leave out preposition + *fact* as the complement of a verb, namely *comfort oneself with* and *hint at*; in all other cases an ungrammatical construction would have been the result.² Some examples of sentences where omission of the prephrase would have been possible are given in (9) (my slashes).

- (9) a. I am grateful /for **the fact/ that** we know so little about his life. (FLOB G08 202)
b. In reality, its merger with another organization was recognition /of **the fact/ that** it carried “high negatives” in public opinion polls. (Frown D09 202)
c. In the closing pages of the text, his attempted self-renunciation hints /at **the fact/ that** to abolish the reflective ego is to cease to be a human subject (FLOB J62 166)

In about one fifth of the sentences, *the fact* was the complement of a preposition with no governing adjective, noun or verb. The most common among these were passive sentences, where *the fact that* took on the function of agent (16 out of a total of 105 sentences). Mair (1988:64f) interprets this as an application of the principle of end-weight, i.e. that long constituents are moved to the end of the clause. Again, it does indeed seem as if the language user sometimes has a choice, so that transforming (10a) into an active clause (10a') does not really produce a more complex structure. In other cases, the principle of given and new in the functional sentence perspective makes the passive stylistically the only viable alternative (cf. 10b and 10b').

¹ This rule does admit a number of different exceptions – indeed far more than might be expected – so that among the prepositions discussed in this section, *except*, *given* and *than* can indeed take *that*-clause complements. Syntacticians normally get around this problem by reclassifying these words, so that *except that* is regarded as a complex subordinator, *besides* and *than* are given a double classification (as a preposition and as a subordinator), and *given* is referred to as a marginal preposition (e.g. Quirk et al. 1985:667) precisely because it may indeed precede a *that*-clause. Further information about *that*-clauses as complements of prepositions can be found in Christophersen (1979) and Seppänen & Granath (forthcoming).

² See Granath 1997 for a detailed account of the omission of prepositions in verb complementation.

- (10) a. The value of the posters must have been enhanced **by the fact that** they were created not by an industry but by an individual. (FLOB E35 213)
 a'. The fact that they were created not by an industry but by an individual must have enhanced the value of the posters.
 b. This vulnerability is underlain **by the fact that** in 1988 66 per cent of goods and services were exported. (FLOB J42 33)
 b'. The fact that in 1988 66 per cent of goods and services were exported underlies this vulnerability.

Other prepositions found with *the fact that* in the two corpora (in order of frequency) were *despite*, *due to*, *given*, *than*, *aside from*, *beside(s)*, and *except for*. Of these, *besides* (but not *beside*), *except* and *given* (and more marginally *than*) will allow a following *that*-clause with no mediating noun (11a). With the majority of prepositions complementing verbs and phrases of various kinds, this possibility does not exist (11b,c).

- (11) a. In the year 600 the civilizations of Spain and Mexico were roughly comparable **except /for the fact/ that** the former had profited from the invention of the wheel. (Frown N13 55)
 b. ... much of this stemmed **from the fact that** he rarely took on the more general questions of the organization (Frown J60 116)
 c. I didn't make anything **of the fact that** we didn't have a particularly good time in bed. (FLOB K21 283)

In summary, out of the 105 sentences where *the fact that* was the complement of a preposition, there were only 15 altogether where it would have been possible to leave out preposition + *fact*. Another consideration is of course what will happen to the meaning of the sentence if this is done. Removing *the fact that* will certainly take away some of the communicative force of the sentence, which can be seen e.g. in the examples in (9) and (11a).

6. *The fact that* as direct object

It has already been pointed out above (section 2) that *that*-clauses are traditionally classified as nominal clauses and thus put on a par with noun phrases syntactically. We would therefore expect that when *the fact that* serves as the object of a transitive verb, a *that*-clause on its own would do, even if it would be less emphatic than a construction with *the fact*. However, when one begins to consider corpus evidence, it is clear that this is not at all the case. There are several issues that deserve description here. First, even a very brief search of one of the big corpora will turn up quite a large number of the structure verb + *the fact that*, which means that this is a common structure in English. Furthermore, it is one that is generally ignored in grammars. Second, whereas there are a number of verbs where *the fact* can be said to be optional syntactically, there are others with which a *that*-clause complement is rare, and yet a further group where this possibility does not seem to exist at all. It is possible that this area of English syntax has been neglected precisely because it has been impossible to investigate the patterns in which these verbs occur without access to large corpora.³

For this part of the paper, several CD-ROMs of British and American newspapers and broadcast news were used. A pilot study of some of these quickly turned up more than 250 verbs followed by *the fact that*. In order to delimit the study somewhat, the results in this section are primarily based on the 120 verbs of this type found in *The Guardian/The Observer* 1999. Altogether, this corpus comprises approximately 40 million words. Other corpora were consulted to determine whether *the fact* could be considered to be optional or obligatory. It is possible that further corpus searches will revise the results somewhat, so that some of the verbs in group 3 (verbs that obligatorily require an NP head of the (appositional) clause) will have to be moved to group 2 (verbs where a *that*-clause complement is an alternative to *the fact that*).

The first group consists of verbs that occasionally take *the fact* as an object head of a *that*-clause, but where the predominant structure is for the clause to follow the verb directly. This applies to roughly one fourth of the 120 verbs in the survey (table 2). One interesting aspect is that when the *Collins Cobuild Dictionary* entries for the 120 verbs were checked, the 28 verbs in this group were all

³ The corpus that Mair (1988) used for his paper comprised about 840,000 words, which was a large corpus at the time, but it is small by today's standards. It is noteworthy that of the nine verbs for which Mair claims that the plain *that*-clause is rejected (*give*, *face*, *ignore*, *bring home*, *disguise*, *resent*, *obscure*, *raise*, *discuss*) (Mair 1988:65), only two have been found not to take a plain *that*-clause in the present study, namely *obscure* and *raise*. This underscores the value of large corpora, since native speaker intuition cannot be trusted. Some speakers may very well reject what other speakers say, and it is impossible even for trained linguists to be able to imagine all the contexts in which a word may be used.

categorised as taking a “report-clause”, i.e. a *that*-clause, whereas the remaining 92 verbs were only marked as taking NP objects.

Table 2. Transitive verbs for which *that*-clause complementation is the normal pattern and *the fact that* less often used

accept	consider	forget	mind	recognise	stress
acknowledge	dispute	grasp	note	regret	take (into account)
add	emphasise	illustrate	notice	report	volunteer
appreciate	establish	lament	proclaim	reveal	
believe	explain	mention	publicise	spot	

In the second group of verbs, two types can be distinguished: for a few verbs *the fact that* occurs with almost the same frequency as a plain *that*-clause, but for the majority of verbs, *the fact that* is in fact the predominant pattern, with the plain *that*-clause a rare alternative. The reason why these two types have been collapsed into one group is that even the relatively large corpora used as the basis of this survey are too small to establish with any certainty how these verbs should be divided up. Thus *reflect* (in the sense ‘mirror’) occurs 46 times with *the fact that* and 6 times with a plain *that*-clause (12a, b); *face* occurs once with a *that*-clause, 13 times with *the fact that*, but in addition to that, it occurs with a number of other NP heads, such as *accusations*, *allegations*, *claims*, *complaints*, *criticism*, *the possibility*, *the realisation* (12c, d). Other verbs, such as *like*, *love*, and *hate* were extremely rare with *that*-clause complements, and the examples that were found of this structure stem from spoken corpora.

- (12)
- The book looks old-fashioned. ... Perhaps this **reflects that** the hey-day of the circus in Britain is long gone. (*The Guardian*, 17 April 1999, p. 9)
 - The Bath course records high scores in the value-added measure of academic teaching. This **reflects the fact that** a high proportion of students with low entry qualifications are graduating with upper seconds and first-class honours. (*The Guardian*, 9 Nov 1999, p. 10)
 - ‘I think it is a damning indictment of our society that my daughter is now having **to face that** you cannot trust any word spoken or written by a politician, let alone a secretary of state or prime minister. (*The Guardian*, 21 Aug 1999, p. 2)
 - It is scary enough to **face the fact that** you yourself are faking it. (*The Guardian*, 30 Jun 1999, p. 8)

Table 3 lists the verbs that were found to take predominantly *the fact that* rather than the plain *that*-clause. This group comprises close to half the 120 verbs, or 51 in all. It needs to be stressed that for some reason, the possibility of using a *that*-clause complement is not indicated in *Collins Cobuild English Dictionary*, even though this is one of the first dictionaries to be based on a large corpus, namely the enormous *Bank of English*.

Table 3. Transitive verbs with which *the fact that* predominates but which occasionally take a *that*-clause complement

absorb	celebrate	disguise	can't help	neglect	rumble
advertise	change	disregard	hide	omit	sell
alter	cite	dread	highlight	overlook	can't stand
avoid	conceal	enjoy	ignore	query	take (into account)
belie	cover up	escape	include	reflect	uncover
bemoan	credit	expose	like	reinforce	underline
bring home, into	criticise	face	love	relish	underscore
focus, out (into	deplore	handle	mark	resent	welcome
the open)	discuss	hate	mask	respect	

The group of verbs that poses the greatest problems to explain in terms of syntactic behaviour is the third group, presented in table 4. In the material consulted for this study, these verbs were found to occur only with a *that*-clause as the apposition of a head noun. *The fact* was obviously found with all these verbs, but for some of them, a number of other head nouns were also found, so that *dismiss*, for instance, was recorded with well over 20 different NPs, such as *speculations*, *accusations*, *hints*, *rumours* etc. This group consists of 41 verbs, that is approximately one third of the verbs studied, but it should be remembered that further research will probably find that at least some of them will indeed also take plain *that*-clause complements.

Table 4. Transitive verbs which only take a *that*-clause complement with a mediating NP such as *the fact*

address	cloud	dodge	invent	obscure	promote
answer (inanimate subject)	concern	drop	italicise	outweigh	raise
begrudge	confront	duck	keep secret	overcome	rue
blame	decry	exploit	laud	overshadow	share
camouflage	defend	forgive	miss	pardon	splash
cloak	dislike	hold sth against	negate	pinpoint	square
	dismiss	illuminate	obfuscate	plug	withhold

It is difficult to account for the reasons why these verbs should behave differently from the verbs in table 2. One possibility is that semantics has something to do with it. Language often works by analogy, so that words that have the same meaning tend to adopt similar syntactic patterns. It is apparent that many of the words are semantically close in meaning, so that we find groups of verbs that are converse terms, such as conceal/disclose, praise/condemn, welcome/regret, admit/deny, like/dislike, support/question etc. If meaning could be used to explain syntactic structure, we would expect that verbs with a similar meaning should behave identically syntactically, but this is not the case. Thus, in group 2 (verbs that take both *the fact that* and a plain *that*-clause) we find *conceal*, *hide*, and *disguise*, whereas in group 3 (verbs that do not take a *that*-clause complement) we find *camouflage*, *cloak*, *cloud* and *obscure*. One possibility that this suggests is that frequency matters, so that less frequent verbs more rarely take plain *that*-clauses. The scope of this paper has not made it possible to look more carefully into this, but even a quick look at table 4 indicates that at least some of the verbs are not at all uncommon in English, for instance *blame*, *concern* and *defend* (example 13). The reader is welcome to try to verify whether it is possible to omit *the fact* in these three sentences.

- (13)
- a. ‘... She shouldn’t **blame the fact that** she’s a woman, or me,’ he gabbles, with more detail than is strictly necessary. (*The Guardian* 31 jul 1998, p 2)
 - b. Mr Clinton’s most serious legal problems **concerned the fact that** he and Ms Lewinsky gave conflicting stories about how and why she gave back the presents he had given her (*The Guardian* 21 Aug 1998, p 13)
 - c. But Bassay asserts that classroom application should not be the overwhelming criterion for research and **defends the fact that** academics make so much use of it. (*The Guardian*, 17 Mar 1998, p 6)

The present paper can only begin to outline the problems encountered in the analysis of these verbs. The problem appears to call for a syntactic rather than a semantic explanation, since a number of the verbs in group 3 will take a plain *that*-clause complement in other languages, such as Swedish. The fact that the structures are in some sense idiosyncratic in English is another reason why it is important to have a fuller account of these verbs than has so far been the case: there is at present no place where learners can find out about the erratic behaviour of these verbs, since not even dictionaries account fully for their colligations.

What we are left with at present is a system that in Matthews’ terms is not codified but rather an area of partial codification, for which there are no explicit rules (Matthews 1981:20-21). To a syntactician, this is of course not the ideal situation, nor is it a happy situation for the language learner. But, as corpus linguists, working with actual language use and not the ideal speaker-hearer, we will probably have to accept that there are areas of a language that cannot be wholly explained in terms of one system or another. Still, the topic of this section, *the fact that* as the complement of transitive verbs, is an area where more research is needed both to provide a fuller account of this usage for learners, and in order to see if in the end perhaps it is possible to find at least some determining factors that can explain what to us looks like largely idiosyncratic behaviour.

7. Internal syntactic variation in the phrase *the fact that*

One aspect of language that has become more obvious to linguists after the introduction of corpus-based research is that speakers only partly have an open choice when it comes to selecting vocabulary, syntactic forms etc. Much of what we say actually consists of pre-fabricated or semi-preconstructed chunks of words, which work together in phrases. This is usually referred to as the idiom principle (Sinclair:1991 110-115). A typical feature of idioms is their frozenness, i.e. lack of syntactic variation in number, tense, voice etc. It seems that *the fact that* would serve well as an example of this, since all mention of the phrase includes the definite article, the subordinator *that*, and *fact* in the singular. However, even though this is indeed the predominant structure of the phrase, some examples will demonstrate that it is in fact not invariable.

Although the definite article is by far the most common determiner, *this* and *that* occur occasionally, both when the *that*-clause is restrictive, as in (14a and b), and when it is non-restrictive, as in (14c):

- (14)
- a. We haven't tried to hide **this fact that**, come November, we would have banned them ourselves. (*The Observer*, 19 Sept. 1999, p. 1)
 - b. He kept this up in meetings and letters, despite **that fact that** we were never together in Paris. (*The Guardian*, 21 July 1998, p. 16)
 - c. Despite **this fact, that** language is so much more than just a huge store of isolated words, what you remember of learning your own native language is likely to be limited to memories of just that – the learning of new words. (Heny 1998:190)

Examples where the indefinite article is used are rare. Instead we find *a fact that* in the phrase *to know for a fact that*, where the *that*-clause must be interpreted as the object of the verb *know*. An extraposed *that*-clause also often follows *a fact* in clauses of the type *It is a fact that (...)*. In neither of these cases is the *that*-clause appositional. What we find instead is that *a fact* is often followed by a relative *that*-clause. The only example in the present investigation where the *that*-clause could be taken to be an appositional clause with the head noun *a fact* was actually one where *fact* was post-modified by a relative clause:

- (15)
- The clash between Schroder and Blair also underscored **a fact that** yesterday's Florence summit of the centre left had been intended to disguise: **that** there remains a huge gulf in thinking between Britain's post-Thatcherite Labour party and America's post-Reaganite Democrats on the one hand and the socialist and social democrat parties on the continent on the other. (*The Guardian*, 23 Nov. 1999, p. 2)

An alternative analysis of the clause structure is to regard the *that*-clause as the direct object of the verb *underscore*, which, as was shown above (table 3), belongs to the group of verbs that may take a *that*-clause with no mediating noun. From a stylistic point of view it also seems logical that the determiner of *fact* is definite, since in the majority of cases it refers to something presupposed and thus given information.

As regards the number of *fact*, the singular by far outnumbers the plural, for the simple reason that in general, only one "fact" is mentioned. Occasionally, we find the plural *facts* with appositive postmodification, something that is also noted by Quirk et al. (1985: 1261). This obviously occurs only when there is more than one appositive clause. However, with two or more appositive *that*-clauses, it is possible to use either the singular *fact* or the plural *facts*, as witness the examples in (16), where (a) and (b) contain *that*-clauses in apposition to the subject, and the ones in (c) and (d) belong to the object of the clause:

- (16)
- a. **The fact that** England did not even practice penalties, **and that** Hoddle characteristically refused to accept this as an error, left the coach looking conceited and complacent (*The Observer*, 4 Oct 1998, p. 5)
 - b. **The facts that** the roads carried 81 per cent of all freight last year **and that** 86 per cent of passenger miles are taken in cars and vans cut no ice. (*The Observer*, 21 Nov. 1999, p. 28)
 - c. He **ignores the fact that** Geffen reputedly offered her \$1million for it, **and that** her previous album *Pretty On The Inside* was a powerful musical landmark two years previously. (*The Observer*, 5 Jul 1998, p. 7)
 - d. But such notions **ignore the facts that** Ramsay has already made three short films **and that** two of them have won prizes at Cannes. (*The Guardian*, 14 Aug. 1999, p. 4)

Finally, the language user also has the option of omitting the subordinator *that* after *fact*. One could assume that *that* would be left out more often in informal than in formal contexts, but corpus data does not confirm this (cf. (17a and b), both of which contain the phrase *lament the fact*, where *that* is omitted in the longer example from the business section (17b)).

- (17)
- a. He **laments the fact that** you can't trust politicians these days. (*The Guardian*, 1 Nov 1999, p. 12)
 - b. Fellow Tory Eric Forth, a former minister, **lamented the fact** the Opposition was 'conniving with his government' to pass the legislation (*The Guardian*, 26 Oct 1999, p. 29)

A cursory glance at the sentences where *that* is omitted after *fact* indicates that this happens more frequently when the embedded clause has a pronoun subject, but it is not restricted to this type of subject, as is shown in (18). Neither does there seem to be a limitation on *that*-omission due to the function of *fact* in the clause, so that we find it both when *the fact* is a subject (18a), an object (18b), and a prepositional complement (18c).

- (18)
- a. **The fact** the President must step down in 2000 ... does not matter. (*The Observer*, 16 May 1999, p. 22)
 - b. These examples of Germany's innovative and corporate success cannot, however, disguise **the fact** the economy inherited by Gerhard Schroder from Helmut Kohl has serious structural problems (*The Guardian*, 10 June 1999, p. 19)

c. After he was discharged the police asked him 'to keep quiet about **the fact** a tank had crushed students ... (The Guardian, 2 June 1999, p. 12)

In conclusion, corpus evidence shows that *the fact that* is not the invariable phrase that it appears to be in grammatical descriptions and according to information given in dictionaries.

8. The semantics of *fact*

The seminal article *FACT* by Kiparsky & Kiparsky (1971) presented the idea that certain predicates are factive, meaning that their complements are presupposed to be true. One criterion used for testing factivity was to negate the matrix verb. The predicate was classified as factive if the presupposition remained intact under negation unless explicitly contradicted (Kiparsky & Kiparsky 1971:351-52, Leech 1974:301-317). Thus *I'm sorry that he lost his job* and *I'm not sorry that he lost his job* both presuppose that he lost his job. In a sense then, such predicates can be said to be open to objective verification, that is, one should be able to show that they are facts. According to Kiparsky & Kiparsky (1971:145), "[o]nly factive predicates can have as their objects the noun *fact* with a gerund or *that*-clause." Typical examples of factive predicates are *regret*, *grasp*, *take into account*, *ignore*, *mind*, and *deplore*. Predicates such as these make up the majority of the verbs in section 6 above, at least in the two groups that predominantly occur with *the fact that*. We also saw that *that*-clauses in subject position tend to be given a factive reading whether or not *the fact* is overtly stated. Therefore, one major conclusion to be drawn from the present investigation is that *the fact* is indeed used because the appositional clause is presupposed to be true.

More interesting, perhaps, is whether *fact* is losing some of its meaning, and whether it is sometimes used for syntactic reasons alone. Even Kiparsky & Kiparsky (1971:147fn) refer to a colleague who had informed them "that for him factive and non-factive predicates behave in most respects alike and that even the word *fact* in his speech has lost its literal meaning and can head clauses for which no presupposition of truth is made." Such usage of *fact* is what language guide books react against, so that a sentence such as *I certainly do not accept the fact that Sir Patrick was remotely influenced by the timing of the leadership election* (Radio 4) receives the comment that "[i]f someone does not accept a fact, then in their eyes it is not a fact" (Blamires 1998:117). The author goes on to suggest that *suggestion* should replace *fact* in this case. The way *fact* is described in grammars, e.g. by being referred to as a marginal subordinator (Quirk et al. 1985:1001-02) also indicates that the word has lost much of its original meaning, and that today, it can be used as a function word devoid of meaning. What indications are there in the corpora that this is true?

Starting from Kiparsky & Kiparsky's list of non-factive verbs, we find that *believe* is a typical example of such a predicate. According to their rule, then, it should never occur with *the fact* as an object. Nevertheless, several examples of this construction were found (19a,b). Included here is also an example with *buy* in the sense 'believe, accept' (19c).

- (19)
- a. Her mother, Amanda, 24, said: 'It's been very traumatic for her and I am just glad she's OK. But I just can't **believe the fact that** she was able to get hold of something like this in the first place. (The Guardian, 15 Dec 1999, p. 6)
 - b. They paid a big price for that, though, and I think they paid it in the sense that Heidstra – he talks about a Bronco-like car leaving the scene of this crime, and if Christopher Darden and the prosecution can make the jury question the timing, but **believe the fact that** there's a car leaving, then his testimony has the effect of suggesting Simpson at the scene. (CNN Morning News, 26 Jul 1995)
 - c. Quite frankly, it is not a tax break for the rich. ... It would be first dollar coverage, it would be a high deductible, it would be very, very affordable for those people and unfortunately I just cannot **buy the fact that** it is a tax break for the rich. (CNN Domestic News, 25 Apr 1996)

If we take a closer look at these examples, it is evident that (19a) expresses a fact that can be verified. The mother's use of *can't believe the fact* is just her way of signalling incredulity. The second example, (19b), might be a little more difficult to explain in terms of factuality. If you are going to make someone believe something, this usually signifies that what you make them believe is not true. However, the context here is the courtroom, and the fact referred to is an event that is a fact for the prosecution, whose task it is to make the jury believe the facts they present.⁴ Neither of these examples

⁴ Greetham (1999:3) outlines three "degrees of truth" from the rhetoric of litigation: *raw facts*, the common ground taken as evidence by both parties, *true facts*, the facts appropriated by only one party, which "had acquired a new level of revealed truth by being frankly partisan, and thus more true than mere facts", and *factoids*, presented by the opposing party but not espoused as facts by one's own side.

can therefore be explained as sentences where *the fact* functions as a prop word devoid of meaning. The last example, on the other hand, is a contradiction, because the speaker states emphatically that “it is not a tax break for the rich”, a “fact” that he immediately says he does not believe in. In this particular case *the fact* is probably more of a rhetorical flourish, used to add emphasis to what the person is saying.

It ought not to be possible to refer to things that are counterfactual (i.e. made up or imagined) as “facts”, but (20) demonstrates precisely this. In both these cases, *fact* fulfils a syntactic function, since neither *lay money on* nor *invent* can take *that*-clause complements.⁵ Thus it does seem that we have two cases here of *the fact that* as a marginal subordinator.

- (20) a. I'll **lay money on the fact that** he's found some other way to work out his aggressions. (Walters, Minette 1999 *The Dark Room* Pan Books, p. 447)
 b. When asked whether he still works behind the bar, Martin says: 'I used to love it, but I don't do it any more. Tim **invented the fact that** he worked behind the bar, but he never did. (*The Observer*, 19 Sep 1999, p. 4)

Two other verbs often mentioned as typical examples of non-factive verbs are *claim* and *assume*. *Claim* is a verb that signals that the complement is the speaker's subjective opinion, and *assume* refers to a hypothetical condition. Consequently, it is not possible to verify the truth of the complement of either of these verbs (21a,b). In (21c), where the predicate is *hear*, it is obvious that the speaker uses *the fact* to refer to something he does not believe in.

- (21) a. Bobby Murcer, Mantle Teammate: Well, I didn't know Gehrig and Ruth – I just heard about those guys. I knew Mickey Mantle, so I can honestly **claim the fact that** I think Mickey Mantle was the greatest. (*CNN World News*, 15 Aug 1995)
 b. Eli Noam: It's true for the present, but if you look at the market share of broadcasters generally, and even **assuming the fact that** they will be able with new technology to squeeze more broadcast channels into their broadcast signals, **even assuming that**, they still are – will be big players, but among many, many other competitors. (*NPR Morning: Business*, 9 Aug 1995)
 c. At Marty's Center Tap, UAW members are not optimistic about reaching a contract agreement and Mark Samp says it will take a tremendous effort to overcome the effects of a strike, which has been terrible for everyone.
 Mark Samp: I get tired of **hearing the fact that** people believe that, well, you know, time heals all wounds. (*NPR Morning Edition: Domestic News*, 9 Jan 1996)

It is significant that all these examples are from spoken language, and the speakers have thus not been able to edit their utterances, which means that they could all be explained as performance errors. However, if we assume that the speakers used *the fact* for some purpose, then it is noteworthy that in none of the sentences in (21), *the fact* is needed for syntactic reasons: *assume*, *claim* and *hear* all normally take *that*-clause complements. So why is *the fact* used here at all? I would like to suggest that in the context, it has a certain pragmatic force, so that in the examples in both (20) and (21) *the fact* is used to strengthen the force of what is said. The fact that examples such as these are found in corpora also supports the suggestion made by Kryk (1981) that the logical analysis in terms of truth conditions of these utterances should be replaced by a pragmatic approach, and the factive/non-factive dichotomy by a scale of factive and “not-so-factive” predicates. Whether *the fact* is necessary to avoid an ungrammatical construction, as in (20), or just added as some kind of rhetorical flourish, as in (21), speakers make use of it in order to get a message across, a message that would certainly be less forceful without *the fact*.

9. Conclusion

The present paper has attempted to demonstrate that the use of *the fact that* in present-day English cannot be explained by referring to its syntactic function alone, but semantic as well as pragmatic meaning needs to be taken into account in a description of its function in utterances. This serves as an example to demonstrate in what way a corpus-based approach to syntax can be used to enhance existing accounts. Traditional syntax can be amended, so that a narrow focus on structure can be replaced by a multifunctional approach. Instead of classifying words as function words or content words, a common procedure in introductory texts in semantics, we will have to acknowledge that very few words act as function words only. The result is a grammar where there is a synthesis of syntactic, semantic and pragmatic rules.

References

- Biber D, Johansson, S, Leech, G, Conrad, S, Finegan, E 1999 *Longman Grammar of Spoken and Written English*. Harlow, Pearson Education.
- Blamires H 1998 *The Cassell Guide to Common Errors in English*. London, BCA.
- Christophersen P 1979 Prepositions Before Noun Clauses in Present-Day English. In Chesnutt, M, Faerch C, Thrane T, Caie G D, *Essays Presented to Knud Schibbye*. Copenhagen: Akademisk Forlag, 229-234.
- Granath S, 1997 *Verb Complementation in English: Omission of Prepositions before that-clauses and to-infinitives*. Göteborg, Acta T. Universitatis Gothoburgensis.
- Greetham D 1999 Facts, Truefacts, Factoids; or, Why Are They Still Saying Those Nasty Things about Epistemology? *Yearbook of English Studies* 29:1-23.
- Hasselgård H, Johansson S, Lysvåg P 1998 *English Grammar: Theory and Use*. Oslo, Universitetsforlaget.
- Heny F 1998 The Structure of Sentences. In Clark V P, Eschholz P A & Rosa A F (eds) *Language: Readings in Language and Culture*. Boston and New York, Bedford/St Martin's, 189-224.
- Kahn J E, Ilson R (eds) 1985 *The Right Word at the Right Time: A Guide to the English Language and How to Use It*. London, Reader's Digest Association.
- Kiparsky P, Kiparsky C 1971 Fact. In Bierwisch, M & Heidolph K E (eds), *Progress in Linguistics: A Collection of Papers*. The Hague, Mouton, 143-173.
- Kryk B 1982 The Relation between Predicates and Their Sentential Complements, A Pragmatic approach to English and Polish *Studia Anblica Posnaniensia: An International Review of English Studies*, 14:103-120.
- Leech G 1974 *Semantics*. Penguin Books.
- Mair C 1988 In Defense of *the fact that*: A Corpus-Based Study of Current British Usage. *Journal of English Linguistics* 21:1, 59-71.
- Matthews P H 1981 *Syntax*. Cambridge, Cambridge University Press.
- Perrin P G 1942 *Writer's Guide and Index to English*. Chicago, Scott, Foresman and Company.
- Seppänen A, Granath S (forthcoming) *That*-Clauses and the Complementation of Prepositions.
- Sinclair J 1991 *Corpus, Concordance, Collocation*. Oxford, Oxford University Press.
- Quirk R, Greenbaum S, Leech G, Svartvik J 1985 *A Comprehensive Grammar of the English Language*. London and New York, Longman.