

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)

# I *Introduction*

---

I have had two complementary aims in view in preparing the present book: to give a selective grammatical description of a corpus of some 135,000 words of written scientific English and to investigate certain areas of the grammar of 'common-core' English – the grammar that is common to all varieties of the language (except possibly a few highly restricted ones). The aims are complementary in that it is obviously impossible to give a non-trivial description that will account for the sentences in a given corpus while excluding all sentences of the language that do not occur therein; a text cannot be analyzed in isolation, and the description will be of interest only to the extent that the grammatical categories established have validity beyond the text. On the other hand texts constitute the linguist's primary data and it is salutary to test one's descriptions by confronting them with a sizeable body of such primary data. This is of course not to deny the value of studies based on data derived from introspection: we need to make use of both kinds of data.

The theoretical framework underlying the description is mainly that of transformational grammar. I therefore assume that the syntactic description of a sentence takes the form of a series of phrase-markers, or labelled bracketings, which represent its structure at different levels: the bracketing represents the constituent hierarchy at the given level, and the labelling represents the classification of the constituents. The first in the series of phrase-markers is said to represent the 'deep structure' of the sentence, the last the 'surface structure'; the first is generated by phrase-structure rules, whereas each of the remaining phrase-markers derives from the immediately preceding one in the series by a transformational rule.<sup>1</sup> I should emphasize, however, that

<sup>1</sup> Transformational grammar is too well known for it to be necessary to give a summary of it here. The fullest account is in Chomsky (1965); it also figures prominently in several recent textbooks of linguistics, e.g. Lyons (1968), Langacker (1967) and Langendoen (1969). Because of my informal approach, and because the book is intended as a contribution to the grammatical description of English, not to general linguistic theory, I have not felt it necessary to commit myself on such controversial issues within transformational theory as that of the generative or interpretative role of semantics (cf. G. Lakoff, forthcoming, and the references cited there).

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)

## 2 Introduction

I have not attempted to produce a formalized generative grammar: this may be regarded as the long-term goal, but in working towards it we can fruitfully discuss in informal terms what we want the formal grammar to say about the structure of particular sentences or classes of sentences.

The choice of scientific English as the variety for textual study was motivated primarily by practical considerations (though there is also the point that this variety of written English is relatively neglected in the major standard handbooks of Jespersen, Curme, etc.): data on scientific English is likely to be of interest to workers in two of the main branches of applied linguistics, automatic language data processing<sup>1</sup> and language teaching. There is a considerable demand for courses designed to teach this type of English to foreign learners, and although this book itself is in no way intended as a teaching grammar I hope it will be of use to those applied linguists who are concerned with preparing such courses.

I should make it clear that the present work is an exercise in 'descriptive linguistics', not 'stylistics', as these terms are used and contrasted, for example, by Crystal & Davy (1969): a stylistic analysis, in their sense, of written scientific English would seek to isolate those linguistic features which distinguish this from other varieties of English. The difference between my approach and theirs thus has to do with aims and priorities: working in stylistics, they regard the grammatical description as a tool for the differentiation of varieties, for the identification of linguistic features 'which are restricted to certain kinds of social context', and they emphasize that their grammar should be evaluated relative to this aim rather than in terms of its adequacy as a linguistic description – my primary concern, on the other hand, has been to produce as adequate a description as I can, judged simply as a (partial) grammar for its own sake. Because of this emphasis on pure description rather than stylistics I have not attempted to compare the corpus with texts from other varieties of English – the comparison in 5.6 between the relative clauses in my corpus and those in a sample of spoken (non-scientific) conversation is one exception, made possible by the availability of readily comparable data on the latter and included as an illustration of what might be done in a larger-scale study. Until further

<sup>1</sup> Cf. the remark of Clarke & Wall (1965: 312) in their 'An economical program for limited parsing of English': 'Perhaps one could hope to select instead the "syntactically most probable" parsing if adequate statistical studies of English grammar were available.'

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)

comparative work of this kind is done one cannot of course tell how far the statistical properties of the corpus reported in the present work are peculiarly characteristic of written scientific English and how far they are generalizable to other varieties; I hope, however, to have provided a solid basis for such comparative study.

As I said at the outset, the description is, naturally, partial and selective. Most of what I have to say relates to the grammar of the clause.<sup>1</sup> The method of presentation I have adopted for the most part is to discuss a given area of the grammar first in general or common-core terms, and then to examine the corpus in the light of the descriptive framework so established. I have not, however, felt it necessary to keep constant from chapter to chapter the relative weight given to common-core and corpus description; thus in the chapter on mood, the textual description occupies quite little space relative to the general discussion, whereas the proportions are reversed in the chapter on the uses of the modal auxiliaries—I have tended to devote relatively more attention to the corpus in those areas of the grammar which are less well understood or where current descriptions are less explicit, such as transitivity (ch. 3) as opposed to complementation (ch. 4).

Full details of the corpus are given in the Appendix. It is made up of 27 texts of 5,000 words each.<sup>2</sup> The texts were taken from three different 'strata', corresponding to different 'levels of brow': the nine 'high' stratum texts come from specialist journals, the nine 'mid' stratum ones from undergraduate textbooks, and the nine 'low' stratum ones from more popular works addressed to the intelligent and well-informed layman. The high and mid texts were also classified according to field or subject matter, the three categories selected being biology, chemistry and physics; it was not found practicable to apply this classification to the low stratum texts. In addition, the corpus was divided into three parts A, B, and C, each being alike in respect of the

<sup>1</sup> The clause may be defined as a simple as opposed to compound sentence; from a distributional point of view simple and compound sentences are virtually identical—wherever a simple sentence occurs it could be replaced by a compound one of the appropriate type—so that while the general category of 'Sentence' appears as a constituent label in phrase-markers, 'Simple sentence' (= clause) and 'Compound sentence' do not. Simple and compound sentences differ in respect of their internal structure: a compound sentence is one consisting of two or more other sentences (typically joined by some kind of conjunction), whereas a clause consists of a subject and predicate (or something similar, depending on one's analysis).

<sup>2</sup> More precisely, each text ends after the orthographic sentence containing the five-thousandth word.

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)

#### 4 Introduction

stratum and field classifications, so that each contains one high stratum biology text, one mid physics text, three low stratum texts, and so on. The reason for this division is that certain areas of the grammar have been examined in detail with respect to just one or two parts of the corpus.

Each quoted example is accompanied by a five-digit reference number. The last three digits give the serial number within the particular text of the first clause in the quotation, while the first two digits identify the text – as first digit, 1, 2 and 3 indicate high, mid and low stratum respectively, and as second digit 1, 2 and 3 are used for biology, 4, 5, and 6 for chemistry, 7, 8 and 9 for physics, insofar as the field parameter is applicable. Thus clause 17001 is the first clause in one of the three high physics texts, 22500 is the five-hundredth clause in one of the three mid biology texts, 32050 is the fiftieth clause in one of the nine low texts. The only changes I have made in the quotations involve: (a) occasional omission of irrelevant material, marked by '[...]'; (b) the replacement of a few complex formulae, equations, etc., containing one or more of the relators '=', '>', '<', '~' by '[R]' ('relation'), with differentiating subscripts where necessary; (c) italicization of the part of the quotation that is particularly relevant to the point being exemplified – the italics of names of plants, etc., of formulae and of Latinisms may be assumed to be the original author's, all the other italics to be mine unless the expression '*ital. sic*' indicates otherwise.

As mentioned in the Preface, the present work is a substantially revised version of my contribution to *Sentence and Clause in Scientific English* (Huddleston *et al.*, 1968). The other sections of the latter are (a) 'The Clause Complex' (= 'compound sentence'), by R. A. Hudson; (b) 'Some Aspects of Cohesion', by E. O. Winter; (c) 'Some Quantitative Issues' (which includes a section on the difference between the strata), by A. Henrici. In addition the full set of statistical tables derived from the first stage analysis of the corpus is kept in the Department of General Linguistics, University College, London, and is available for consultation there by interested scholars.

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)

## 2 *Mood*

---

### 2.1 **Mood and illocutionary force**

It is important for a number of reasons to distinguish between the grammatical mood of a sentence and the illocutionary force of an utterance (in the sense of Austin, 1962; Searle, 1969). I shall use the terms declarative, interrogative, imperative and exclamative exclusively for types of sentence classified according to grammatical mood, whereas assertion, question, order, exclamation and various other terms will refer to the illocutionary force of different kinds of speech act. The contrast between sentence and utterance (or speech act) in this formulation will suggest that mood is a matter of competence, illocutionary force of performance. But this is not the crucial issue, for we should expect the illocutionary force of an utterance to be largely if not wholly explicable in terms of the semantic–syntactic description of the associated sentence: it is reasonable to require that a competence description should account for at least the illocutionary *potential* of a sentence. My reasons for distinguishing mood and illocutionary force do not therefore depend on the competence–performance contrast; they are as follows:

(a) Firstly there is the practical point that although it is a reasonable requirement that a grammar account for the illocutionary potential of sentences our current descriptions come nowhere near satisfying this requirement. The classification of sentences as declarative, interrogative, imperative and exclamative is quite well-established (which is not of course to say that all grammarians who apply it use exactly the same criteria – or the same terminology) and certainly seems to be valid for some stage in the deep to surface structure progression. Detailed work on illocutionary force is much more recent, and it still remains somewhat programmatic and inexplicit. We do not know what the categories are (the works mentioned above suggest there is quite a large number: suggestion, advice, entreaty, invitation, wager, promise, warning, threat, insult, etc., besides the usual assertion, question, command and exclamation), to what extent they are discrete and mutually exclusive, just what the ‘illocutionary force indicators’ are (Searle, 1969: 30 ff.) and so on.

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)6 2.1 *Mood and illocutionary force*

It will be expedient therefore to use mood as the basis for our discussion; the use of the terms declarative, interrogative, etc. will make it clear that this classification does not claim to account directly for illocutionary force/potential.

(b) Secondly the domains of the two classificatory systems do not wholly coincide. Consider for example the following sentences:

- (i) *a* John didn't know that Bill was coming  
       *b* John didn't know who was coming

As far as the matrix sentences are concerned we shall say that both are declaratives and both are assertions (i.e. would typically be used to make an assertion). But when we turn to the embedded constituent sentences *that Bill was coming* and *who was coming* we see that mood is relevant – the former is declarative, the latter is interrogative – but that the notion of illocutionary force is not applicable, for in any utterance of (i) *a* and *b* *that Bill was coming* and *who was coming* are not taken as separate speech acts. There are compelling syntactic reasons for grouping together embedded sentences like *who was coming* in (i) *b* and independent ones like *who is coming?*, and this grouping is supported by semantic considerations too: I develop these points in 2.2.3.

This is not to say that the mood system of complement sentences is exactly the same as that of independent sentences. In particular the former does not include imperatives. Thus in *John requested/ordered Bill to leave* it seems to me quite inappropriate to classify the complement sentence (*Bill*) *to leave* as an imperative. Nor is it an 'indirect' or 'reported' imperative: the matrix sentence as a whole may be said to report a request or order, but the latter may have been expressed as an imperative (*leave!*), a declarative (*I would like/I order you to leave*) or, in the case of requests rather than orders, an interrogative (*would you like to leave?*). And what makes the matrix a reported request/order is clearly not the infinitival form of the complement but the presence of the specific verb *request* or *order*: compare *John persuaded Bill to go*. I will leave until later the question of whether the domain of the mood system is the sentence or the clause: for the present I shall continue to discuss it in relation to the more general category of sentence.

(c) Finally there is the phenomenon of echo-questions. Echo-questions are of two main types: 'yes/no' and *wh*. Thus an imperative such as *give him £5* might be 'echo-questioned' as *give him £5?* (yes/no: 'is that what you said?') or *give who £5?* (*wh*: 'who did you say to

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)

2.1 Mood and illocutionary force 7

give £5 to?’). The illocutionary force of these echoes is clearly question not request/command (except insofar as a question is a special kind of request, but this is not relevant to my present point). But to account for the grammatical form of such sentences we need to treat them as both imperatives and interrogatives: the natural way to do this is to distinguish between what we may call basic mood and second-order mood. In basic mood, the terms declarative, interrogative, imperative, and exclamative are mutually exclusive; second-order mood, with terms neutral versus echo-interrogative, cuts across the basic mood system, giving a paradigm such as that shown in table 2:1.

TABLE 2: 1. *Basic and second-order mood*

		Second-order mood		
		Neutral	Echo-interrogative	
			yes/no	wh
Basic mood				
Declarative	He went with Bill.	He went with Bill?	He went with who?	
Interrogative:				
yes/no	Did he go with Bill?	Did he go with Bill?	Did he go with who?	
wh	Who went with Bill?	Who went with Bill?	Who went with who?	
Imperative	Go with Bill!	Go with Bill?	Go with who?	
Exclamative	What a crowd went with Bill!	What a crowd went with Bill?	What a crowd went with who?	

The range of elements that can be questioned is greater in echo-interrogatives than in basic ones – *you’re intending to what?* can only be an echo, for example. Yes/no echoes are characterized by ‘question intonation’, but certain types of sentence are ambiguous according as their second-order mood is neutral or yes/no echo. This is so with yes/no basic interrogatives with inverted word-order and rising intonation like *has he finished?* and also with sentences with normal word order and rising intonation like *you’re going with her?* This latter is certainly not necessarily an echo of *I’m going with her*. In its neutral second-order mood interpretation it is probably best treated as an interrogative, since it has the illocutionary force of a question and we could regard the intonation as the illocutionary force indicator. This would be to treat *has he gone?* and *he’s gone?* (with ‘question intonation’) as variants of a single type contrasting with declarative *he’s gone* (non-question

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)8 2.1 *Mood and illocutionary force*

intonation). One difficulty with this analysis arises with such sentences as *I suppose he's gone?* (again with question intonation). For this is clearly not equivalent to *do I suppose he's gone?* – its meaning is more like that of *I suppose he's gone, hasn't he?*<sup>1</sup>

I shall represent the mood of sentences by means of features assigned to the appropriate S node. This may be regarded as a matter of notational convenience: it avoids the necessity for *ad hoc* constituents like the Q(uestion) and Imp(erative) morphemes of Katz & Postal (1964) or the T(ype) morpheme of Rosenbaum (1968), and I have preferred not to commit myself at this stage to a constituent structure analysis involving abstract performative verbs (cf. R. Lakoff, 1968) – such that all command-imperatives, for example, would have an abstract verb of commanding in the underlying structure. My concern is with the internal structure of the various mood types of sentence and the analysis should carry over whatever means we decide on to represent the mood itself in structural terms.<sup>2</sup>

I shall not deal with echo-interrogatives: all examples will be neutral in respect of second-order mood. Declaratives I regard as the unmarked mood category, and no separate section will be devoted to them: interrogatives, imperatives and exclamatives are characterized by various positive structural properties, declaratives by the absence of all of these.

<sup>1</sup> Notice, however, that such a sentence as *do I take it you've finished?* is in effect equivalent to *have you finished?* in the sense that the latter is the question the speaker wants to have answered. It might be possible to formulate rules accounting for the interpretation of various types of sentence in ways that differ from their 'literal' meaning – and in this case we might be satisfied to treat *I suppose he's gone?*, with question intonation, as the interrogative of *I suppose he's gone*, with statement intonation. A similar rule of non-literal interpretation might be envisaged to account for the fact that such sentences as *will you (would you like/care to) help me?* have the illocutionary force of requests; from the point of view of the present analysis, these are regarded as interrogative in mood.

<sup>2</sup> The abstract verb approach would probably give a more satisfactory account of echoes than that suggested above. In such a yes/no echo as *go with her?* there would be two such abstract verbs, one for the interrogative, the other for the imperative. The fact that it is the second-order mood here that correlates with the illocutionary potential of the sentence would be explained by the interrogative abstract verb's being higher in the structural tree than the imperative one, which would be embedded within its complement – cf. 'I ask you whether you request(ed) me to go with her?'



Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)

## 2.2 Interrogatives

**2.2.1 Disjunctive interrogatives.** Interrogative sentences are normally subdivided into the two main types exemplified in:

- (1) This can be split up into two questions, namely, how thick *can [ital. sic] this layer be*, and is this thick enough to do any good? (17464)

Numerous pairs of terms have been proposed for the two classes (cf. Jespersen, 1924: 303), but the terminology is clearly of less interest than the criteria used for distinguishing the classes. It is possible to find three different sets of criteria that have been invoked, explicitly or implicitly, in the classification: one has to do with the presuppositions of the question, a second with the presence or absence of an interrogative word (a *wh*-word), a third with the type of answer expected. These criteria give similar but by no means identical results, so that it is worth considering them in turn.

Consider first then a classification based on the presuppositions of the question: it seems reasonable to suppose that this criterion is implicit in the 'total' versus 'partial' interrogative distinction. Such a sentence as *when did John arrive?* is a partial interrogative in that it presupposes that John arrived and asks only about the time, whereas *did John arrive?* makes no equivalent presupposition and is thus treated as a total interrogative. Notice, however, that in an interrogative like *did John arrive last week?* spoken with contrastive stress on *last week*, John's arrival is presupposed just as much as in the *when* example above. Moreover, in the *wh*-type it is possible to avoid making presuppositions by introducing an *if*-clause: *what, if anything, are you going to tell him?* Thus although in the simple cases yes/no interrogatives can be regarded as total, and *wh*-ones as partial, we can also find partial yes/no and total *wh*-interrogatives. And while the analysis of presuppositions is undoubtedly of considerable grammatical importance, it clearly involves not just interrogatives but all sentences, irrespective of mood. Thus in such a pair as *did John mow the lawn?* and *John mowed the lawn* with contrastive stress on *John* in both cases, it is presupposed as given that someone mowed the lawn; the declarative gives the new information that John was the someone, while the interrogative asks for new information concerning whether or not it was John (see Halliday, 1967*c*, for a discussion of this area of the grammar).

There is a sense however in which presuppositions are relevant to the

Cambridge University Press

978-0-521-11395-3 - The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts

Rodney D. Huddleston

Excerpt

[More information](#)10 2.2 *Interrogatives*

quite different classification of questions into rhetorical versus ordinary? It might be maintained that the answer to a rhetorical question contains *no* new information, which is why it is arguable whether it has the illocutionary force of a question at all: note that Searle (1969: 66) gives as one of the conditions for a (non-examination) question that the speaker does not know the answer, a condition that is clearly not met by rhetorical questions. But I'm not sure how far the rhetorical/non-rhetorical distinction belongs in a grammar of competence, for it raises the difficult problem of 'conduciveness' – cf. Bolinger (1957: 97): 'The devices for creating a conducive or leading Q(uestion), one that shows that a given answer is expected or desired, are intonational, gestural, and verbal. [...] few are determinate enough to make a Q unambiguously conducive without the support of one or more others.'

The presence or absence of a *wh*-word is regarded as a matter of surface structure by Katz & Postal (1964): they claim that there is a *wh*-element in the underlying structure of all interrogatives. I shall return to their arguments below; for the present it is enough to say that as far as independent interrogatives are concerned the presence or absence in surface structure of a *wh*-word correlates quite closely with a significant distinction in deep structure, whether or not this deep structure distinction is itself expressible in terms of the presence or absence of *wh*. To see this let us turn to the third criterion: the type of answer.

The analysis of 'expected', 'appropriate', 'proper' or 'possible' answers belongs of course to the study of competence. In actual performance the addressee may evade the question ('Why should I tell you?', 'I don't know'), challenge the questioner's presuppositions ('When is she coming to Paris?' – 'Who said she *was* coming?'), and so on; nevertheless I take the view that intuitions about linguistically appropriate answers provide as valid data for linguistic analysis as do intuitions about grammaticality, etc.

The most usual classification based on the type of appropriate answer opposes yes/no interrogatives to all others. There are two cases where this classification yields different results from that based on the presence or absence of *wh*. The criteria conflict firstly in the case of Bolinger's (1957: 7) continuation or complementary questions like *his reason being?*, and *John?*, but *later?*. The fragmentary nature of the last two is perhaps such that we would not wish to generate them directly, but I see no reason of principle for treating the first as ungrammatical. This type is not exemplified in the corpus, however, and I shall not have anything further to say about it in this study.