REVIEW ARTICLE

A new Gray's Anatomy of English grammar[1]

GEOFFREY LEECH
Lancaster University
(Received 30 September 2003)

Rodney Huddleston and Geoffrey Pullum, *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press, 2002. Pp. xviii + 1842. Hardback, ISBN 0 521 43146 8. £120.

This article reviews Huddleston & Pullum (2002) from the viewpoint of a co-author of Quirk, Greenbaum, Leech & Svartvik (1985). (This author, however, makes no claim whatsoever to represent the views of the other authors of Quirk et al.) Particular attention is paid to some of the more controversial aspects of Huddleston & Pullum's analysis. It is argued that the two grammars, although similar in their comprehensively wide coverage of English, are not strictly comparable, in that Huddleston & Pullum's grammar is more theory-oriented and Quirk et al.'s grammar is more observation-oriented. These different orientations go with different strengths and weaknesses. In some areas Huddleston & Pullum's more up-to-date account has manifest advantages over that of Quirk et al., but there are also arguably areas where Huddleston & Pullum have not moved with the times.

1 Preliminaries

Every so often, there appears a book which is important enough to fill the reviewer with something like awe. Such a book may be a brilliant ground-breaking book such as Chomsky's *Syntactic Structures*. A different but no less awesome production is the present book, whose strength lies more in being a consolidation and synthesis of existing linguistic theory and description. But to suggest that *The Cambridge Grammar of the English Language* is backward-looking in any sense is misleading, as it also contains a great deal that is new, if not daringly provocative, in its reworking of the well-tilled territory of English grammar. The depth and richness of detail, as well as the breadth of coverage, are extraordinarily impressive, so that there is scarcely a topic that grammatical old-timers like myself cannot read without fresh insight and understanding.

What first of all attracts notice is the work's massive scale and comprehensive coverage. Whoever thought Quirk et al.'s *A Comprehensive Grammar of the English Language* (1985) contained the whole of English grammar within a single pair of covers

---

[1] In writing this review article, I have gratefully received comments and corrections from Bas Aarts, Costas Gabrielatos, Magnus Levin, Joybrato Mukherjee, and Randolph Quirk.

will have to think again. This tome has 1,842 pages compared with 1,779 in Quirk et al., and its information is typographically more densely packed page by page. After some rough-and-ready calculations, I reckoned that the *Cambridge Grammar*, with c. 890,000 words, would not be far short of the length of Jespersen's seven-volume *A Modern English Grammar on Historical Principles*, if Jespersen's first volume on phonetics and spelling (topics largely excluded from the *Cambridge Grammar*) were disregarded.

Secondly, consider the authorship and provenance of the book. Whereas Quirk et al. were a compact authorial team of four, as many as fifteen names appear on the title page of the *Cambridge Grammar*, those of the two authors (both, interestingly, expatriate Englishmen) and of thirteen collaborators, most of whom co-authored one or more chapters: Laurie Bauer, Betty Birner, Ted Briscoe, Peter Collins, David Denison, David Lee, Anita Mittwoch, Geoffrey Nunberg, Frank Palmer, John Payne, Peter Peterson, Lesley Stirling, and Gregory Ward. This distinguished authorial team has also greater geographical spread than the Quirk et al. 'gang of four': it includes representatives from Australia, Israel, and New Zealand, in addition to the UK and the USA. However, unquestionably the chief architect and leading author of the book is Rodney Huddleston, who is named as sole author of seven chapters, and joint author of the remaining thirteen chapters.

The gestation period has matched the size of the book. Intriguingly, the book's CUP webpages inform us that Huddleston conceived its writing soon after publishing his review in *Language* of Quirk et al. in 1988 (with Pullum joining later in 1995). This together with a handsome acknowledgement of Quirk et al. in the Preface[2] perhaps warrants angling this review article to the comparison of the two grammars. As one of the authors of Quirk et al., I will avoid looking at the work entirely through Quirkian spectacles, but cannot refrain, at this initial stage, from a petty complaint: the title of this book reduces acronymically to *CGEL*, which happens also to be the established abbreviation for *A Comprehensive Grammar of the English Language* – as used, among many others, by Huddleston in his 1988 review of that book. Perhaps time will find a solution to this difficulty, but meanwhile, for the purposes of this article, I will use H&P as an abbreviation for the *Cambridge Grammar* and Q et al. as the abbreviation for the *Comprehensive Grammar*.

Considering the many authors of this book, one might expect it to be stylistically uneven and lacking in intellectual coherence. On both these scores, however, the work achieves admirable consistency. It is also consistently well written, so that, in spite of the density of information and argument, it achieves a reasonable standard of clarity and accessibility throughout. On the book's readability, one reservation I have, which will need discussion later (in section 10), is the authors' profligacy in introducing unfamiliar terminology or unfamiliar uses of existing terminology.

---

[2] 'Special mention should be made … of the work of Randolph Quirk and his colleagues … Although the present work often pursues a very different theoretical approach and analysis from that of Quirk et al., their grammar proved an indispensable source of data and ideas' (p. xvi).

Thirdly, let's turn to the content and organization of the book. Its twenty chapters correspond fairly closely to the eighteen chapters plus three appendices of Q et al., chapters 19 and 20 corresponding to appendices I and III (on word formation and punctuation respectively in Q et al.). Apart from chapter 12 of Q et al., on 'Pro-forms and ellipsis', whose contents are dispersed in H&P, appendix II, on stress, rhythm, and intonation, is the only significant area of Q et al. without counterpart coverage in H&P. However, the spread of the material in the remaining chapters and appendices is by no means identical, as table 1 shows. The book ends with a Further Reading section (containing a short bibliography) and two indexes: a lexical index and a conceptual index.

Table 1. *Chapter correspondences between H&P and Q et al.*

Chapters of H&P	Rough correspondences with Q et al. chapters
1. Preliminaries	1. The English language
2. Syntactic overview	2. A survey of English grammar
3. The verb	3. Verbs and auxiliaries, and 4. The semantics of the verb phrase
4. The clause: complements	10. The simple sentence (?)
5. Nouns and noun phrases	5. Nouns and determiners, 6. Pronouns and numerals, and 17. The noun phrase
6. Adjectives and adverbs	7. Adjectives and adverbs
7. Prepositions and prepositional phrases	9. Prepositions and prepositional phrases
8. The clause: adjuncts	8. The semantics and grammar of adverbials
9. Negation	[Part of 10. The simple sentence]
10. Clause type and illocutionary force	11. Sentence types and discourse functions
11. Content clauses and reported speech	[Parts of 14. The complex sentence, and 15. Syntactic and semantic functions of subordinate clauses]
12. Relative constructions and unbounded dependencies	[Parts of 17. The noun phrase, 14. The complex sentence, and 15. Syntactic and semantic functions of subordinate clauses]
13. Comparative constructions	[Parts of 15. Syntactic and semantic functions of subordinate clauses]
14. Non-finite and verbless clauses	[Parts of 15. Syntactic and semantic functions of subordinate clauses, and 16. Complementation of verbs and adjectives]
15. Co-ordination and supplementation	13. Co-ordination
16. Information packaging	18. Theme, focus and information processing
17. Deixis and anaphora	[Parts of 5. Nouns and determiners, and 19. From sentence to text]
18. Inflectional morphology and related matters	[Distributed in various chapters]
19. Lexical word-formation	Appendix I: Word-formation
20. Punctuation	Appendix III: Punctuation

Table 1 cannot be taken, of course, as a guide to the extent to which particular topics are covered in the two grammars. For example, although the noun phrase seems

to be covered by H&P in only one chapter (chapter 5), compared with three in Q et al., this chapter 5 is a book in itself – a mega-chapter of 202 pages. A similar observation can be made about the treatment of the verb: the single verb chapter, chapter 3, roughly equates with both chapters 3 and 4 of Q et al., and its 142 pages are comparable to both those chapters rolled into one (bearing in mind also that the treatment of irregular verbs is postponed in H&P to chapter 18 on inflectional morphology). Among areas of grammar dealt with in much more depth and detail in H&P should be mentioned comparative constructions and negation (each accorded a separate chapter). Other grammatical topics developed in depth in H&P – but not in Q et al. – include unbounded dependencies in chapter 12 (curiously called 'pushdown elements' in Q et al., and handled in a somewhat fragmentary way). A further addition in the analytic apparatus of H&P is *supplementation*, supplements being 'elements . . . which are not integrated into the structure of a sentence' (p.1350) being set off by a prosodic break in speech, or normally by punctuation marks in writing. Supplements include parenthetical main clauses, nonrestrictive relative clauses, and nonrestrictive apposition.

While noting ways in which H&P go beyond Q et al., I should mention some other areas – e.g. the passive – which are less satisfactorily dealt with. I should also mention the extent to which H&P make use of the apparatus of logic and Gricean pragmatics (tense logic, entailment, conversational implicature, and the like) in explaining the semantics of, for example, tenses, quantifiers, and conditional clauses. Although I will not have room to discuss these, they illustrate respects in which H&P impressively achieve greater depth and sophistication than Q et al. in relating syntax to meaning and use.

Not surprisingly, in spite of all this, there is much to argue about in H&P, and I take up some controversial issues in the following sections.

2   Theory versus description

There is no doubt that H&P's grammar is much more theory-laden than that of Q et al. Many of their ideas and analyses show adherence to monostratal phrase structure models of theoretical syntax – but with an added emphasis on functional categories such as predicator, complement, modifier, and, above all, head (the headedness of constructions is a pervasive principle). However, H&P disclaim theory to the extent that they aim (perhaps too optimistically) to urge analyses so persuasively that 'even someone with a different . . . theory of syntax would have to come to a conclusion tantamount to ours if they considered all the facts' (p.19). Thus H&P do not espouse a named syntactic theory, nor do they attempt the kind of formal explicitness that would enable a testable computational model of this grammar to be built. I would attribute this to two factors:

(a) The comprehensiveness achieved by H&P, or indeed by any such detailed and wide-coverage grammar, is not realistically consistent with such explicitness.[3] It is

---

[3] This follows what can be called 'Zadeh's uncertainty principle' that 'In general, complexity and precision bear an inverse relation to one another in the sense that as the complexity of a problem increases, the possibility of

worth noting here that H&P are sparing in tree diagrams or any kind of diagrammatic or tabular display, considering the range of constructions exemplified. Consequently one remains in the dark over many puzzling details of constituent analysis, such as how to deal with some of the noncanonical clause structures (existential and cleft constructions, for example), how to represent the relation between the parts of discontinuous structures, or how to analyse the nicely named structures of 'nonce-constituent co-ordination' – p. 1341.

(b) From time to time the descriptive analyses implicitly suggest an element of doubt or indeterminacy (e.g. p. 172 on modality, p. 619 on 'complex prepositions', p. 1114 on *than/as* + single element). There is informal use of nondeterminate cognitive notions of 'prototype', 'salience', and 'default'. These can be useful to grammarians, in allowing exceptional, dispreferred or peripheral cases to be alluded to without specific commitment. Another favourite term, 'blurring', implies recognition of the fact that grammatical analyses are not always of an all-or-nothing, clear-cut nature.

These imprecisions are wholly understandable (and I would add, desirable) in a grammar 'aim[ing] for as comprehensive coverage as space allows' (p. xv). They sanction a characterization of H&P not as *theoretical grammar* but as a *theory-oriented descriptive grammar*. But H&P dislike indeterminacy, and their theoretical orientation is evident in the space and effort they devote to explanatory arguments in favour of one analysis or against another. For example, one of the longest arguments is on whether in certain cases a reduced clause follows comparative *as* or *than* (pp.1114–17). On the page, these pieces of argumentation show up as blue-shaded boxes, which vary in length from a few lines to four pages.[4]

H&P's determination to arrive at a single correct analysis contrasts with Q et al., who are more apt to accept alternative analyses, or to build gradience into their description.[5] This may suggest that Q et al. are less rigorous than H&P, or rather that Q et al. provide a rather different kind of descriptive grammar – primarily a *reference grammar*, generalizing over observables, aiming to make the description of the language accessible to a wide readership, not just syntacticians.[6] Although H&P maintain that they 'do not assume any familiarity with theoretical linguistics on the part of the reader' (p. xv), in practice the intensity of detail, including much unfamiliar terminology, is likely to deter the nonspecialist.

analysing it in precise terms diminishes' (Zadeh, 1972). The task of writing a comprehensive grammar of a language is, most would agree, exceedingly complex.

[4] As H&P put it, 'The reader will therefore find much more discussion of grammatical concepts and much more syntactic argumentation than is usually found in grammars of English. It is supplied, however, not to establish some wider theoretical point applying to other languages, but simply to persuade the reader that our description is sound' (pp. 19–20).

[5] A well-known case of alternative analyses is the treatment of prepositional verbs as either transitive (verb + prep.) or intransitive (verb separated from the following prepositional phrase) (Q et al.: 91). H&P, on the other hand, persuasively reject both these analyses in favour of verb + prepositional phrase complement (p. 277).

[6] Perhaps the term 'reference grammar' is more readily employed by publishers than by authors. I do not think H&P use this term of their own book, although it is called such on the CUP website.

To elaborate this distinction between H&P as more theory-oriented and Q et al. as more observation-oriented, I will digress a little, arguing for a scale of abstraction running from the most data-oriented (inductive) view of grammar to the most theoretical (deductive) view. (The *corpus-driven* approach, rejecting 'superimposed' grammatical concepts in favour of corpus-derived ones, is a prime example of the former – see Sinclair, 2001, Tognini-Bonelli, 2001; whereas, say, the minimalist program of Chomsky – see Adger, 2003 – is a prime example of the latter.) As H&P point out, if any generalizations are to be made, even the most data-oriented view of grammar requires some model or theory – minimally, say, an account of parts of speech. It is also true, at the other extreme, that a theoretical grammar requires some minimal attention to data: what manifestations of language can occur in given languages. But between these extremes lies a broad spectrum of grammatical perspectives, varying in their degree of attention to theoretical or evidential concerns. The perspective of *descriptive grammar*, a term accepted by both H&P and Q et al., occupies a broad middle ground of this spectrum, and aims at coverage of the phenomena of the language above all, as suggested by figure 1. Far from the naive view that different grammars of the same language are trying

DATA ORIENTATION	DESCRIPTIVE ORIENTATION	THEORY ORIENTATION
corpus-driven or corpus-informed or using other kinds of empirical evidence; variational in focus; nondeterministic (in using frequency, gradience, alternative analyses)	working towards a comprehensive description of individual languages, informed both by evidence (especially corpora) and by theory	using rigorous and explicit formulations; argumentative, with explanatory focus on language; answerability to overarching theory of universals and typology

Figure 1. A scalar view of grammars

to do the same job, it is evident that even different descriptive grammars of a language may have different orientations, making them difficult to compare. This is true of Q et al. and H&P. I would place Q et al. towards the left of the centre of figure 1, and H&P towards the right.

There is also a difference in grammars according to the user-community they address. Theoretically inclined linguists will find H&P's book a goldmine: they will recognize much of the terminology, and the kinds of generalizations which are being made through the use of such categories as 'nominal' as a major constituent of the noun phrase. They will appreciate the 'deep' trees (there is rarely any exception to singular and binary branching with H&P), which are more explanatory of co-ellipsis, scope phenomena, and, above all, co-ordination[7] than the 'shallow trees' of Q et al.

---

[7] However, be it noted that co-ordination is one of the exceptional cases where H&P allow multiple branching – see p.1279. Although co-ordination will be neglected in this article, I should not neglect to mention that chapter 15, mainly on co-ordination, is one of the most impressive and innovative in the book.

On the other hand, many users of Q et al.'s book will find it difficult to adapt to the frequently abstruse terminology of H&P (see section 10 below), together with some radical solutions including a redrawing of the boundaries between word classes such as prepositions and adverbs (see sections 3–5 below), and between major constituents such as phrases and clauses. They may also be happier to stick with the shallow trees such as (21b) below, which, although less explanatory of certain syntactic phenomena, are more usable both for textual analysis and for mapping from syntactic structures to syntactic functions (such as subject, object) and to semantic arguments.

The differentiation between H&P and Q et al. as more theory-oriented and more data-oriented respectively should not be interpreted simplistically. Q et al. are data-oriented in a sense that grew out of the pioneering evidential investigation of grammar at the Survey of English Usage (SEU) in the 1960s and 1970s. H&P's book is not devoid of good illustrative use of corpora, but their use of corpora is exemplificatory rather than an integral part of the analysis. A footnote on p.11 acknowledges the use of three small written corpora, the Brown (American), ACE (Australian), and LOB (British) corpora, together with the *Wall Street Journal*.[8] But these sources (two of them some forty years old) could not have been used to investigate the most important dimensions of language variety: in particular, they contain no spoken English.

Advances in corpus linguistics, and in the availability of corpus resources for English, have been so enormous since Q et al.'s grammar was written, that H&P would not have found it difficult to exploit corpora – not just for finding suitable examples, but for verifying degrees of acceptability, for investigating or checking usage in different varieties, for making or validating statements of frequency, and so on. On the positive side, however, H&P have indeed made good use of attested examples (not just from electronic corpora), and their invented examples are on the whole natural-sounding – seldom falling into the wooden artificiality of exemplification associated with much syntactic theorizing.[9] But for the corpus linguist, however, they show too strong a tendency to judge acceptability in clear-cut ways which suit the analytic point being illustrated. The desire to seek a decisive answer to all research questions is too strong, in particular, when examples of borderline acceptability are judged to be either fully grammatical or fully ungrammatical. The following (presumably unattested) examples from the book are here presented without the asterisk which precedes half of them in H&P. The reader is invited to judge where the asterisk, according to H&P, should or should not be placed:

(1) both Kim, Pat and Alex . . . (p. 361)
(2) Have finished before I return. (p. 932)
(3) To who are you referring? (p. 465)

---

[8] The misaffiliation of LOB as the 'London/Oslo/Bergen' corpus (London replacing Lancaster) is particularly painful to the present reviewer!

[9] Old acquaintances from theoretical textbooks, such as *Sue* and *Max,* make only rare appearances in H&P's examples, although another old friend, *Kim*, makes her (his?) appearance with preternatural frequency.

(4) Don't have eaten all the pizza by the time I get back. (p. 802)

(5) He returned to me the key. (p. 1018)

(6) He often isn't there when you call him, isn't he? (p. 801)

(7) Two serious harms were done to the project's prospects. (p. 337)

(8) Have you loved me and I been so inconsiderate as to make myself unworthy of your love? (p. 97)

(9) We expected all along an improvement. (p.1180)[10]

My guess is that most readers who try this test will not achieve better results than if they had tossed a coin. H&P are somewhat too ready to brand an example as ungrammatical or accept it as grammatical, without allowing for the in-between cases. (Although they do allow for the marking of examples with [?] as of questionable grammaticality, this symbol is rarely used.)[11]

However, it is also true that, despite little use of corpus data, H&P covers variation between BrE and AmE well.[12] Not surprisingly, considering that most of the book was written in Australia, some attention is also given to Australian English, as well as – occasionally – New Zealand English. (This is a clear improvement on Q et al. who, as far as regional variety is concerned, were almost exclusively concerned with BrE and AmE, with only token references to other national varieties.) Although often analytically *avant-garde*, H&P, like Q et al., are culturally conservative in assuming (in the terms of Kachru's concentric circle model – Kachru, 1985) that the 'inner circle' of predominantly native-English-speaking nations determines standard English. In view of the growing prominence of the new Englishes of the outer circle (Singaporean, Indian, West African English, etc.) and the growing use of 'international English' as a *lingua franca* among non-native speakers (Seidlhofer, 2001), it is becoming more questionable today than in 1985 that native speakers' English from the 'first world' is all that a reference grammar need take note of. H&P accept as uncontroversial the notion of a standard international English based on inner-circle language, splitting it into

---

[10] In H&P, the even-numbered examples in this list are judged acceptable, and odd-numbered ones are marked with asterisks. Measured against corpus evidence, some of these judgements are questionable: for example the plural noun *harms* (excluded by H&P) occurs 17 times in the British National Corpus (BNC), in perfectly ordinary-seeming contexts such as (7). Similarly, the use of *both . . . and* with three conjuncts is not particularly unusual: I found more than a few examples in the BNC, mainly in published texts. An exhaustive search of the BNC has revealed no examples of imperatives beginning *Have* + past participle or *Don't have* + past participle (cf. (2) and (4)).

[11] Apart from * and [?], other warning symbols used to label examples are [%] ('grammatical in some dialect(s) only'), [#] ('semantically or pragmatically anomalous') and [!] ('non-standard').

[12] One weakness in H&P's coverage is on the AmE avoidance of plural agreement with singular collective nouns (pp. 501–2), a well-known area of difference between AmE and BrE. Although the authors acknowledge that the plural agreement is 'more common in BrE than in AmE', they go on: 'It must be emphasised, however, that the plural construction is unquestionably fully grammatical in Standard English' (presumably for both AmE and BrE). This statement is typical of H&P's preference for clear-cut grammaticality solutions (see section 2), but overlooks accumulating corpus evidence (see Levin, 2001; Depraetere, 2003: 103–4) of significant differences between AmE and BrE. This is not just a matter of frequency (although, from figures in Levin, 1999: 26–7, it appears that plural verb agreement with collectives is over twice as common in written BrE as in written AmE). It is also a question of lexical and semantic factors: for example, singular names of sports teams, as in *Tottenham is/are winning*, are not attested with plural verbs in AmE (Biber et al., 1999: 189).

two major dialect areas springing from standard educated BrE and AmE respectively (pp. 4–5).

Further, again like Q et al., H&P conservatively follow millennia-long grammatical tradition in placing the written language in the foreground of attention (pp. 11–12). Extensive and varied corpora are now available (see Leech, 2000) for the study of the spoken language, and research on the real data of spontaneous spoken dialogue (e.g. Biber et al. 1999: 1038–125; Carter & McCarthy, 1995) has revealed how far the observed typical forms of spoken language stray from the assumedly 'canonical' forms of written syntax.[13] While H&P acknowledge spoken varieties of standard English, they seem to be thinking of those public, prepared varieties of speech closest to the written medium:

> we are describing the kind of English that is widely accepted in the countries of the world where English is the language of government, education, broadcasting, entertainment, and other public discourse. (H&P: p. 4)

What this leaves out of account is the most commonly exercised of all varieties of English – that of spontaneous conversational dialogue. This big gap in coverage, less excusable today than in the 1980s, raises the question of whether we can accept conversational English as belonging to the standard variety at all. While sympathetic to the construct of 'standard international English' (SiE) for the written medium, I feel that by largely excluding spontaneous speech from their territory, H&P have managed to let this construct of SiE conveniently stand proxy for the language as a whole, ignoring the grey areas that inevitably come into consideration when one tries to apply the term 'standard English' to the spoken medium. Consequently, the clear-cut view of grammar that H&P put forward is aided and abetted by a clear-cut view of the standard language: for example, they dismiss relative clauses with resumptive pronouns as ungrammatical in English (p.1091), although some resumptive-pronoun clauses are not uncommon (and seem peripherally standard) in speech.[14] Nonclausal elements (minor clauses, interjections) characteristic of speech are given little attention, and the much-investigated use of words like *okay, well, like*, etc. as discourse markers is ignored. By saying 'the only legitimate basis for an absolute judgement of incorrectness in a usage manual is that what is being rejected is not in the standard language' (p.10 – in criticism of the prescriptivist tradition), H&P reveal an uncompromising strain in their own thinking.

---

[13] Admirable though H&P's treatment of interrogatives is in chapter 10, close attention to dialogic forms like interrogatives doesn't necessarily redress the general balance in favour of spoken English. Biber et al. (1999: 211) report, in their corpus-based grammar, that nearly half the questions in speech are fragmentary – tags or nonclausal structures – the latter scarcely commented on in chapter 10's treatment. So even for interrogatives, with their built-in bias towards the spoken medium, there is a tendency to overlook what occurs frequently in conversation, and to focus on what is canonically explicit and well-formed.

[14] Prince (1990) gives examples such as *There are always guests who I am curious about what they are going to say*, and comments: 'relative clauses with resumptive pronouns are officially ungrammatical in English . . . However, they are in fact not uncommon in speech.' See also examples in Biber et al. (1999: 622).

Why does this kind of thinking matter in grammar? It shows up in the rather arbitrary way that certain usages are accepted or rejected, without recognition of their problematical status. For example, the 'hypercorrect' use of nominative pronouns after prepositions as in %*It would be an opportunity for you and I to spend some time together* (p. 463) is marked as 'acceptable in some standard English dialects but not in others', in contrast with !*Me and Larry are going to the movies* – marked as nonstandard. However, in terms of frequency and range of use (the criteria H&P claim to be using), there is little doubt that the 'stigmatized' subject construction of *Me and Harry* or *Harry and me* is more widespread than the hypercorrect form of *for Harry and I* (cf. Biber et al., 1999: 337–9). There may be some covert linguistic gentility operating here – preference for the hypercorrect rather than the hypocorrect form – but my main point is that the drawing of lines between standard and nonstandard usage can be fraught with difficulty in speech, and once again there is need to recognize some kind of gradient – in this case a gradient of 'standardness'.

3   Part-of-speech classification

My plan now is to move on to some descriptive issues where H&P are likely to be judged as innovative, if not daring, in their departure from grammatical traditions. Throughout the book, H&P avoid overt critical reference to other grammarians and positions in the published literature – as is understandable, for space reasons alone, in a work of this scope. Instead, they make use of a conveniently generalized aunt sally – known as 'traditional grammar'. This term covers a multitude of sins: it is 'traditional grammar', for example, that insists that there are six forms of *took* – one for each combination of three persons and two numbers – and it is also 'traditional grammar' that gives us the major part-of-speech categories which are more or less universally adopted by English dictionaries: e.g. in treating *if* as a subordinating conjunction, and *then* as an adverb. I would argue though that these are two very different notions of 'traditional grammar'. One comes from a procrustean Latinate grammar that had its heyday in the eighteenth century, while the other is still 'mainstream', in that there will have to be a wholesale revision of not only dictionaries, but grammars, part-of-speech tagged corpora, and ELT textbooks, if H&P manage to persuade the world that their analysis is right.

As a case in point, H&P redefine some major word classes: they redefine prepositions as a much expanded and numerous category, to the detriment of adverbs and subordinators (or subordinating conjunctions), which are much reduced in number. Other categories which undergo substantial change are determinatives (the term 'determiner' is reserved for the corresponding functional category, so that H&P have unhelpfully reversed the use of these terms 'determiner' and 'determinative' in Q et al.). Auxiliary verbs, pronouns, and numerals disappear from the inventory of word classes at the highest level. Auxiliary verbs become in effect main verbs; pronouns become a subcategory of nouns, and numerals are subsumed under the determinative category.

4   Prepositions, adverbs, and subordinators

The two most far-reaching of these recategorizations are those of preposition, adverb, and conjunction (which are interrelated, and can therefore be best considered together) and those of determinative and pronoun (see the next section).

In the H&P treatment, there is a fascinating inverse relation between prepositions and adverbs. Whereas grammars like Q et al. define prepositions as close to their etymological meaning – words which are *placed before* a noun phrase or other prepositional complement (barring, of course, preposition stranding as a special case), H&P extend the class of prepositions to include *post*positions like *ago*, and 'intransitive prepositions' like *about, afterwards, ashore, downstairs, east, else, heavenward(s), here, home, indoors, north, now, south, then*, *west, westward(s), when*, *where, whenever, whence*, as well as clause-introducing prepositions such as *if*, *because* and *while*. Although 'pre-position' is no longer a suitable label, H&P uncharacteristically stick with the established term *preposition* (p. 602) despite its loss of appropriateness.

On the other hand, H&P diminish the class of adverbs so that it matches more closely its etymological meaning:

> Adverbs characteristically modify verbs and other categories except nouns, especially adjectives and adverbs. (p. 563)

As for prepositions, their definition as an extended class is so vague that it has roughly changed places with the woolly definition of adverbs found in 'traditional grammar':

> PREPOSITION: a relatively closed grammatically distinct class of words whose most central members characteristically express spatial relations and serve to mark various syntactic functions and semantic roles. (p. 603)

Just about the only definite thing this tells us about prepositions – if one can ignore hedging expressions such as 'most central members' and 'characteristically' – is that they are associated with spatial relations. Surely this is a reversion to the old-fashioned notional definition of parts of speech, rightly castigated by H&P in their opening chapter (pp. 29–31). But in any case it doesn't help us with words like *then* and *if*, neither of which is spatial.[15]

H&P's case for considering words like *then* and *if* prepositions depends on the argument of parallelism with verbs (and other categories) subclassified according to their complementation, as shown in table 2. *See* and *since* are versatile words which can accept different complement patterns. But consider a verb like *elapse* which can only be intransitive, a verb like *beat* which can only be monotransitive, or a verb like *retort* which can only have a clausal complement. If these are all verbs, goes the argument, then on the same grounds both *then* (intransitive) and *if*

---

[15] One of the difficulties of H&P's expansion of the preposition category at the expense of adverbs is that a word standing alone as head of an adjunct or modifier constituent can belong to either category. Hence the authors sometimes find difficulty deciding which of the two classes a word belongs to – see their discussion of *else* (p. 615 n.).

Table 2. *Comparable complementation of verbs and prepositions*

	(a) verb	(b) preposition
(10) no complement	*I <u>see</u>.*	*I've met her <u>since</u>.*
(11) NP as complement	*I <u>see</u> the problem.*	*I've met her <u>since</u> the game.*
(12) clause as complement	*I <u>see</u> you're alone.*	*I've met her <u>since</u> she left.*

(with clausal complement) can be deemed prepositions just as *with* can. This argument gains force from the fact that items such as *since* can be given a unitary description as prepositions, instead of being classified as an adverb in (10b), a preposition in (11b), and a subordinator in (12b). According to H&P, this sort of unitary assignment avoids 'a quite pointless complication of the grammar' (p. 1015), since the 'multiplication of categories for a single word with a single meaning makes no sense'. However, this argument is undermined somewhat by the fact that the causative meaning of *since* is peculiar to its use as a subordinator, whereas the temporal meaning applies to all three items. Also, the parallel between (a) and (b) sentences in table 2 ignores the fact that inflectional morphology provides a strong reason for regarding verbs as verbs, whatever their complementation, whereas there is no such morphological argument for a 'super-class' of prepositions.

Another of H&P's arguments is that traditionally prepositions are allowed to take clausal complements of certain kinds, but not others: e.g. a standard preposition like *on* can be followed not only by an NP, as in *it depends on <u>the weather</u>*, but also by a clause in *on <u>arriving at the hospital</u>* and *it depends on <u>whether you vote</u>*. They see this as muddying the distinction between prepositions and conjunctions that 'traditional grammar' upholds (pp. 1012–13). But they miss an important generalization captured in Q et al. and similar grammars, in the category of nominal clause – a clause, that is, which has the privilege of occurrence of an NP. Interrogative clauses such as *what you say* or *whether you vote* above are nominal, and so are *-ing* clauses such as *arriving at the hospital*. Prepositions that can accept 'nominalization' NPs as complement can also accept such nominal clause complements. Here are illustrations with *in, by, of, about, at*, and *to*:

interested in a house	intrigued by her boots
interested in buying a house	intrigued by your eating goldfish
interested in where you'd buy a house	intrigued by whether it would work
it's part of the deal	talked about the new alliance
it's part of growing up in the city	talked about getting a new car
it's part of whether you succeed or not	talked about when we'd leave
surprised at the mayhem	with a view to marriage
surprised at making a profit	with a view to getting married
surprised at how long it lasted	with a view to which venue was best

On the other hand, subordinators such as *if, while, because, although*, and *unless* do not take such complements, except in the case of what Q et al. call 'abbreviated clauses'. These abbreviated clauses involve ellipsis of the subject of the clause and a finite form of *be*, but retain the subordinator. Notice it is possible for an abbreviated clause to contain a predicative adjective phrase, noun phrase or prepositional phrase as predicative complement, as well as an *-ed* clause: *if still working, if unsound, if a male model, if at home, if accepted*, etc.

Consequently, in spite of an overlap between prepositions and subordinators in the area of *-ing* clauses, there is a clear difference between the array of sequences that can occur in a prepositional phrase and an abbreviated clause:

PREPOSITIONAL PHRASE:	ABBREVIATED CLAUSE:
(13) despite their arrival	*although their arrival
(14) despite dating from the 1400s	although dating from the 1400s
(15) despite who said what	*although who said what
(16) *despite careful with money	although careful with money
(17) *despite made of solid steel	although made of solid steel
(18) *despite they arrived on time	although they arrived on time
(19) *despite under new management	although under new management

(The asterisk for *although their arrival* could be removed if the following noun phrase could be interpreted as a predicative complement: *although a woman of fortune*.)

The paradigms displayed above represent quite a powerful generalization differentiating the syntactic potential of prepositions and conjunctions. This generalization would, of course, disappear if the two categories were merged[16] in the way H&P propose (although H&P could still handle the same phenomena by the subcategorization properties of individual items). It is also worth bearing in mind that stranded prepositions occur in English grammar, while stranded conjunctions do not:

a bus that I'm waiting for   *a true champion that he'll win if

It is true that a few words have the syntactic privileges of both classes: *before, after, since, notwithstanding, as*, and some descriptive disadvantage arises from their separation. But surely this is less than the advantage of keeping the classes separate. Similarly, there is some disadvantage in keeping separate the cases like *except* (preposition) and *except that* (conjunction), while H&P would treat *that* alone as the optional conjunction, introducing a declarative content clause. What I feel sure of is that the arguments for merging are not so 'compelling' (p. 1013) as H&P claim.

---

[16] The merger is not complete, but traditional subordinating conjunctions are abolished in H&P except for a 'rump' category of subordinators consisting of *that*, *whether*, and *if* (= 'whether') (p. 600). Tellingly, the authors here condone the fault they criticize in traditional grammar, of splitting a word unnecessarily into two different word classes: for traditional grammar, conditional *if* (for H&P a preposition) and *if* (= 'whether') (for H&P a conjunction) are both subordinators. (A similar split of *how* into adverb and subordinator categories is allowed for on p. 954.)

A problem of defining the expanded class of prepositions is that it lacks morphological clues. This leads H&P to rely on a strange mixture of tests – a negative inflectional criterion (lack of *-ly*) and a positive syntactic criterion (ability to take a complement) – as distinctive of prepositions:

> This provides, then, a reasonably clear basis for distinguishing between prepositions and adverbs. If a word not ending in the *-ly* suffix licenses a complement, it is not an adverb; if a word other than those of the type covered in [30–31] fails to license a complement, it is not a preposition. (p. 617; [30–31] refers to lists of spatial and nonspatial prepositions)

This definition statement may be 'clear', but is it motivated? Implicitly H&P seem to have fallen back on the traditional idea that a preposition needs a prepositional complement. As they say earlier: 'Prototype prepositions have NP complements, and other items will be admitted into the preposition category only if there is positive evidence to support such an extension beyond core members' (p. 612). It is significant that H&P here apply the nondeterministic notion of prototype to a syntactic category, whereas elsewhere they tend to confine prototype thinking to semantic categories. Altogether, the attempt to elevate the preposition to a major word category is unconvincing, and seems to hark back historically to Chomsky's feature theory of word categories (Chomsky, 1970), now largely discarded (see Baker, 2003: 2).
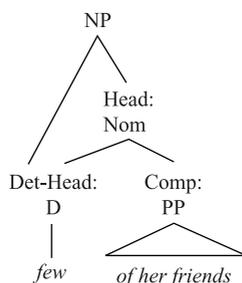
5   Trade-offs between lexicon and grammar

H&P have comparatively little to say about the lexicon or dictionary their grammar implies – perhaps someone should publish such a lexicon. However, the word-class conflations mentioned above do not entail a simplification of the grammar. If anything, it is a simplification of the lexicon if one word is assigned to one word class rather than two or three. But of course the different frames or functions into which a word fits still have to be specified within one entry of the lexicon. H&P generally seem to follow a principle of simplifying the mapping from lexemes or word forms to word classes as much as possible. If we consider further instances: *both* is considered a determinative whether it functions in a simple NP or as a correlative co-ordinator; and the words *some*, *any*, and *that* (except for *that* as a subordinator) are assigned to a single class of determinative. This same label applies to them, whether (in Q et al. terms) they function as head (*That is mine*), determiner (*That hat is mine*) or adverb (*that tall*). This simplifies the lexicon in one direction, but passes on the complexity to the mapping between grammatical words and their grammatical functions. For H&P, the three functions illustrated by these examples are determiner, determiner-head, and modifier.

The second of these functions exemplifies an important theoretical innovation by H&P: the **fused-head** analysis permitting a single word to span more than one syntactic function in the same NP. This introduces a further complexity into the

lexicon–grammar mapping, constituting a radical breach of the single-mother condition normally presumed to apply in phrase structure grammars. For here we have a single word, say *few* in *few of her friends* (see (20)), as daughter of two higher nodes, nominal and NP.

(20)

```
                    NP
                   / \
                  /   Head:
                 /    Nom
                /    /    \
          Det-Head:      Comp:
             D            PP
             |           / \
            few      of her friends
```

(from Huddleston & Pullum, 2002: 412)

An advantage of this, from H&P's point of view, is that it enables them to assign a single part-of-speech analysis to the items spelt *this, these, those, some, any, several*, etc., whereas in Q et al., for example, these have dual roles as determiner and pronoun. H&P can do this, because they see determinatives as having a dual function – both as determiner and as determiner-fused-with-head. They argue (pp. 420–1) unconvincingly, in my view, the necessity of this analysis, by rejecting alternatives relying on (a) multiple part-of-speech assignment, (b) ellipsis, (c) dual functional analysis of these words as head and determiner – though why (c) would be a worse solution than the dual functionality of determiner and determiner+head which the authors adopt is unclear. Similarly, regarding (b), they accept ellipsis in other aspects of the NP, so why not here? There are a number of unconvincing aspects of this fused-head analysis. One is that having argued for the absence of the genitive inflection from fused-head items, H&P go as far as to include compound pronouns like *somebody* and *everyone* as determinatives – despite their readiness to accept genitive forms (*somebody's*, *everyone's*). One argument they give for treating these compound pronouns as determinatives is that 'because determiner and head are fused it is not possible for these forms to take internal pre-head dependents' (p. 423) such as **sensible anybody* – but of course this would also apply if these items were treated as pronouns as most grammarians assume. It is also a strange consequence of the fused-head analysis that demonstratives are determinatives (not pronouns), although they frequently behave similarly to the pronoun *it* in deictic/anaphoric use: '*Hello – who's <u>that</u>?*' '*<u>It</u>'s Hilary*', etc.

Another consequence of this analysis is that the pronoun category, like those of adverb and subordinator, is reduced to the core class of personal pronouns and a few outlying groups such as reciprocal, interrogative, and relative pronouns plus, unexpectedly, *yesterday, today*, and *tomorrow*. (These last three are argued to be pronouns 'by virtue of their inability to take determiners' – but why does this same argument not apply to *somebody, everything*, etc.?)

The fusion of constituents comes in more convincingly with what H&P call the **fused relative construction** ('nominal relative clauses' in Q et al.), as in *I spent what she gave me*. The by-now mainstream argument that *what she gave me* here is an NP, not a clause, is strongly supported by H&P, who assign to *what* in this construction the combined functions of head of the NP and prenucleus in the relative clause. Nevertheless, H&P acknowledge that 'it is by no means a straightforward matter to distinguish between the fused relative and subordinate interrogative [clause] constructions' (p. 1077). Although semantically the fused relative construction is undeniably NP-like, syntactically its clause-like structure is likewise difficult to deny – without the special mechanism of fused constituency.

The enlargement of the preposition category, as already noted, leads to a reduction of subordinators (or subordinating conjunctions) to a residual category of five: *that, whether, if* (= whether), and *for* and *to* (introducing infinitive constructions). This goes with the reduction of the categories of [finite] subordinate clauses to three: content clauses, relative clauses, and comparative clauses. These are defined by their structure: content clauses, like main clauses, are subdivided into declarative, interrogative, and exclamative types. Content clauses are also treated as the 'default' category having the least constraint of defining structural characteristics and occurring in varied positions corresponding to Q et al.'s adverbial clauses, *that*-clauses, *wh*-clauses, etc.[17] Certainly there is economy in this analysis of subordination, both at the word-class level (i.e. few subordinators) and at the clause level (i.e. few clause types), but this is at the expense of both formal and functional economy. Content clauses, in particular, are so variable in their form and function that they seem totally unconstrained. They can function as subject, object, adjunct, clausal complement, predicative complement, complement of nouns, adjectives, adverbs or preposition, postposed subject or object, supplement. Structurally, they can have normal declarative order, various types of preposing, subject–auxiliary inversion, an infinitive main verb, and so on.

One might summarize the effects of H&P's innovations noted in this and the preceding section as 'swings and roundabouts'. By assigning most subordinating conjunctions to the preposition category, H&P have in effect put more of the grammar into phrases, and less into clauses. This is good from the point of view of constraining the form–function interrelations: phrases are associated with particular word classes as head (e.g. prepositions as head of prepositional phrases), and on the whole the functions of phrases in clauses are well defined. However, the negative side is (a) at clause level, and (b) at word-class level. At clause level, the relation between subordinate clause types and their functions in higher constituents is left unclear. Although the ultimate head of a clause is said to be a verb, this does not seem to constrain the positions of

---

[17] Declarative content clauses typically have an SVO order of constituents, but this doesn't apply to some constructions H&P deem content clauses. For example, the proportional clause of *The more we talked, the more I liked her* (p. 970) is syntactically similar to a comparative or relative clause – but is assigned 'by default' to the content clause category.

clauses in other clauses and in phrases. And the content clause (see last paragraph) is a complete chameleon category.

At word-class level, the problem is again that, although assignment of words to multiple word classes is reduced (conspicuously for words like *before* and *since*), the function of these word classes in higher constituents is allowed to proliferate. As an example, we might take the words *both* and *either*, which H&P consider to be determinatives not only in their usual NP-based functions, but also in their function as first element of a correlative co-ordination construction (e.g. *both at home and abroad*; *either you like it or you don't*). In this respect H&P seem to turn a blind eye to multiplicity of functions, and it is particularly difficult for the user of the grammar to keep a check on such multiplicity, as the lexical index (with several exceptions) gives no information about either the word classes or the functional distribution of words. This is particularly unhelpful with a word like *either*, which is simply listed with twenty-five undifferentiated references in the index, although in dictionaries and other grammars it is likely to occur with four grammatical labels (determiner, pronoun, correlative co-ordinator, adverb), signalling its functional potential – information not recoverable from H&P.

6   Constituent structure: deep trees versus shallow trees

As noted in section 2 above, H&P adopt a deep-tree rather than a shallow-tree style of constituent structure, with a large number of nodes per sentence and a small number of daughter nodes per parent node. Q et al., on the other hand, tend to favour shallow trees, although they also occasionally offer the deep-tree analysis as an alternative. See (21a and b) for contrasting tree diagrams of the simple sentence: *They must have been eating popcorn*.
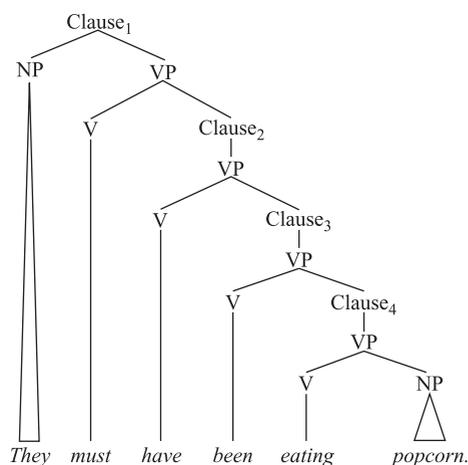
(21a)   (21b)
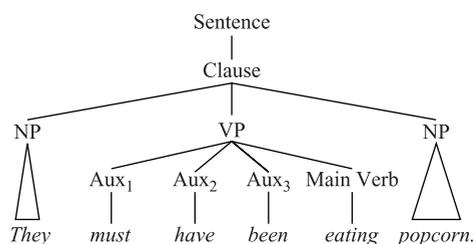
Deep-tree structure shows up strongly in the analysis of both VPs and NPs. For NPs, H&P adopt a multi-level structure involving intermediate *nominal* nodes, whereas in VPs, they adopt the multi-level auxiliary = main verb analysis (see (21a) above). This has advantages in being more explanatory of semantic relations and syntactic functions within the phrase. Also, much of it is well motivated by potential constituency under co-ordination and ellipsis. On the other hand, it has the disadvantage of being unwieldy for practical tree representation. Intermediate nodes and multi-word heads are routine, so that single-word heads in the sense of Q et al. have to be characterized as 'ultimate heads'. H&P themselves resort to practical short-cuts in representing tree structure, such as ignoring the nominal node level or the AdjP level where there is singular branching, in a simple noun phrase like *a woman* (p. 55) or a simple case of adjectival modification: *an excellent result* (p. 57): 'we simplify the tree diagrams by omitting the higher-level constituents if they consist of just a head element'. Another drawback is that with more than one modifier the structure of more complex NPs can be semantically over-determined, so that *a recent experiment on rats* could be analysed either as *a* [[*recent experiment*] [*on rats*]], or as *a* [*recent* [*experiment* [*on rats*]]]. A difference of constituent structure doesn't always go with a difference of meaning.

Turning now to VPs (which are analysed in H&P, unlike Q et al., as including complements and adjuncts of the verb), H&P espouse the *catenative-auxiliary* analysis (p.104) whereby in effect auxiliary verbs are considered main verbs.[18] More exactly, for H&P, auxiliary verbs (except the copula *be*, which is also an auxiliary – strangely called 'non-core' in spite of its overwhelming frequency) are a subset of catenative verbs – verbs taking nonfinite clauses as complements (such as *want, promise, begin*, etc.). H&P contrast their analysis with the *dependent auxiliary* analysis found in Q et al.[19] and elsewhere, and discuss at length (pp. 1210–20) its advantages relating to negation, temporal specification, constituency, preposing, position of adjuncts, and co-ordination. The consequence of this analysis is that subordinate clause constituents are recursively proliferated, as shown in (21a) above, as contrasted with the dependent auxiliary analysis à la Q et al. illustrated in (21b).

H&P characteristically adopt 'categorial thinking' in following the logic of the catenative analysis to its ultimate, while Q et al. adopt 'noncategorial thinking' in postulating a gradient between the analysis of (21b) (for auxiliaries) and (21a) (for catenatives), recognizing quasi-modals, semi-auxiliaries, and catenatives as intermediate classes on a gradient between auxiliaries and main verbs. However, with uncharacteristic tolerance of alternative analyses, H&P recognize the practical merit, in text analysis, of adopting the dependent auxiliary analysis. For the VP in (21b) they

---

[18] An important aspect of H&P's catenative analysis is that the complements of a catenative verb are not assigned a more specific function, such as object, predicative, or adjunct.

[19] However, strictly Q et al. do not commit themselves to the position that every phrase has a head – thus, for them, neither the auxiliary nor the main verb is a head.

use the term 'verb group', and comment: 'This ad hoc term reflects our view that the category is not theoretically justified but may have some practical descriptive value' (p. 1213, note).

To provide one more illustration of the deep-tree analysis, we turn to prepositional phrases. The interest here focuses on 'complex prepositions' such as *in front of*, where H&P (pp. 620–5) generally opt for an analysis which embeds one PP in another, as in (22).
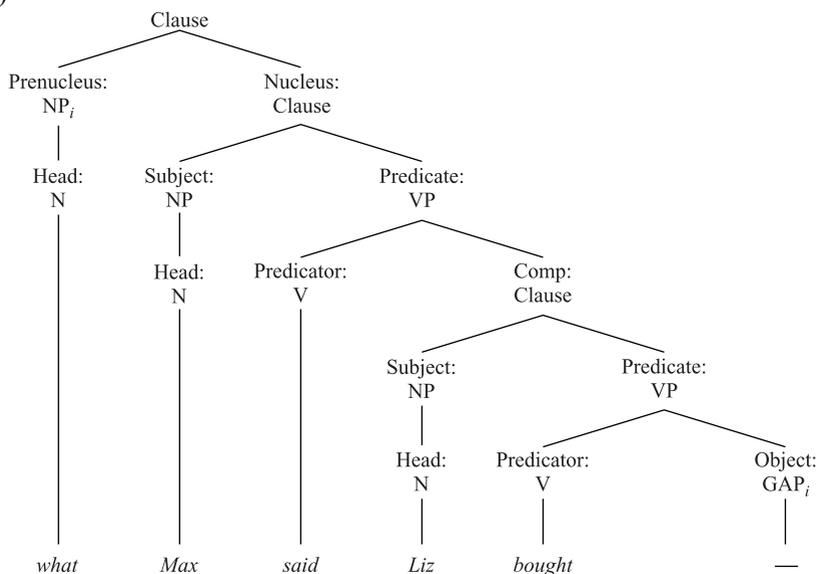
(22)
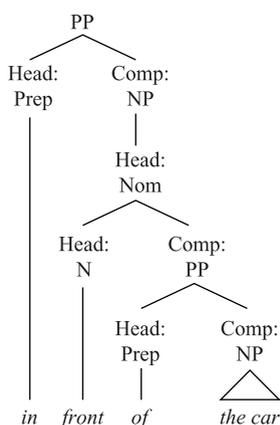
(from Huddleston & Pullum, 2002: 620 – example (9a))

What is notable here is H&P's clear determination to demarcate syntactic considerations from lexico-semantic ones. In terms of grammaticalization theory, *in front of* is evidently well on the road to coalescence, lexically and semantically, into a single unit, and it might be argued that this affects syntax (in making *in front of* comparable with its antonym *behind*). But H&P dismiss the single-constituent 'complex preposition' analysis except in a few cases. Again, this contrasts characteristically with Q et al., who opt for an analysis of the same phenomena in terms of gradience – a gradient between the single-preposition and preposition + noun + preposition analyses.

7   Gaps versus movement

In keeping with their nontransformational phrase structure approach, H&P scrupulously avoid any implication of processes linking one tree structure with another. Instead, they make extensive use of gap constituents à la GPSG, and of coindexing (*i* . . . *i*) between gaps and other constituents. This works pre-eminently well for unbounded dependency constructions (see (23)).

(23)

```
                                Clause
                    ┌─────────────┴──────────────┐
              Prenucleus:                    Nucleus:
                 NP_i                         Clause
                  │              ┌──────────────┴──────────────┐
               Head:         Subject:                    Predicate:
                 N              NP                           VP
                  │             │              ┌─────────────┴────────────┐
               Head:       Predicator:                           Comp:
                 N              V                                Clause
                                            ┌──────────────────────┴──────────┐
                                        Subject:                         Predicate:
                                           NP                                VP
                                            │              ┌──────────────────┴──────────┐
                                         Head:        Predicator:                    Object:
                                           N               V                         GAP_i
                                            │               │                          │
                what      Max      said     Liz        bought                         —
```

(from Huddleston & Pullum, 2002: 49 – example (7))

This strategy enables the canonical declarative order of major constituents to be preserved while accounting for mismatches between the underlying semantic structure and overt syntactic structure (e.g. in process parlance, dealing with fronting, postponing, raising, or omission of constituents). All implication of movement or of other processes linking an underlying and a surface phrase marker is studiously avoided.[20] Very occasionally a process term is used: for example, 'raised complement', 'preposing', and 'PP fronting' – but no concession to transformational thinking is made in the way such terms are defined. Thus a *raised complement* is defined simply as 'one which belongs semantically in a lower clause than that in which it functions syntactically' (p. 65). However, the terminological consequences of avoiding transformational thinking can be less than happy, as when the phenomenon termed 'transferred negation' in Q et al. is described as 'conventionalisation of specificity increase' (p. 842).

H&P invoke gap analysis more questionably when they apply it to initial adjuncts as in *If you pay me_i, I'll do it___i* (p. 1092) or to preposing of NPs as in *Anything you don't eat put ___ back in the fridge* (p. 1372) – where ___ signals the gap. In the latter case there is no coreferential indexing and the gap simply marks the position that the object of *put* would canonically take in the deep-tree analysis. But here H&P may be said to have their cake and eat it, stipulating that 'prenuclear elements that are linked to a gap are interpreted as having the function of that gap' (p. 1038). Hence in

---

[20] A similar nontransformational stance is maintained by Q et al., who define a bi-directional relation of 'systematic correspondence' as a means of making generalizations which could otherwise be expressed transformationally. However, Q et al. are somewhat less systematic (or less pedantic?) than H&P in avoiding transformational implications of terminology.

this manner of speaking, *put* can be said to have two objects – of which one is a gap constituent, and one a preposed NP termed an object 'in a secondary, derivative sense' (ibid.). On the other hand, in the shallow-tree analysis of Q et al., *Anything you don't eat* would be the only object of *put*, although in a marked position, and no gap would be needed.

8   Present and preterite modal auxiliaries

The last controversial topic I will pick out is that of present-tense and preterite modals, one of the topics of chapter 3, which in general offers rich, thorough, and revealing analyses of the syntax and semantics of the verb. H&P regard the present and preterite forms of the modals as forming a single lexeme: hence there are four central modals only: ***can***, ***may***, ***shall***, and ***will*** (with one or two extra ones without present/preterite alternation).[21] I see this as illustrating once again H&P's seeking clear-cut solutions and also simplifying the grammatical interface of the lexicon. After all, it is simpler and more regular to have a grammar with four-plus modals alternating between present and preterite inflections, than one with nine or so modals, with *can*, *could*, etc. as separate lexemes.

But obviously the truth lies somewhere between these two solutions – in some respects the 'preterite' modals are analogous in form and meaning to the 'present-tense' forms, but in other respects they are totally different. Furthermore, they are inflectionally extremely irregular or defective, and their 'preterite' forms have senses which are unparalleled by the present-tense form. This is particularly the case with *should* – a recent corpus study of this modal revealed that less than 5 per cent of *should*s could be semantically considered to be past/remote forms of *shall* (Leech, 2003: 233). The same study showed that *shall* has declined in frequency by about 40 per cent in BrE and AmE between 1961 and 1991, so that *shall* is roughly four times less frequent than *should* even in written English – while in speech it is becoming a rarity, particularly in AmE. As H&P admit (p. 196), the past forms of the modals contrast markedly with the preterites of other verbs, in that their use to express modal remoteness is far more common than their use for past time, while with other verbs the situation is completely the opposite. Furthermore, modally remote (i.e. hypothetical, tentative) uses of the preterite modals can occur in reference to nonpast time without obvious constraint: *We might see her tomorrow* contrasts in this respect with **We liked to see her tomorrow*.

Over a long period of history, the past-form modals have diverged from their present-tense analogues to the stage where it finally makes little sense to call one the preterite form and the other the present form. Most of the generalizations one can make about the preterite in English do not apply to these words. On the other hand, it would also be wrong to claim that there is now no relatedness between the present and preterite forms.

---

[21] H&P use ***bold italics*** in citing a lexeme in abstraction from its inflectional forms. Thus *can*, here, subsumes both *can* and *could* (and also *can't* and *couldn't* – since H&P consider the reduced negative *n't* to be a suffix, not a clitic).

The divergence, arising partly through the gradual process of grammaticalization, is still in progress, and I maintain that it would be better if the two possible analyses, each with its strong and weak points, were acknowledged.

Here I focus particularly on *would*, which H&P (given their stance on *would* as the preterite of *will*) take to be semantically analogous to the present form *will*. Accordingly H&P reject the 'mood marker' meaning (Q et al. p. 234) whereby *would* simply conveys hypotheticalness without expressing any of the predictive (epistemic) or volitional senses of *will*. This means that for H&P, there is no modally remote equivalent of the simple present, as in (24):

(24) If she's here, she's in her office. (OPEN CONDITION) (example 17v on p. 742)

Instead, H&P argue (p. 752) that the hypothetical *would* of (25) is the modally remote equivalent of (26), with *will* in the apodosis:

(25) If she was/were here, she *would be* in her office. (REMOTE CONDITION)

(26) If she's here, she *will be* in her office. (OPEN CONDITION)

Interestingly, the simple-present protasis in (24) and (26) is the most common type, tensewise, in *if* conditions, and the most common apodosis following it has the simple present again, as in (24) (with the *will* apodosis in (26) being the second most common). So H&P give us no hypothetical equivalent of the most common type of open condition. It is also noticeable that the time reference of (25) is most likely to be unrestricted present, like that of (24), whereas the time reference of (26) is more oriented to the future.[22]

Table 3. *Comparison of H&P and Q et al. on the hypothetical use of* would

Huddleston & Pullum:	OPEN	REMOTE		Quirk et al: (implicit)	OPEN	REMOTE
	Simple pres.	(NONE)			Simple pres.	*would* + V
	will + V	*would* + V			*will* + V	

The positions in table 3 are those taken by H&P (p. 752) and Q et al. (pp. 228–35); however, as often happens, Q et al. are less explicit than they could be on this matter. Essentially, Q et al.'s position is that sometimes *would* behaves like the preterite of *will*, and sometimes not.

However, H&P appear to contradict their own position and to go over to Q et al.'s position (on p. 748) when they give (28) as the remote equivalent of (27):

(27) If you *are* under 18 you *need* parental approval.

(28) If you *were* under 18 you *would need* parental approval.

---

[22] The future use of epistemic *will* appears to be much more common than the present ('predictability') one – see Coates (1983: 171).

There are other arguments involving infrequency and marginal acceptability. The perfect *will have* + p.p. is infrequent (see note 23 below), and cannot easily be regarded as the counterpart of the past remote construction:

> (29) If I said that, I *was telling* a lie.   (OPEN CONDITION − PAST)
>
> (30) If I'd said that, I'*d have been telling* a lie.   (REMOTE CONDITION − PAST)
>
> (31) If I said that, I'*ll have been telling* a lie.   (OPEN CONDITION − PAST)

It is entirely natural and reasonable to regard (30) as the remote (counterfactual) equivalent of (29). But H&P's nonremote version of (30) would be (31), substituting *will have* + p.p., and thus replacing the preterite by an infrequent construction[23] which very rarely occurs in any *if* conditional. In fact (31) seems to be only marginally acceptable.

Now consider some other constructions where the simple present-tense construction corresponds to *would* + infinitive. *Do you mind (if) . . . .?* is a form of indirect request, but it is also possible to make the request more tentative and polite by using hypothetical *would*: *Would you mind (if)?*[24] A corpus search I carried out (using the BNC) showed, unsurprisingly, that these two alternatives are vastly more common than *Will you mind (if)?* Moreover, the option with *will* was not, in the BNC corpus data, an indirect request, but rather a question about the addressee's future feelings, as in *Will you mind dining alone?* Similarly, *I suggest* and *I would suggest* are common ways of introducing the speaker's opinion, one more tentative than the other. But *I will suggest* is uncommon, and is not used for that purpose (again, this is based on a BNC search). Another example: *It seems that . . .* and *It would seem* are both common, but *It will seem* occurs with microscopic frequency in the BNC.[25]

9 Frequency

My arguments above for the implausibility of H&P's position on *will* and *would* depend considerably on frequency, and this is something that needs clarification. I maintain that information about frequency can be important for a descriptive grammar – particularly for its more practical applications, e.g. in language teaching and NLP – although there is a popular theoretical argument that frequency is irrelevant to grammar, since frequency is a matter of the use of the grammar rather than of the grammar itself. It is indicative of H&P's position being more towards the 'theoretical' end of the spectrum of figure 1 that they have little use for frequency, although they mention facts of frequency from time to time.[26] This means they fail to give frequency information useful to the user

---

[23] In a random 740 examples from the written BNC, only one *will* + Perfect apodosis occurred. (There were 14 other modal + Perfects, including 6 of *could* – Costas Gabrielatos, personal communication.)

[24] In the BNC, *Do you mind* occurs 204 times, *Would you mind* 124 times, and *Will you mind* 5 times.

[25] In the BNC, *It seems* occurs 2,694 times, *It would seem* 279 times, and *It will seem* twice.

[26] Although H&P occasionally risk frequency statements, these cannot always be trusted. For example, H&P claim (p.1594) that for nouns like *virtuoso* 'the regular plurals are much more frequent than the foreign ones

of the language, such as the fact that short (or agentless) passives are more than four times as frequent as long passives.[27]

Although advocating frequency information drawn from corpora as an input to descriptive grammar, I would not go as far as to argue that acceptability or non-acceptability in the language can be determined solely by presence or absence of phenomena in a corpus. Of course, the contexts in which grammatical forms occur need to be inspected: a corpus can contain aberrant forms, dialect forms, etc. On the other hand, if a corpus shows a grammatical phenomenon as occurring a considerable number of times, in contexts which suggest normal usage, it is reasonable to take this as evidence for its acceptability. The negative side of this is that if a form does not occur, or occurs with extreme infrequency, in a large and representative corpus such as the BNC, conclusions as to the currency of that form can reasonably be drawn: the evidence of the corpus invites such an explanation. More generally, as in the examples above relating to *would*, more extreme indications of frequency or infrequency can add to or detract from the plausibility of an analysis. The modals are one area of grammar where great differences of frequency are observable, and where changes particularly in the direction of infrequency have recently been taking place (Leech, 2003). It is an omission, from this point of view, that H&P do not mention the often-noticed low (and declining) frequency of *may* in the sense of permission, and underplay the obsolescence of *shall* in the 'condescending' sense of *You shall have your money back tomorrow* (p. 194).[28]

Still on the modals, but turning to their negative forms, H&P note the virtual disappearance of *mayn't* and the rarity of *shan't*, but it is also worth noting the extreme rarity of *mustn't*, *mightn't*, and *needn't* in spoken AmE. In fact, I found no examples of *needn't* in the 2.5-million-word corpus of American conversation used in Biber et al. (1999). Ignoring more extreme cases of infrequency can distort a descriptive account of the language. For example, on a point about scope of negation in reference to modal *need*, H&P make the claim – scarcely tenable in view of its lack of attestation in spoken AmE – that *You needn't attend the lectures* is 'perfectly acceptable to all speakers' (p. 183). H&P can scarcely be blamed for omitting frequency findings which are only just emerging from current research, but the modals are a special case where current decline in frequency usage leads to marginalization particularly of the already less frequent modals *shall, need, ought*, and to a lesser extent *must* and *may* (Leech, 2003).

---

in *i*'. However, in the 100-million-word British National Corpus (BNC) *virtuosos* occurs only 4 times and *virtuosi* 17 times.

[27] A comparison with Q et al. on the issue of frequency will reveal that, while Q et al. give more attention to frequency than H&P, most of their frequency statements suffer from vagueness, words like 'often', 'sometimes', and 'rarely' being heavily used. An excuse for this, in the 1980s, was that corpora were less developed and grammatical frequency investigations were more difficult to conduct. This lack of reliable frequency data was remedied to a considerable extent in Biber et al. (1999), a corpus-based wide-coverage grammar using virtually the same descriptive apparatus as Q et al. In this sense, Q et al. is frequentially complemented by Biber et al.

[28] Coates (1983: 191) found no examples of this in her extensive corpus study of both written and spoken BrE.

Against this background, a grammatical description which largely ignores frequency can give a false impression of uniformity in the language system, as I argue H&P do in the case of the modals.

10   Terminology

As already hinted, this grammar carries a heavy weight of terminology, which is often new and difficult. But first, two positive things about H&P's terminology are (a) its admirable consistency of use, and (b) its language independence. By language independence, I mean that H&P, as a matter of policy, avoid terminology which is tied to the English language, and could not be used of a comparable grammatical phenomenon in another language (see p. 83 note). Some choices of new terms – such as the 'present futurate' for the use of the present tense in reference to the future – are to be welcomed, as is also the rehabilitation of 'preterite' in preference to the ambiguous use of 'past' to refer to tense, as well as to time reference.

Too often, however, the attempt to use a precise and non-language-specific terminology leads to cumbersome obscurity. The *by*-phrase with passives is given the obscure and unspecific label 'internalised complement'. Instead of the simple term *ing*-form, or *ing*-participle, H&P use the composite term 'gerund-participle'. Instead of *wh*-forms or *wh*-words, they prefer the term 'unbounded dependency words'. Other familiar terms incorporating English words, such as *if*-clauses and *that*-clauses, disappear at least in part for descriptive reasons: for H&P, *if*-clauses become phrases, and *that*-clauses are simply declarative content clauses which have been 'expanded' by the addition of *that*. Acceptably enough, instead of the *were*-subjunctive, the term 'irrealis' is used.

Some divergences between Q et al. and H&P are purely fortuitous. The concepts of 'end focus' and 'end weight' in Q et al. are well described by H&P, but without using those terms. Talking of the linear position of adjuncts, H&P use 'central' rather than Q et al.'s 'medial'. 'Preposing' is used for Q et al.'s 'fronting', although 'PP fronting' is used for the pied-piping construction. Whereas H&P use 'finite' and 'nonfinite' for clauses, they switch to 'primary' and 'secondary' as terms for Q et al.'s finite and nonfinite verbs.

In other cases, the choice of unfamiliar and sometimes verbose terminology appears to be descriptively motivated. Among such cases is the distinction between 'integrated' and 'supplementary' relative clauses – corresponding to, but conceptually different from, 'restrictive/non-restrictive' or 'defining/non-defining' clauses in Q et al. and elsewhere. The term 'phrasal verb' is not used, as all such verb+particle constructions come under H&P's definition of 'prepositional verbs'. The term 'nonassertive' in Q et al. is replaced by 'non-affirmative' (p. 60) or more lengthily by 'negatively-oriented polarity-sensitive items' (PPIs for short), which contrast with 'positive-oriented PPIs' (largely ignored by Q et al.).

11 Conclusion

These and many other differences between the terminologies of H&P and Q et al. demonstrate in no uncertain terms that these are very different ways of 'doing grammar', available to be consulted and studied by different bodies of students, teachers, and researchers. In fact, going back to the cline of grammar in figure 1, there is a diglossia in the dialect of grammatical discourse – pre-existing but accentuated by the publication of H&P's monumental volume. In fact, of course, there are many overlapping terminologies: one might argue for 'triglossia', thinking of the vast grammatical user-community of TEFL, where (for example) the progressive aspect is replaced by the 'continuous tenses'. But the divergent terminologies of H&P and Q et al. also reflect a pronounced fault-line across the cline of theory-description-observation in figure 1. In the final analysis, these two books present much of the same information, but packaged in a very different way. The serious student of grammar, although likely to be annoyed by terminological discrepancies, will continue to learn from both.

Author's address:
Department of Linguistics and Modern English Language
Lancaster University
Lancaster LA1 4YT
g.leech@lancaster.ac.uk

References

Adger, D. (2003). *Core syntax: a minimalist approach*. Oxford: Oxford University Press.

Baker, M. (2003). *Lexical categories*. Cambridge: Cambridge University Press.

Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan (1999). *Longman grammar of spoken and written English*. London: Longman.

Carter, R. & M. McCarthy (1995). Grammar and the spoken language. *Applied Linguistics* **16**: 141–58.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1970). Remarks on nominalization. In Jacobs, R. A. & P. S. Rosenbaum (eds.), *Readings in English transformational grammar*. Waltham, MA: Ginn & Co. 184–221.

Coates, J. (1983). *The semantics of the modal auxiliaries*. London: Croom Helm.

Depraetere, I. (2003). On verbal concord with collective nouns in British English. *English Language and Linguistics* **7**: 85–128.

Jespersen, O. (1909–47). *A modern English grammar on historical principles*. Copenhagen: Munksgaard.

Kachru, B. (1985). Standards, codification and sociolinguistic realism: the English language in the outer circle. In Quirk, R. & H. G. Widdowson (eds.), *English in the world*. Cambridge: Cambridge University Press. 11–30.

Leech, G. (2000). Grammars of spoken English: new outcomes of corpus-oriented research. *Language Learning* **50**: 675–724.

Leech, G. (2003). Modality on the move: the English modal auxiliaries 1961–1992. In Facchinetti, R., M. Krug & F. Palmer (eds.), *Modality in contemporary English*. Berlin: Mouton de Gruyter. 223–40.

Levin, M. (1999). Concord with collective nouns revisited. *ICAME Journal: Computers in English Linguistics* **23**: 21–34.

Prince, E. G. (1990). Syntax and discourse: a look at resumptive pronouns. In Hall, K. et al. (eds.), *Proceedings of the 16th Annual Meeting of the Berkeley Linguistics Society*. 482–92.

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik (1985). *A comprehensive grammar of the English language*. London: Longman. (Abbreviated as 'Q et al.' in this article.)

Seidlhofer, B. (2001). Closing the conceptual gap: the case for a description of English as a lingua franca. *International Journal of Applied Linguistics* **11**: 133–58.

Sinclair, J. (2001). Review of D. Biber et al., *Longman grammar of spoken and written English*. London: Longman (1999). *International Journal of Corpus Linguistics*. 339–59.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: Benjamins.

Zadeh, L. A. (1972). A fuzzy-set theoretic interpretation of linguistic hedges. *Journal of Cybernetics* **2**: 4–34.