1 Introduction: 'grammar blindness' in the recent history of English?

Surprising though this may be in view of a vast and growing body of literature on recent and ongoing changes in the language, there is very little we know about grammatical change in written standard English in the twentieth century. No one would seriously doubt that grammar constitutes a central level of linguistic structuring, and most people would agree that standard English, while being one variety among many from a purely descriptive-linguistic point of view, has nevertheless been the most studied and best documented one because of its social and cultural prominence. What, then, are the causes of this apparent 'grammar blindness'?

1.1 Grammar is more than an arbitrary list of shibboleths

Among lay commentators on linguistic change what we have is not really complete blindness but an extreme restriction of the field of vision. Rather than see grammar as the vast and complex system of rules which helps us organize words into constituents, clauses and sentences, the term is restricted to refer to a collection of variable and disputed usages which have been selected arbitrarily in the course of almost 300 years of prescriptive thinking about good grammar and proper English.

Let us illustrate this restriction of the field of vision with a first example. English has a complex and highly differentiated inventory of noun-phrase post-modification by means of relative clauses. This inventory comprises several types of finite and non-finite clauses which differ greatly in grammatical structure, in logical status (as 'restrictive' or 'non-restrictive' specification of the head) and also in stylistic connotation. All the highlighted structures listed in (1) below would be considered part of this system. The first is an authentic instance from a standard digital reference corpus of present-day written English; the others are variations on the theme:

(1) a. Interestingly, Mr John Major is acquiring a high profile as a foreign statesman *to whom more and more heads of state are willing to turn, and*

whose voice is regularly listened to in international councils. [F-LOB B06][1]

b. Mr John Major is acquiring a high profile as a foreign statesman *who(m) more and more heads of state are willing to turn to* [. . .]
c. Mr John Major is acquiring a high profile as a foreign statesman *that more and more heads of state are willing to turn to* [. . .]
d. Mr John Major is acquiring a high profile as a foreign statesman *for heads of state to turn to* [. . .]
e. Mr John Major is acquiring a high profile as a foreign statesman *to turn to* [. . .]
f. Mr John Major is acquiring a high profile as a foreign statesman *to be turned to* [. . .]

In the history of English, not all these forms are of equal age and spread, and there is no reason to assume that historical developments in this fragment of the grammar of English should have come to a halt in the twentieth century. Many interesting questions arise which might well be worth exploring. For example, we might ask whether non-finite relative clauses are spreading, possibly at the expense of finite alternatives, as this would be an expected development in view of a general tendency for non-finite clauses to become more important in the recent history of English (see, e.g., Chapter 9 of the present book and Mair 2006b: 119–140). Or we might look at the statistical or semantic relationships between active and passive infinitives in examples such as (1e) and (1f) above.

However, most discussions on recent changes in the use of relative clauses in English will instantly home in on one issue, namely the choice between *who* and *whom* as a relative pronoun in object function. Similar variability between the two forms is, of course, found in independent and dependent interrogative clauses (cf., e.g., *Who(m) did you ask?*; *I didn't know who(m) to ask*), so that – unless indicated otherwise – the following comments on *who* and *whom* can be taken to refer to both types of constructions. Usually, the issue is framed around the question of whether English is losing a traditionally 'correct' form, *whom*, and whether the resulting loss of distinction between the subject and object uses of this relative pronoun should be seen as a desirable simplification – the minority view – or as a sign of possible decay in the language.

At this stage, we do not want to anticipate the results of a detailed investigation of the use of relative clauses in present-day English, which will be offered in Chapter 10 (section 10.5) of the present book. However, we

[1] When quoting examples from standard corpora or digital databases, the usual conventions are followed. In this particular example, which is from the F-LOB (Freiburg–Lancaster–Oslo/Bergen) Corpus of written British English, 'B' refers to the textual category, in this case 'Press/ Editorial' and '06' is the number of the 2,000 word text sample the quote is taken from. Readers unfamiliar with corpus-linguistic conventions and/or the corpora used for the present study are referred to section 1.2 below and Chapter 2 for more information.

would like to use the example to point out the most important ways in which prescriptivism tends to narrow our field of vision in the study of linguistic change in progress and in some instances even promotes positions which are at odds with the facts of language history.

As for the use of *whom* in questions, the prescriptive tradition has identified the historical developments correctly in very general terms. *Who* and *whom* go back to the Old English interrogative pronouns *hwā* and *hwām*, which functioned as the nominative and dative case forms, respectively.[2] The use of uninflected *who* in object function is a historically younger development, which the *OED* (2nd edn., 1989: s.v. *who* 5) labels as ungrammatical but as 'common in colloquial use'. The same *OED* entry, however, also shows very clearly that English is not losing the form *whom* now (as is commonly alleged), but lost it in informal spoken English long ago. The first of many instances of the colloquial use given in the entry is from a letter written in 1450 (*Paston Lett.* I. 112: *I rehersyd no name, but me thowt be hem that thei wost ho I ment* 'I mentioned no name, but felt that they knew who I meant'), and the usage is attested continuously to the present day.

The facts are a little more complicated in the case of relative clauses, as both *who* and *whom* were added to the inventory of English relative pronouns relatively late, in the thirteenth and fourteenth centuries. While it is plausible to assume that the distribution of the two forms was governed by inflection and *whom* was the primary choice for objects, the historical record shows hardly a time lag between the first attestation of relative *who* in restrictive clauses (1297, *OED*, s.v. *who* 9) and the first possible case of the modern 'ungrammatical' use in a fourteenth-century work (*OED*, s.v. *who* 13).[3] Not unexpectedly, the use is attested in Shakespeare. For example, Macbeth can bewail the fall of him 'who I myself struck down' (*Macbeth*, iii.1). In view of this, it is difficult for prescriptivists to construct an argument for the historical priority of *whom* over *who* as a relative pronoun.

For both the interrogative and the relative uses it seems that the past few centuries have seen little genuine grammatical change, as the facts have been clear and stable. In all the examples below the (a) options have been the normal ones in spoken and informal English, and the (b) variants have been available as additional options in written and formal spoken English.

(2) a. *Who* did she come to see? [F-LOB P06]
    b. *Whom* did she come to see?
(3) a. 'There is Doris Jones, for instance, *who* I go away with, and Mary Plumb, and the Fosters –' [F-LOB L02]

[2] The generalized use of *whom* for all kinds of objects is a later development.
[3] 'Qua þat godd helpis wid-all, Traistli may be wend ouer-all' (= 'whom God helps . . . '). Note that *be* may be a misreading here for *he*, and that the use of the nominative might be prompted by the continuation of the sentence, in which 'the one who is helped by God' functions as subject.

b. 'There is Doris Jones, for instance, *with whom* I go away, and Mary Plumb, and the Fosters –'

This being so, any statistical shifts in usage which we might observe in twentieth-century language data would not be due to direct grammatical change. The grammar, seen as the system of rules and options underlying usage, has been very stable for the past few centuries. What might have changed, though, are stylistic conventions or expectations of formality. For example, a writer of a sports feature in a newspaper had both options available in the year 1900 as well as in 2000. If a corpus analysis were able to show late twentieth-century sportswriters to favour the informal (a) options more often than their predecessors, it would be an interesting finding – not about the evolution of the grammar of English, but about the evolution of newspaper writing style in a changing market. Of course, there is an obvious relation between style change and grammar change in the long term. If, for example, a linguistic form becomes marginal generally or across a very broad variety of genres, it will eventually disappear as an option from the structural system and either die out or live on as a fossilized expression in the lexicon.

If we are looking for clear-cut grammatical change in the use of *whom*, we have to concentrate on a very specific syntactic environment, namely the one illustrated in example (3b). Currently, the position immediately following a preposition (cf. (3b) – *with whom*) is the only one in which grammatical descriptions of present-day English regard the use of the inflected form as obligatory, and this – in addition to an occasional desire on the part of speakers and writers to sound formal and elegant – is probably what has protected it from extinction. Real grammatical change would be demonstrated if we were able to show that relative clauses of the type:

(4) 'There is Doris Jones, for instance, *with who* I go away, and Mary Plumb, and the Fosters –'

were not possible a hundred years ago, are being used now and are possibly becoming more frequent. We will return to this question in section 1.2 below.

The most heated phases in the arguments over the proper use of *who* and *whom* are, it is safe to say, behind us, and even conservative commentators on the state of the English language may have begun to acquiesce in the 'ungrammatical' use of *who* as an oblique form – much as they have got used to *it is me* instead of *it is I* or the use of *will* instead of *shall* to refer to the future with the first persons singular and plural.

However, the satisfactory conclusion that this particular debate has found does not mean that we are generally living in an enlightened age which has moved beyond such linguistic prejudice and merely needs to wonder about the curiosities of a misguided past. Even today, prescriptive rules are being enforced which are as unfounded in fact as any eighteenth-century traditional recommendation but advocated with no less vigour than their predecessors.

As it happens, a case in point is provided by another instance of variable usage in the field of relative pronouns, namely the choice between *which* and *that*. Especially in the United States, the prevailing opinion among educators and editors is that *that* is the only legitimate way of introducing a restrictive relative clause with a non-human antecedent and that *which* should not be used for this purpose. However, an unprejudiced look at historical data shows beyond doubt that *which* has not been confined to introducing non-restrictive relative clauses at any period in the history of English. In fact, it has served as a frequent alternative to *that* in restrictive relative clauses in educated usage – throughout the entire history of the English language in North America and for almost a thousand years in British English.[4] Of course, a neat one-to-one mapping of form and function – *which* for non-restrictive and *that* for restrictive post-modification, as in (5a) and (5b) below – appears tidy and makes theoretical sense (at least on the not unproblematical assumption that the logic of natural languages follows formal logic rather closely):

(5) a. Already he was asking Hemingway about his next book of stories, a book *that* Pound strongly advised against. [Frown G38]
b. Already he was asking Hemingway about "Men without Women," *which* Pound strongly advised against.

However, this distribution has never been obligatory in any variety of English past or present.[5] Instead, there is an untidy asymmetry. *That* cannot normally be used for non-restrictive post-modification, but *which* is normal in restrictive relative clauses.

(5) c. Already he was asking Hemingway about his next book of stories, a book *which* Pound strongly advised against.
d. *Already he was asking Hemingway about "Men without Women," *that* Pound strongly advised against.

Interestingly enough, American usage manuals and US editorial practice for almost a century now have been based on the fiction that a clear functional separation between *that* and *which* should exist – which is either an interesting case of a collective illusion taking hold among educated members of a speech community or a modern-day revival of the eighteenth-century impulse to bring natural language into line with logic and thus remove its perceived defects. Whatever its motivation, prescriptive teaching in this case has not been without effect: a comparison between matching British and American databases undertaken in Chapter 10 shows restrictive *which* to be seriously under-represented in American English in comparison to British English.

[4] The earliest *OED* attestations date from the twelfth century. Use of *that* as a relative pronoun is attested from Old English times.
[5] Indeed, the American Frown corpus itself contains numerous examples of restrictive *which*, for instance the following one from a – presumably professionally edited – newspaper source: 'That's the verdict which repeatedly emerges from the polls.' [Frown A10]

6 Introduction: 'grammar blindness'

Here we shall conclude by referring our readers to an instructive jeremiad on this issue in which eminent linguist Arnold Zwicky, after referring to an episode in which the 'sacred That rule' generated considerable extra income for the legal profession,[6] summarizes the many but usually futile battles he has fought in order to get instances of restrictive *which* past avid but misguided American editors.

Every so often, I've had to deal with editors from presses who are genuinely puzzled by the passion I have invested in protesting the That Rule. It's just a matter of house style, they say; it has nothing to do with syntax. You say how capitalization works, you tell people what fonts to use and how paragraphing is indicated and all that. And you tell people which subordinators to use in restrictive relative clauses. Why are YOU getting your knickers in a twist? I mean (they say), this is basically all arbitrary stipulation, the only function of which is to create and maintain consistency in the press's publications. (Some writers, like Louis Menand, even revel in arbitrary 'rules' for their own sake.)

Twice, my aggressive truculence about the That Rule (and a collection of other zombie rules) has prompted editors to cave in to my craziness and let me do whatever I want. Me. Not anyone else, just me, for this one book. They were then baffled that I didn't view this response as really satisfactory. I pointed out that the scholarly books their firms published on English grammar uniformly failed to subscribe to the That Rule, so that their presses looked like packs of hypocrites and fools. They simply didn't get it. For them, one thing is scholarship, the other thing is practice. They're just different. ('Language Log', posting by Arnold Zwicky at 22 May 2005; http://itre.cis.upenn.edu/~myl/languagelog/archives/002291.html#more)

In this connection it is interesting to note that a major recent reference grammar of English explicitly condemns this ill-founded rule in one of its 'prescriptive grammar notes' (Huddleston and Pullum 2005: 191), which are otherwise devoted to more traditional shibboleths such as the use of 'singular' *they*, the split infinitive or the choice between *I* and *me*.

[6] Zwicky points to a disturbing legal case in which the perfectly obvious meaning of a sentence was turned into its opposite in court: 'The Texas statute furthers no legitimate state interest which can justify its intrusions into the personal and private life of the individual'. [US Supreme Court, Lawrence v. Texas] In debating technicalities of a complex judgement, legal experts seriously, and in print, appealed to the 'That rule' to support their reading of the *which*-clause as non-restrictive – never even minding the fact that, as Zwicky points out, a non-restrictive reading is not even possible in this example because 'no legitimate interest' is not a referential noun phrase. The possibility of a completely absurd misinterpretation of the statement, with *which* introducing a sentential relative (with a paraphrase such as, roughly, 'The Texas statute furthers no legitimate state interest, and this can justify its intrusions into the personal and private lives of the individual') was fortunately never pursued.

1.2 Grammatical changes: proceeding slowly and invisible at close range?

Grammar is probably the level on which the English language has changed most radically in the course of its recorded history, and this is noted in treatments of Old and Middle English. By contrast, studies of change in the more recent past generally place much more emphasis on phonological and lexical phenomena than on grammatical ones (cf., e.g., the small number of pages devoted to grammar in standard treatments of changes in present-day English such as Barber 1964, Foster 1968 or Potter 1975). Barbara Strang, herself the author of a classic history of the English language, has noted this imbalance, arguing that it is most likely not rooted in the facts of language history but in our ability to perceive and analyse them:

> One possible explanation can hardly be proved false, but should be entertained only as a last resort: namely, that although there has been considerable grammatical change in the past, English grammar in our own lifetime is somehow uniquely stable and free from change.
>
> The most promising direction of search for an explanation would seem to lie in the assumption that there is grammatical change in progress at the moment, as in the past, but that we are considerably less perceptive of it than of other kinds of linguistic change. (1970: 59–60)

What is it that makes grammatical change difficult to perceive? For a lay observer, especially in a language such as English with its largely analytical grammar, part of the difficulty may lie in the fact that so little of the grammar is audible/visible directly – for example in the form of inflectional endings on words – and so much of it is abstract, involving, for example, the position of elements in a clause relative to each other or, as in the case of re-analysis, the development of a new underlying form for an established surface sequence. Thus – to take an instance of a simple 'visible' change – it does not take a degree in linguistics to note that the plural of *postman* remains irregular (*postmen*) in present-day English, while the plural of *Walkman* tends to be *Walkmans*.

The following example, by contrast, raises a few more complicated issues about the status of *following*:

(6) *Following* the signing of the peace treaty and British recognition of American independence, Washington stunned the world when he surrendered his sword to Congress on Dec. 23 1783 and retired to his farm at Mount Vernon. [Frown G13]

Following looks like a present participle, and indeed similar constructions would make decent enough non-finite adverbial clauses in many syntactic contexts, for example in *Following the suspicious stranger, they ended up in a rather unpleasant part of town*. Such an analysis, however, is not available

here, and at least in this example and similar ones *following* is therefore most appropriately analysed as a deverbal preposition roughly equivalent to *after*. The gradual expansion of some participles into the prepositional domain is by no means a unique phenomenon, but illustrates a well-trodden path of grammaticalization. Earlier instances from the history of English include *regarding*, *concerning*, *barring* or even *during* and *notwithstanding*, and similar phenomena are common in other languages. However, the long time taken by such shifts, their gradual nature, the involvement of abstract grammatical categories rather than concrete words and morphemes, and not least the structural ambiguity of many relevant examples all make it very difficult for lay observers to spot such changes and to make explicit the linguistic processes involved.[7]

For lay and expert observers alike, an additional difficulty in perceiving grammatical change, in particular grammatical change at close range, is that it generally proceeds more slowly than lexical and phonetic change. While a lifetime devoted to observing lexical or phonetic developments in English will generally be enough to arrive at a fair number of definitive conclusions, the same timespan is insufficient to allow testable statements about the direction and speed of grammatical trends. For grammatical changes, therefore, even linguistically trained observers will need more solid orientation than their own necessarily subjective and partial observations provide. As David Denison has made clear in his magisterial study of grammatical change in nineteenth- and twentieth-century English, practically all grammatical change involves a gradual and statistical element during the long process in which an innovation establishes itself in the community of speakers (or, conversely, a formerly common but now obsolescent form is phased out):

> Since relatively few categorial losses or innovations have occurred in the last two centuries, syntactic change has more often been statistical in nature, with a given construction occurring throughout the period and either becoming more or less common generally or in particular registers. The overall, rather elusive effect can seem more a matter of stylistic than of syntactic change, so it is useful to be able to track frequencies of occurrence from EModE through to the present day. (Denison 1998: 93)

[7] Minimally, the person would have to have the metalinguistic competence necessary to conduct standard linguistic re-formulation tests and interpret their results. For example, an analysis of *following* as a verbal participle is unlikely because the construction cannot be expanded into a finite adverbial clause which shares its subject with the main clause (in this case 'Washington'):

> *When he *followed* the signing of the peace treaty and British recognition of American independence, Washington stunned the world when he surrendered his sword to Congress on Dec. 23, 1783 and retired to his farm at Mount Vernon.

For a more detailed analysis of this particular instance of grammaticalization, see Olofsson (1990).

In view of this, there is no way around the systematic compilation of statistics and frequencies which are based on large machine-readable bodies of textual data.

The present work is thus based on the following three premises, namely that (1) the systematic study of such corpora will refine our understanding of recent and ongoing grammatical change in standard English, that (2) such research will help us to correct current misperceptions and that (3) the method will occasionally point us towards interesting developments in the language which have not even been noticed before.

The corpora used for the present study are first and foremost the four matching one-million-word corpora of British and American English known as the 'Brown family' (after the pioneering Brown corpus which set the pattern for many similar ones subsequently compiled). The Brown corpus, named after Brown University in Providence, Rhode Island, where it was compiled by W. Nelson Francis and Henry Kučera in the 1960s, is – as its official title describes it – a 'Standard Corpus of Present-Day Edited American English, for Use with Digital Computers'. It contains about a million words of text, sampled in 500 extracts of c. 2,000 words each spanning a range of 15 different textual genres, and representing the state of written American English in the year 1961.[8] The LOB (Lancaster–Oslo/Bergen) corpus was compiled under the direction of Geoffrey Leech and Stig Johansson in the 1970s to provide a matching database for British English. In the 1990s, F-LOB (the Freiburg update of the LOB corpus) and Frown (Freiburg update of the Brown corpus) were compiled under the direction of Christian Mair at the University of Freiburg, in order to bring the comparison of British and American English closer to the present and, even more importantly, to make possible the systematic corpus-based study of how regional variation interacts with short-term diachronic change. The 'Brown family' of corpora has spawned a considerable amount of research on grammatical change in progress in present-day English, both by the authors of the present book and by others. Most of this research has been based on the plain-text versions of the corpora, with the obvious limitations on linguistically sophisticated access to the material that such a restriction entails.

However, the present book is not merely a continuation and summary of previous research, but represents a new departure in at least two respects. First, it is now possible to complement research on the plain-text corpora with investigations of versions of the corpora which have been grammatically annotated for parts of speech. As will be shown, this opens up interesting possibilities of accessing the material in novel ways, and studying aspects of ongoing grammatical change which have never been covered before. To give an illustration: a study of inflectional and analytical comparison of

[8] See the Preface, Chapter 2 and Appendix I for further information on this corpus and other corpora used for the present study.

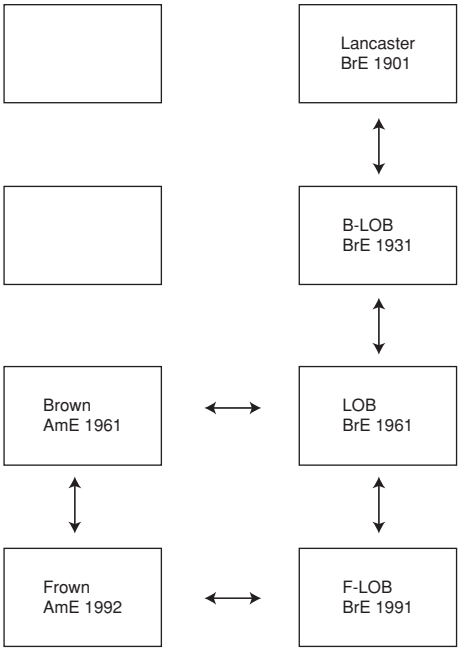

Figure 1.1 Matching one-million-word corpora of written English

adjectives (see section 11.6.3) based on untagged corpora is confined to searching for individual pairs such as *politer* vs *more polite*, or *commoner* vs *more common*. It would not be possible to search for the category 'inflectionally graded adjective' as a whole, nor would we be able to determine the share of comparative and superlative forms as a proportion of the total number of adjectives (i.e. the forms which are potential carriers of the marking investigated). In other words, we would be almost certain to miss many important generalizations about ongoing change in this fragment of the grammar. Second, work on the Brown family of corpora was often hampered by the fact that in a timespan of a mere thirty years it is difficult to differentiate directed diachronic developments from random fluctuation. To remedy this, the two UK-based authors of the present book have started work on compiling matching corpora documenting the development of British English in 1931 ('B-LOB,' for 'before LOB') and in 1901 ('Lancaster 1901').

The relationship between these corpora is visually represented in Figure 1.1.

As the blanks in the 'American' half of the diagram show, the symmetry is not perfect yet, and much of Lancaster 1901 remains to be completed. Nevertheless, the corpus-linguistic working environment illustrated generally makes it possible to sketch the development of high- and medium-frequency