John Algeo

British or American English?

A handbook of word and grammar patterns
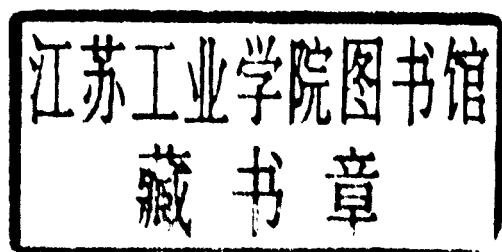
CAMBRIDGE

British or American English?

A Handbook of Word and
Grammar Patterns

JOHN ALGEO
University of Georgia

CAMBRIDGE
UNIVERSITY PRESS

Preface

The study on which this book is based began about forty years ago as a casual interest in the subject engendered by Thomas Pyles's history textbook, *The Origins and Development of the English Language* (now in its fifth revised edition, Algeo and Pyles 2004). It was focused during a year (1986–7) the author spent in the Survey of English Usage at University College London as a Fulbright Senior Research Scholar and a Guggenheim Fellow. In those days, the Survey was only beginning to be converted into electronic form, so at first research involved hunting through paper slips and copying information by hand. Later, as the Survey was computerized, electronic searches became possible, initially only at the Survey office and later through a CD anywhere.

The present study later benefited from the collection of citations made by Allen Walker Read for a historical dictionary of British lexical items. My wife, Adele, and I then set out to supplement Read's files with citations we collected from more recent material than he had used, including citations for grammatical as well as lexical matters. Our own corpus of British citations is now about three million words in size. That is not large for a contemporary data file, but it consists entirely of citations that we had reason to suspect exemplified British use.

Work on this book was delayed by a variety of other duties to which its author had fallen heir. It is now presented, with painful awareness of its limitations, but, as the French are fond of saying, *faute de mieux*. Undoubtedly, British and American English are grammatically different in ways not reported here. And some of the grammatical differences reported here may be less certain than this book suggests because of difficulties in identifying and substantiating those differences or because of the misapprehension of the author. Nevertheless, I hope that it will be helpful in pinpointing various areas of structural difference between the two major national varieties of the language.

Acknowledgments

Contents

Introduction

British and American as national varieties

There are many varieties of English other than British (here the English of the United Kingdom) and American (here the English of the United States). All of those other varieties are intrinsically just as worthy of study and use as British and American. But these two varieties are the ones spoken by most native speakers of English and studied by most foreign learners. They have a special status as the two principal national varieties of the language simply because there is more material available in them than in any other variety.

British is the form of English now used in the country whence all other forms of English have ultimately derived. But present-day British is not the origin of any other variety of the language; rather it and all the other varieties are equally descendant from a form of English spoken in the British Isles in earlier times. In some respects, present-day British is closer to the common ancestral form of the present-day varieties than is American or other varieties; but in other respects the reverse is true, and American, for instance, preserves older uses that became obsolete in British use. To mistake present-day British for the ancestor of all other forms of English is a logical and factual error.

The focus of this study is on how contemporary British English differs from American. That is, in comparing two varieties of a language, it is convenient to take one as the basis for comparison and to describe the other by contrast with it. This study takes American as its basis and describes British in relation to that basis. The reason for this approach is that American has more native speakers than British and is rapidly becoming the dominant form of English in non-native countries other perhaps than those of Western Europe. Much European established academic bias favors British as a model; but evolving popular culture is biased toward American. This widespread dissemination of the American variety makes it a reasonable basis for describing British.

Differences between British and American

The most obvious difference between British and American is in the "tune" of the language, that is, the intonation that accompanies sentences. When a Briton or an American talks, they identify themselves primarily by the tunes of their respective varieties. In singing, the prose tune is overridden by the musical tune, making it much harder to distinguish British and American singers.

Other pronunciation differences exist in stress patterns and in consonant and vowel articulation and distribution. Those differences have been described in fine detail. Vocabulary differences have been very widely noted between the two varieties, and they are fairly extensive, although also often subtler than most lists of supposed equivalences account for. Popular awareness probably centers more on lexical differences than on any other sort, partly perhaps because they are the easiest for the layperson to notice. Subtle differences of national style also exist, but have been but little and only incidentally noted (Algeo 1989, Heacock and Cassidy 1998).

Grammatical differences have been treated, but mainly by individual scholarly studies focused on particular grammatical matters. Extensive and comprehensive treatment is rare. Popular writers on grammar are aware that British and American differ in their morphosyntax but tend to be sketchy about the details. Anthony Burgess (1992), who is one of the linguistically best informed men of letters, settled on a few verb forms as illustrations. The grammatical differences between the two principal national varieties of the language are, however, manifold. Some general treatments of British-American grammatical differences, from various standpoints, are those by Randolph Quirk et al. (1985), John Algeo (1988), Michael Swan (1995), Douglas Biber et al. (1999), Rodney Huddleston and Geoffrey Pullum (2002), Gunnel Tottie (2002, 146–78), Peter Trudgill and Jean Hannah (2002), and Pam Peters (2004).

Although many, few of the grammatical differences between British and American are great enough to produce confusion, and most are not stable because the two varieties are constantly influencing each other, with borrowing both ways across the Atlantic and nowadays via the Internet. When a use is said to be British, that statement does not necessarily mean that it is the only or even the main British use or that the use does not occur in American also, but only that the use is attested in British sources and is more typical of British than of American English.

The basis of this study

A distinction is often drawn between intuition and data as the basis for statements about language. That dichotomy, like most others, is false. Intuition is needed to identify matters to comment on, and data is (or, as the reader prefers, are) needed to substantiate intuition. My wife and I have spent twenty years

gathering citations of what intuition told us were British uses. Then I set out to substantiate those intuitions by consulting corpora of data. In most cases, our intuitions proved correct, and the corpora yielded statistics to support our hunches. In some cases, however, what intuition told us was a Briticism turned out to be nothing of the sort, but instead just to be a rare or peculiar use – rare and peculiar in both British and American English. And in a few cases, we were spectacularly wrong. Linguistic intuition is invaluable but unreliable.

Corpus data is likewise invaluable, but it has its own unreliability. The statistics from any corpus should be used with care and reservations, especially in comparing statistics from different corpora or even statistics derived from the same corpus but in different ways. A bit of folk wisdom has it that there are three kinds of lies: lies, damned lies, and statistics. The problems with statistics based on language corpora include the fact that two corpora may not be comparable because they are of different sizes or because they are composed of different kinds of texts. Academic printed texts and conversational oral texts will have strikingly different characteristics.

The way one phrases a search in a corpus can also produce different results; for example, if the search engine is sensitive to capitalization, asking for examples and statistics of a form with a lower-case initial letter may produce rather different results than a query asking for the same information of the same form, but with an upper-case initial letter. In this study, capitalization was taken into consideration when it seemed potentially influential, but not otherwise.

Moreover, many grammatical items are difficult to find in a corpus unless it has been extensively and accurately tagged, and few corpora, especially the larger ones, have the sort of tagging that would make grammatical searches easy. Instead, one must come up with ways of asking the corpus about instances of something that its search engine can find and that will give at least implicit, albeit incomplete, information about grammatical structures. Thus if one wants information about the form of negation in sentences with indefinite direct objects (*They had no money*) versus those with definite direct objects (*They didn't have the money needed*), barring sophisticated grammatical tagging, it is necessary to ask about particular constructions (such as those just cited) and extrapolate a generalization from them. This study generally eschews such broad extrapolation, but some was unavoidable.

Finally, however, one relies on whatever is available. For the entries in this study, such evidence as was convenient to extract from corpora has been cited. But when that evidence was not readily available, intuition was still used. Any entry with no substantiating evidence is an intuitional guess, as far as its Britishness is concerned. In those, as well as other, cases it is advisable to keep in mind the wise words of Oliver Cromwell to the General Assembly of the Church of Scotland: "I beseech you, in the bowels of Christ, think it possible you may be mistaken." The author intones those words as a mantra.

Sources of comparative statistics and citations

Statistics

In the body of this work, several corpora have been used and are cited by name, but the one most used, especially for comparative statistics, is the Cambridge International Corpus (CIC). Statistics from it are sometimes cited as ratios or percentages; in those cases, the base number is of a size to make such form of citation appropriate and easy to follow. CIC statistics are also sometimes cited by an arcane abbreviation: "iptmw," that is, "instances per ten million words," which is the way the CIC reports frequencies from its nearly two hundred million words. The accompanying table shows the composition of this great corpus and the relative sizes of its component parts. As can be seen, the British corpus totals 101.9 million words, of which 83 percent are written texts and 17 percent spoken texts; the American corpus totals 96.1 million words, of which 77 percent are written texts and 23 percent spoken texts.

CAMBRIDGE INTERNATIONAL CORPUS				
corpus group	*corpus name*	*million words*	*number of cites*	*contents*
British written	BRNEWS25	25.0	60224	mixed newspapers 1988 – June 2000
	BRWRIT2	25.4	26915	fiction, nonfiction & magazines etc.
	BNCWRIT1	25.1	901	British National Corpus part 1 (1979–1994)
	ACAD_BR	9.2	1260	British academic journals & nonfiction
		84.7		
British spoken	BRSPOK2	7.1	1652	spoken (lexicography) incl. Cancode/Brtrans
	BNCSPOK	10.1	911	British National Corpus spoken (1980–1994)
		17.2		
American written	AMNEWS25	25.0	45026	mixed newspapers 1979–1998
	AMNW01_2	22.0	23042	newspapers 2001
	AMWRIT2	23.8	28453	fiction, nonfiction & magazines etc.
	ACAD_AM	3.6	41	American academic journals & nonfiction
		74.4		
American spoken	AMLEXI	6.2	764	spoken (lexicography) incl. Naec/Amspok
	AMSPPROF	1.9	17	spoken professional (lexicography)
	AMTV	13.6	60881	TV & radio (lexicography & research)
		21.7		

In consulting the CIC, all textual categories were weighted equally, even though only 17 percent of British texts and 23 percent of American texts are spoken versus written, and 11 percent of British written texts and 5 percent of American written texts are academic versus general. That equal weighting emphasizes disproportionately the fewer spoken over written texts and academic over general writing. Different weightings would very likely have produced at least somewhat different results.

Because the focus of this study is not on speech versus writing or academic versus general style, and because British and American are treated alike in this respect, ignoring the differences in text types probably does not greatly affect the general conclusions concerning British versus American use. Thus a statement such as "*daren't* is 13.9 times more frequent in British than in American" refers to a combination of spoken and written texts in both varieties, although it is in the nature of things that contractions are more frequent in speech than in writing. That, however, is not the concern of this study.

The CIC is especially useful for a statistical comparison of British and American because of its large size and because it has roughly comparable samples of British and American texts. As mentioned above, statistics from it are often cited in terms of "instances per ten million words" (iptmw). When some form or construction is cited as occurring X times more or less often in one variety than in the other, or in percentages, the basis for that comparison seemed adequate, and that style of comparison easier to understand.

Citations

In keeping with the focus on British English mentioned above, all of the illustrative citations are of British use. Most of them are drawn from a corpus of British examples compiled by Adele and John Algeo over a period of some twenty years. That corpus consists of British citations gathered because they were suspected to contain characteristically British features, chiefly lexical but also some grammatical ones. Most of the citations are from newspapers or popular fiction. The corpus is stored electronically in word–processor format.

Illustrative quotations are generally limited to one for each entry. In many cases the files that underlie this study contain a great many more, but space was not available for them. Several of the chapters depend heavily on prior studies by the author and draw both examples and exposition from articles reporting those studies.

The sources cited are heavily in the genre of mystery novels and other light fiction, chosen because the initial reading was for lexical purposes, and those genres have a rich store of colloquialisms and informal language (in which British-American differences are most pronounced) whereas serious fiction contains fewer such items.

British fiction that has been adapted for American readers provides a useful source to document the words and expressions that publishers change for the American market. In the case of the *Harry Potter* books, a website (www.hp-lexicon.org/) provides a list of such changes. Quotations from these books in this work note the American adaptation when it was recorded on that site.

Many of the quotations cited here were computerized by graduate assistants at the University of Georgia. They sometimes made mistakes in transcribing a quotation that suggest the quotation's use was at variance with their own native

use; such mistakes are occasionally noted as evidence for the Britishness of a particular form.

Examples cited from publicly available corpora are identified appropriately. Those cited from the Survey of English Usage (SEU) have corpus identification numbers preceded by either "s" for spoken or "w" for written.

Conventions and organization of this study

Illustrative quotations are abridged when that can be done without distortion or losing needed context. Matter omitted in the middle of a quotation is indicated by ellipsis points; matter omitted at the beginning of a quotation is indicated only if the retained matter does not begin with a capital letter; matter omitted at the end of a quotation is not indicated.

In the illustrative quotations, periodical headlines have arbitrarily been printed with initial capital letters for each word, as a device to facilitate their recognition.

The abbreviation "iptmw," which is widely used, has been explained above as meaning "instances per ten million words" in the CIC texts. An asterisk before a construction (as in *go sane*) means that the construction is impossible in normal use. A question mark before a construction (as in ?*They dared their friends solve the puzzle*) means that the construction is of doubtful or disputed possibility in normal use. Cross-references from one chapter to another use the symbol §; thus § 2.2.2.3 means "chapter 2 section 2.2.3". Abbreviations of titles of dictionaries, grammars, and corpora are explained in the bibliographies of scholarly works and of citation sources.

Studies and dictionaries are cited either by title abbreviations (e.g., *CGEL*), which are identified in the bibliography, or by author and year (e.g., Peters 2004). Citation sources are cited by date and author (e.g., 1977 Dexter) and short title, if necessary (e.g., 1937 Innes, *Hamlet*) or by periodical date and title (e.g., 2003 June 12 *Times* 20/2; for location in a periodical, "2 4/2–3" means "section 2, page 4, columns 2 to 3").

In headwords and glosses to them, general terms representing contextual elements are italicized, e.g., **pressurize** *someone* means that the verb *pressurize* takes a personal object.

A comment that a construction is "rare" means that the Algeo corpus contains few examples, often only one, and that CIC has no or very few instances of it. Such constructions are included because they illustrate a pattern. The term "common-core English" designates usage common to the two varieties, British and American, and not differing significantly between them.

Of the seventeen following chapters, the first ten deal with parts of speech, and the final seven with matters of syntax or phrase and clause constructions. Because the verb is central to English grammatical constructions, it is considered in Chapter 1. Thereafter, the elements of the noun phrase are taken up: determiners, nouns, pronouns, and adjectives. Adverbs and qualifiers (i.e., adverbs of degree) follow, succeeded by prepositions and conjunctions, with the highly

miscellaneous category of interjections coming last in the chapters on parts of speech.

In the chapters on syntactic constructions, no effort is made to treat all matters of English syntax, most of which vary little between British and American use. Instead, chapters have been devoted to those relatively few syntactic matters that do show significant differences between the two national varieties: complementation (*agree [on] a plan*), mandative constructions (*insisted he was/be there*), expanded predicates (*have/take a bath*), concord (*the team have/has won*), propredicates (*I haven't finished but I could [do]*), tag questions (*he would, wouldn't he?*), and other constructions, such as focusing (*it's right tasty, is Webster's*).

I

Parts of Speech