

THE
MOTHER
TONGUE

ENGLISH &

HOW IT GOT THAT WAY

B I L L B R Y S O N

W I L L I A M M O R R O W

A N D C O M P A N Y , I N C .

N E W Y O R K

10.

ORDER OUT OF CHAOS

HOW BIG IS THE ENGLISH LANGUAGE? That's not an easy question. Samuel Johnson's dictionary contained 43,000 words. The unabridged *Random House* of 1987 has 315,000. *Webster's Third New International* of 1961 contains 450,000. And the revised *Oxford English Dictionary* of 1989 has 615,000 entries. But in fact this only begins to hint at the total.

For one thing, meanings in English are much more various than a bald count of entry words would indicate. The mouse that scurries across your kitchen floor and the mouse that activates your personal computer clearly are two quite separate entities. Shouldn't they then be counted as two words? And then what about related forms like *mousy*, *mouselike*, and *mice*? Shouldn't they also count as separate words? Surely there is a large difference between something that is a mouse and something that is merely mousy.

And then of course there are all the names of flora and fauna, medical conditions, chemical substances,* laws of physics, and all the other scientific and technical terms that don't make it into ordinary dictionaries. Of insects alone, there are 1.4 million named species. Total all these together and you have—well, no one knows. But certainly not less than three million.

So how many of these words do we know? Again, there is no simple answer. Many scholars have taken the trouble (or more

* One of which, incidentally, is said to be the longest word in the English language. It begins *methianylglutaminyl* and finishes 1,913 letters later as *alynalanylthreonilarginyls-erase*. I don't know what it is used for, though I daresay it would take some rubbing to get it out of the carpet.

probably compelled their graduate students to take the trouble) of counting the number of words used by various authors, on the assumption, one supposes, that that tells us something about human vocabulary. Mostly what it tells us is that academics aren't very good at counting. Shakespeare, according to Pei and McCrum, had a vocabulary of 30,000 words, though Pei acknowledges seeing estimates putting the figure as low as 16,000. Lincoln Barnett puts it at 20,000 to 25,000. But most other authorities—Shipley, Baugh and Cable, Howard—put the number at a reassuringly precise 17,677. The King James Bible, according to Laird, contains 8,000 words, but Shipley puts the number at 7,000, while Barnett confidently zeroes in on a figure of 10,442. Who knows who's right?

One glaring problem with even the most scrupulous tabulation is that the total number of words used by an author doesn't begin to tell us the true size of his vocabulary. I know the meanings of *frangible*, *spiffing*, and *cutesy-poo*, but have never had occasion to write them before now. A man of Shakespeare's linguistic versatility must have possessed thousands of words that he never used because he didn't like or require them. Not once in his plays can you find the words *Bible*, *Trinity*, or *Holy Ghost*, and yet that is not to suggest that he was not familiar with them.

Estimates of the size of the average person's vocabulary are even more contentious. Max Müller, a leading German philologist at the turn of the century, thought the average farm laborer had an everyday vocabulary of no more than 300 words. Pei cites an English study of fruit pickers, which put the number at no more than 500, though he himself thought that the figure was probably closer to 30,000. Stuart Berg Flexner, the noted American lexicographer, suggests that the average well-read person has a vocabulary of about 20,000 words and probably uses about 1,500 to 2,000 in a normal week's conversation. McCrum puts an educated person's vocabulary at about 15,000.

There are endless difficulties attached to adjudging how many words a person knows. Consider just one. If I ask you what *incongruent* means and you say, "It means not congruent," you are correct. That is the first definition given in most dictionaries, but

that isn't to say that you have the faintest idea what the word means. Every page of the dictionary contains words we may not have encountered before—*inflationist*, *forbiddance*, *moosewood*, *pulsative*—and yet whose meanings we could very probably guess.

At the same time there are many words that we use every day and clearly know and yet might have difficulty proving. How would you define *the* or *what* or *am* or *very*? Imagine trying to explain to a Martian in a concise way just what *is* is. And then what about all those words with a variety of meanings? Take *step*. *The American Heritage Dictionary* lists a dozen common meanings for the word, ranging from the act of putting one foot in front of the other to the name for part of a staircase. We all know all these meanings, yet if I gave you a pencil and a blank sheet of paper could you list them? Almost certainly not. The simple fact is that it is hard to remember what we remember, so to speak. Put another way, our memory is a highly fickle thing. Dr. Alan Baddeley, a British authority on memory, cites a study in which people were asked to name the capital cities of several countries. Most had trouble with the capitals of countries like Uruguay and Bulgaria, but when they were told the initial letter of the capital city, they often suddenly remembered and their success rate soared. In another study people were shown long lists of random words and then asked to write down as many of them as they could remember. A few hours later, without being shown the list again, they were asked to write down as many of the words as they could remember then. Almost always the number of words would be nearly identical, but the actual words recalled from one test to another would vary by 50 percent or more. In other words, there is vastly more verbal information locked away in our craniums than we can get out at any one time. So the problem of trying to assess accurately just how much verbal material we possess in total is fraught with difficulties.

For this reason educational psychologists have tended to shy away from such studies, and such information as exists is often decades old. One of the most famous studies was conducted in 1940. In it, two American researchers, R. H. Seashore and L. D. Eckerson, selected a random word from each left-hand page of a Funk & Wagnalls standard desktop dictionary and asked a sam-

pling of college students to define those words or use them in a sentence. By extrapolating those results onto the number of entries in the dictionary, they concluded that the average student had a vocabulary of about 150,000 words—obviously very much larger than previously supposed. A similar study carried out by K. C. Diller in 1978, cited by Aitchison in *Words in the Mind*, put the vocabulary level even higher—at about 250,000 words. On the other hand, Jespersen cites the case of a certain Professor E. S. Holden who early in the century laboriously tested himself on every single word in *Webster's Dictionary* and arrived at a total of just 33,456 known words. It is clearly unlikely that a university professor's vocabulary would be four to six times smaller than that of the average student. So such studies would seem to tell us more about the difficulties of framing tests than about the size of our vocabularies.

What is certain is that the number of words we use is very much smaller than the number of words we know. In 1923 a lexicographer named G. H. McKnight did a comprehensive study of how words are used and found that just forty-three words account for fully half of all the words in common use, and that just nine account for fully one quarter of all the words in almost any sample of written English. Those nine are: *and, be, have, it, of, the, to, will, and you.*

By virtue of their brevity, dictionary definitions often fail to convey the nuances of English. *Rank* and *rancid* mean roughly the same thing, but, as Aitchison notes, we would never talk about eating rank butter or wearing rancid socks. A dictionary will tell you that *tall* and *high* mean much the same thing, but it won't explain to you that while you can apply either term to a building you can apply only tall to a person. On the strength of dictionary definitions alone a foreign visitor to your home could be excused for telling you that you have an abnormal child, that your wife's cooking is exceedingly odorous, and that your speech at a recent sales conference was laughable, and intend nothing but the warmest praise.

The fact is that the real meanings are often far more complex than the simple dictionary definitions would lead us to suppose. In

1985, the department of English at the University of Birmingham in England ran a computer analysis of words as they are actually used in English and came up with some surprising results. The primary dictionary meaning of words was often far adrift from the sense in which they were actually used. *Keep*, for instance, is usually defined as to retain, but in fact the word is much more often employed in the sense of continuing, as in “keep cool” and “keep smiling.” *See* is only rarely required in the sense of utilizing one’s eyes, but much more often used to express the idea of knowing, as in “I see what you mean.” *Give*, even more interestingly, is most often used, to quote the researchers, as “mere verbal padding,” as in “give it a look” or “give a report.” [London *Sunday Times*, March 31, 1985]

In short, dictionaries may be said to contain a certain number of definitions, but the true number of meanings contained in those definitions will always be much higher. As the lexicographer J. Ayto put it: “The world’s largest data bank of examples in context is dwarfed by the collection we all carry around subconsciously in our heads.”

English is changing all the time and at an increasingly dizzy pace. At the turn of the century words were being added at the rate of about 1,000 a year. Now, according to a report in *The New York Times* [April 3, 1989], the increase is closer to 15,000 to 20,000 a year. In 1987, when Random House produced the second edition of its masterly twelve-pound unabridged dictionary, it included over 50,000 words that had not existed twenty-one years earlier and 75,000 new definitions of old words. Of its 315,000 entries, 210,000 had to be revised. That is a phenomenal amount of change in just two decades. The new entries included *preppy*, *quark*, *flexitime*, *chairperson*, *sunblocker*, and the names of 800 foods that had not existed or been generally heard of in 1966—*tofu*, *piña colada*, *chapati*, *sushi*, and even *crêpes*.

Unabridged dictionaries have about them a stern, immutable air, as if here the language has been captured once and for all, and yet from the day of publication they are inescapably out of date. Samuel Johnson recognized this when he wrote: “No dictionary of a living tongue can ever be perfect, since while it is hastening to

publication, some words are budding, and some are fading away." That, however, has never stopped anyone from trying, not least Johnson himself.

The English-speaking world has the finest dictionaries, a somewhat curious fact when you consider that we have never formalized the business of compiling them. From the seventeenth century when Cardinal Richelieu founded the Académie Française, dictionary making has been earnest work indeed. In the English-speaking world, the early dictionaries were almost always the work of one man rather than a ponderous committee of academics, as was the pattern on the Continent. In a kind of instinctive recognition of the mongrel, independent, idiosyncratic genius of the English tongue, these dictionaries were often entrusted to people bearing those very characteristics themselves. Nowhere was this more gloriously true than in the person of the greatest lexicographer of them all, Samuel Johnson.

Johnson, who lived from 1709 to 1784, was an odd candidate for genius. Blind in one eye, corpulent, incompletely educated, by all accounts coarse in manner, he was an obscure scribbler from an impoverished provincial background when he was given a contract by the London publisher Robert Dodsley to compile a dictionary of English.

Johnson's was by no means the first dictionary in English. From *Cawdrey's Table Alphabeticall* in 1604 to his opus a century and a half later there were at least a dozen popular dictionaries, though many of these were either highly specialized or slight (*Cawdrey's Table Alphabeticall* contained just 3,000 words and ran to barely a hundred pages). Many also had little claim to scholarship. *Cawdrey's*, for all the credit it gets as the first dictionary, was a fairly sloppy enterprise. It gave the definition of *aberration* twice and failed to alphabetize correctly on other words.

The first dictionary to aim for anything like comprehensiveness was the *Universal Etymological Dictionary* by Nathaniel Bailey, published in 1721, which anticipated Johnson's classic volume by thirty-four years and actually defined more words. So why is it that Johnson's dictionary is the one we remember? That's harder to answer than you might think.

His dictionary was full of shortcomings. He allowed many spelling inconsistencies to be perpetuated—*deceit* but *receipt*, *deign* but *disdain*, *hark* but *hearken*, *convey* but *inveigh*, *moveable* but *immovable*. He wrote *downhil* with one *l*, but *uphill* with two; *install* with two *l*'s, but *reinstal* with one; *fancy* with an *f*, but *phantom* with a *ph*. Generally he was aware of these inconsistencies, but felt that in many cases the inconsistent spellings were already too well established to tamper with. He did try to make spelling somewhat more sensible, institutionalizing the differences between *flower* and *flour* and between *metal* and *mettle*—but essentially he saw his job as recording English spelling as it stood in his day, not changing it. This was in sharp contrast to the attitude taken by the revisers of the Académie Française dictionary a decade or so later, who would revise almost a quarter of French spellings.

There were holes in Johnson's erudition. He professed a preference for what he conceived to be Saxon spellings for words like *music*, *critic*, and *prosaic*, and thus spelled them with a final *k*, when in fact they were all borrowed from Latin. He was given to flights of editorializing, as when he defined a *patron* as "one who supports with insolence, and is paid with flattery" or *oats* as a grain that sustained horses in England and people in Scotland. His etymologies, according to Baugh and Cable, were "often ludicrous" and his proofreading sometimes strikingly careless. He defined a *garret* as a "room on the highest floor in the house" and a *cockloft* as "the room over the garret." Elsewhere, he gave identical definitions to *leeward* and *windward*, even though they are quite obviously opposites.

Even allowing for the inflated prose of his day, he had a tendency to write passages of remarkable denseness, as here: "The proverbial oracles of our parsimonious ancestors have informed us, that the fatal waste of our fortune is by small expenses, by the profusion of sums too little singly to alarm our caution, and which we never suffer ourselves to consider together." *Too little singly?* I would wager good money that that sentence was as puzzling to his contemporaries as it is to us. And yet at least it has the virtue of relative brevity. Often Johnson constructed sentences that ran to 250 words or more, which sound today uncomfortably like the

ramblings of a man who has sat up far too late and drunk rather too much port.

Yet for all that, his *Dictionary of the English Language*, published in two volumes in June 1755, is a masterpiece, one of the landmarks of English literature. Its definitions are supremely concise, its erudition magnificent, if not entirely flawless. Without a nearby library to draw on, and with appallingly little financial backing (his publisher paid him a grand total of just £1,575, less than £200 a year, from which he had to pay his assistants), Johnson worked from a garret room off Fleet Street, where he defined some 43,000 words, illustrated with more than 114,000 supporting quotations drawn from every area of literature. It is little wonder that he made some errors and occasionally indulged himself with barbed definitions.

He had achieved in under nine years what the forty members of the Académie Française could not do in less than forty. He captured the majesty of the English language and gave it a dignity that was long overdue. It was a monumental accomplishment and he well deserved his fame.

But its ambitious sweep was soon to be exceeded by a persnickety schoolteacher/lawyer half a world away in Connecticut. Noah Webster (1758–1843) was by all accounts a severe, correct, humorless, religious, temperate man who was not easy to like, even by other severe, religious, temperate, humorless people. A provincial schoolteacher and not-very-successful lawyer from Hartford, he was short, pale, smug, and boastful. (He held himself superior to Benjamin Franklin because he was a Yale man while Franklin was self-educated.) Where Samuel Johnson spent his free hours drinking and discoursing in the company of other great men, Webster was a charmless loner who criticized almost everyone but was himself not above stealing material from others, most notably from a spelling book called *Aby-sel-pha* by an Englishman named Thomas Dilworth. In the marvelously deadpan phrase of H. L. Mencken, Webster was “sufficiently convinced of its merits to imitate it, even to the extent of lifting whole passages.” He credited himself with coining many words, among them *demoralize*, *appreciation*, *accompaniment*, *ascertainable*, and *expenditure*, which in

fact had been in the language for centuries. He was also inclined to boast of learning that he simply did not possess. He claimed to have mastered twenty-three languages, including Latin, Greek, all the Romance languages, Anglo-Saxon, Persian, Hebrew, Arabic, Syriac, and a dozen more. Yet, as Thomas Pyles witheringly puts it, he showed "an ignorance of German which would disgrace a freshman," and his grasp of other languages was equally tenuous. According to Charlton Laird, he knew far less Anglo-Saxon than Thomas Jefferson, who never pretended to be an expert at it. Pyles calls his *Dissertations on the English Language* "a fascinating farrago of the soundest linguistic common sense and the most egregious poppycock." It is hard to find anyone saying a good word about him.

Webster's first work, *A Grammatical Institute of the English Language*—consisting of three books: a grammar, a reader, and a speller—appeared between 1783 and 1785, but he didn't capture the public's attention until the publication in 1788 of *The American Spelling Book*. This volume (later called the *Elementary Spelling Book*) went through so many editions and sold so many copies that historians appear to have lost track. But it seems safe to say that there were at least 300 editions between 1788 and 1829 and that by the end of the nineteenth century it had sold more than sixty million copies—though some sources put the figure as high as a hundred million. In either case, with the possible exception of the Bible, it is probably the best-selling book in American history.

Webster is commonly credited with changing American spelling, but what is seldom realized is how wildly variable his own views on the matter were. Sometimes he was in favor of radical and far-reaching changes—insisting on such spellings as *soop*, *bred*, *wimmen*, *groop*, *definit*, *fether*, *fugitiv*, *tuf*, *thum*, *hed*, *bilt*, and *tung*—but at other times he acted the very soul of orthographic conservatism, going so far as to attack the useful American tendency to drop the *u* from *colour*, *humour*, and the like. The main book with which he is associated in the popular mind, his massive *American Dictionary of the English Language* of 1828, actually said in the preface that it was "desirable to perpetuate the sameness" of American and British spellings and usages.

Many of the spellings that he insisted on in his *Compendious Dictionary of the English Language* (1806) and its later variants were simply ignored by his loyal readers. They overlooked them, as one might a tic or stammer, and continued to write *group* rather than *groop*, *crowd* rather than *croud*, *medicine* rather than *medicin*, *phantom* for *fantom*, and many hundreds of others. Such changes as Webster did manage to establish were relatively straightforward and often already well underway—for instance, the American tendency to transpose the British *re* in *theatre*, *centre*, and other such words. Yet even here Webster was by no means consistent. His dictionaries retained many irregular spellings, some of which have stuck in English to this day (*acre*, *glamour*) and some of which were corrected by the readers themselves (*frollick*, *wimmen*). Other of his ideas are of questionable benefit. His insistence on dropping one of the *l*'s in words such as *traveller* and *jeweller* (which way they are still spelled in England) was a useful shortcut, but it has left many of us unsure whether we should write *excelling* or *exceling*, or *fulfilled*, *fullfilled*, or *fulfiled*.

Webster was responsible also for the American *aluminum* in favor of the British *aluminium*. His choice has the fractional advantage of brevity, but defaults in terms of consistency. *Aluminium* at least follows the pattern set by other chemical elements—*potassium*, *radium*, and the like.

But for the most part the differences that distinguish American spelling from British spelling became common either late in his life or after his death, and would probably have happened anyway.

In terms of pronunciation he appears to have left us with our pronunciation of *schedule* rather than the English “shedjulle” and with our standard pronunciation of *lieutenant* which was then widely pronounced “lefftenant” in America, as it still is in England today. But just as he sometimes pressed for odd spellings, so he called for many irregular pronunciations: “deef” for *deaf*, “nater” for *nature*, “heerd” for *heard*, “booty” for *beauty*, “voloom” for *volume*, and others too numerous (and, I am tempted to add, too laughable) to dwell on. He insisted that *Greenwich* and *Thames* be pronounced as spelled and favored giving *quality* and *quantity* the short “ä” of *hat*, while giving *advance*, *clasp*, and *grant* the broad

“ah” sound of southern England. No less remarkably, Webster accepted a number of clearly ungrammatical usages, among them “it is me,” “we was,” and “them horses.” It is a wonder that anyone paid any attention to him at all. Often they didn’t.

Nonetheless his dictionary was the most complete of its age, with 70,000 words—far more than Johnson had covered—and its definitions were models of clarity and conciseness. It was an enormous achievement.

All Webster’s work was informed by a passionate patriotism and the belief that American English was at least as good as British English. He worked tirelessly, churning out endless hectoring books and tracts, as well as working on the more or less constant revisions of his spellers and dictionaries. In between time he wrote impassioned letters to congressmen, dabbled in politics, proffered unwanted advice to presidents, led his church choir, lectured to large audiences, helped found Amherst College, and produced a sanitized version of the Bible, in which Onan doesn’t spill his seed but simply “frustrates his purpose,” in which men don’t have testicles but rather “peculiar members,” and in which women don’t have wombs (or evidently anything else with which to contribute to the reproductive process).

Like Samuel Johnson, he was a better lexicographer than a businessman. Instead of insisting on royalties he sold the rights outright and never gained the sort of wealth that his tireless labors merited. After Webster’s death in 1843, two businessmen from Springfield, Massachusetts, Charles and George Merriam, bought the rights to his dictionaries and employed his son-in-law, the rather jauntily named Chauncey A. Goodrich, to prepare a new volume (and, not incidentally, expunge many of the more ridiculous spellings and far-fetched etymologies). This volume, the first Merriam-Webster dictionary, appeared in 1847 and was an instant success. Soon almost every home had one. There is a certain neat irony in the thought that the book with which Noah Webster is now most closely associated wasn’t really his work at all and certainly didn’t adhere to many of his most cherished precepts.

* * *

In early February 1884, a slim paperback book bearing the title *The New English Dictionary on Historical Principles*, containing all the words in the language (obscenities apart) between A and *ant* was published in Britain at the steepish price of twelve shillings and six pence. This was the first of twelve volumes of the most masterly and ambitious philological exercise ever undertaken, eventually redubbed the *Oxford English Dictionary*. The intention was to record every word used in English since 1150 and to trace it back through all its shifting meanings, spellings, and uses to its earliest recorded appearance. There was to be at least one citation for each century of its existence and at least one for each slight change of meaning. To achieve this, almost every significant piece of English literature from the last 7½ centuries would have to be not so much read as scoured.

The man chosen to guide this enterprise was James Augustus Henry Murray (1837–1915), a Scottish-born bank clerk, school-teacher, and self-taught philologist. He was an unlikely, and apparently somewhat reluctant, choice to take on such a daunting task. Murray, in the best tradition of British eccentrics, had a flowing white beard and liked to be photographed in a long black housecoat with a mortarboard on his head. He had eleven children, all of whom were, almost from the moment they learned the alphabet, roped into the endless business of helping to sift through and alphabetize the several million slips of paper on which were recorded every twitch and burble of the language over seven centuries.

The ambition of the project was so staggering that one can't help wondering if Murray really knew what he was taking on. In point of fact, it appears he didn't. He thought the whole business would take a dozen years at most and that it would fill half a dozen volumes covering some 6,400 pages. In the event, the project took more than four decades and sprawled across 15,000 densely printed pages.

Hundreds of volunteers helped with the research, sending in citations from all over the world. Many of them were, like Murray, amateur philologists and often they were as eccentric as he. One of the most prolific contributors was James Platt, who specialized in

obscure words. He was said to speak a hundred languages and certainly knew as much about comparative linguistics as any man of his age, and yet he owned no books of his own. He worked for his father in the City of London and each lunchtime collected one book—never more—from the Reading Room of the British Museum, which he would take home, devour, and replace with another volume the next day. On weekends he haunted the opium dens and dockyards of Wapping and Whitechapel looking for native speakers of obscure tongues whom he would query on small points of semantics. He provided the histories of many thousands of words. But an even more prolific contributor was an American expatriate named Dr. W. C. Minor, a man of immense erudition who provided from his private library the etymologies of tens of thousands of words. When Murray invited him to a gathering of the dictionary's contributors, he learned, to his considerable surprise, that Dr. Minor could not attend for the unfortunate reason that he was an inmate at Broadmoor, a hospital for the criminally insane, and not sufficiently in possession of his faculties to be allowed out. It appears that during the U.S. Civil War, having suffered an attack of sunstroke, Dr. Minor developed a persecution mania, believing he was being pursued by Irishmen. After a stay in an asylum he was considered cured and undertook, in 1871, a visit to England. But one night while walking in London his mania returned and he shot dead an innocent stranger whose misfortune it was to have been walking behind the crazed American. Clearly Dr. Minor's madness was not incompatible with scholarship. In one year alone, he made 12,000 contributions to the *OED* from the private library he built up at Broadmoor.

Murray worked ceaselessly on his dictionary for thirty-six years, from his appointment to the editorship in 1879 to his death at the age of seventy-eight in 1915. (He was knighted in 1908.) He was working on the letter *u* when he died, but his assistants carried on for another thirteen years until in 1928 the final volume, *Wise to Wyzen*, was issued. (For some reason, volume 12, *XYZ*, had appeared earlier.) Five years later, a corrected and slightly updated version of the entire set was reissued, under the name by which it has since been known: the *Oxford English Dictionary*. The com-

pleted dictionary contained 414,825 entries supported by 1,827,306 citations (out of 6 million collected) described in 44 million words of text spread over 15,487 pages. It is perhaps the greatest work of scholarship ever produced.

The *OED* confirmed a paradox that Webster had brought to light decades earlier—namely, that although readers will appear to treat a dictionary with the utmost respect, they will generally ignore anything in it that doesn't suit their tastes. The *OED*, for instance, has always insisted on *-ize* spellings for words such as *characterize*, *itemize*, and the like, and yet almost nowhere in England, apart from the pages of *The Times* newspaper (and not always there) are they observed. The British still spell almost all such words with *-ise* endings and thus enjoy a consistency with words such as *advertise*, *merchandise*, and *surprise* that we in America fail to achieve. But perhaps the most notable of all the *OED*'s minor quirks is its insistence that Shakespeare should be spelled Shakspeare. After explaining at some length why this is the only correct spelling, it grudgingly acknowledges that the commonest spelling "is perh. Shakespeare." (To which we might add, it cert. is.)

In the spring of 1989, a second edition of the dictionary was issued, containing certain modifications, such as the use of the International Phonetic Alphabet instead of Murray's own quirky system. It comprised the original twelve volumes, plus four vast supplements issued between 1972 and 1989. Now sprawling over twenty volumes, the updated dictionary is a third bigger than its predecessor, with 615,000 entries, 2,412,000 supporting quotations, almost 60 million words of exposition, and about 350 million keystrokes of text (or one for each native speaker of English in the world). No other language has anything even remotely approaching it in scope. Because of its existence, more is known about the history of English than any other language in the world.