The Higher Plagiarism

by JONATHON GREEN

To steal from one author is plagiarism,
if you steal from many, it's research

Wilson Mizner (1876-1933)

And to steal from not just 'many' but literally (in both senses of that word) from thousands - what's that? Well, I fear that's lexicography, the making of dictionaries, especially when the dictionary in question, for example the multi-volumed *OED*, is what is called 'historical', i.e. offers not merely headwords, definitions and etymology, but tens, even hundreds of thousands of supporting citations - as far as possible from first use onwards - for every lemma.

The nature of lexicography, while we might wish it otherwise, is inevitably that of successive plagiarisms. The vocabulary does not appear fresh-minted for every new compiler. (Dictionaries of neologisms exist, but they are are hardly general-purpose. Similarly a vocabulary, that of computing, may change completely after so seismic a shift as the arrival of the PC in 1981, but even then the old 'mainframe' language still exists) The lexicologist Sidney Landau is perhaps unduly cruel to categorise such agglomeratioms as 'a recital of successive and often successful acts of piracy' but in esence he is correct. Each dictionary draws upon its predecessors, displays its own additions, whether stylistic or lexical, and waits until a new dictionary or new generation overlays it with fresh contributions, which in turn will be subsumed as new lexicographers attempt to pin down the latest version of a given language. Like Isaac Newton, the lexicographer can, or should, declare to his forbears, 'If I have seen further... it is by standing on the heads of giants' On that basis nearly every general-purpose dictionary is 'research'. The question remains: is it plagiarism too?

To take Mizner's definition more seriously than perhaps he meant it, one can see what has been termed plagiarism as a twofold concept: 'one author', or out-and-out plagiarism, and 'many authors', or scholarly derivation from previous sources. Lexicography offers examples of both. It is surprising, given the huge number of dictionaries, that there seems to have been less 'one author' plagiarism than might be expected. The hagiographers of certain pivotal figures have excoriated as 'plagiarists' those who acknowledged the excellence of their work by integrating it into their own, but this is often unfair. A classical example is that of Thomas Blount and Edward Phillips, whose works appeared respectively in 1656 and 1658. There are, undeniably, a good many similarities, and Blount was perhaps justified in his howls of fury. But in the end death cut short the argument and Phillips' work, entirely recast by a third lexicographer, went on to form the basis of a series of important advances in a new century. The work of an earlier author, John Rider, was allegedly stolen by another, Francis Holyoake. But Holyoake's accuser was not Rider, who was dead and in any case had pretty much abandoned the work that took his name as soon as the first edition was off the press, but the executors of yet another, even earlier, compiler: Thomas Thomas, whose work, they claimed, had been plundered by Rider himself. This argument too failed to take wing and petered out with no resolution. Indeed, no dictionary-maker seems ever to have taken the ultimate step – a visit to court – to stop publication of another's work.

Johnson too suffered alleged plagiarism, although he never complained. Not until Philip Gove, editor of *Webster's Third International* (1961), itself the source of even greater controversies (albeit unconnected with any form of literary theft), wrote about the marketing of Johnson's and Scott's dictionaries was battle joined on Johnson's behalf. According to Gove, Johnson was stripped bare by Joseph Nicol Scott, in his revision of Nathaniel Bailey's *Dictionarium Britannicum*. But Johnson in turn had worked with a copy of Bailey open at his desk. Given the bitterness that informed the 19th century 'dictionary war' between America's Noah Webster and Joseph Worcester, it is inevitable that plagiarism ranked among the mutual insults. But theirs was a battle of nationalisms as much as of literature, and if the anglophile Worcester had really plundered the anglophobe Webster, it is hard to see how this would have advanced his cause. Finally, although few of its compilers seemed to object, the collection of slang, certainly in its first two centuries, was rarely more than

barefaced, repetitious copying. But given that that the 'slang' of those early glossaries was essentially no more than a core glossary of criminal jargon or cant, it cannot be that surprising that one lexicographer copied his immediate predecessor.

'Many authors' plagiarism, what modern scholarship enjargons as 'intertextuality', is another matter. A variety of linguistic 'food-chain' a lexical Pac-man as it were, or one of those fabled meals where a swan encases a turkey, the turkey a goose, the goose a capon and so on down, perhaps to a wren, it lies at the foundation of lexicography. It is not excessive to see dictionary-making, from Greece to Rome, Rome to Byzantium and thence onwards to the present day, as a continuous thread of lexicographical accretion. Certain figures, undoubtedly, were more seized upon than others. Ambrosio Calepino's monumental efforts ensured that his eleven-language *Dictionarium* (1502) directly influenced the whole range of European dictionary-making for the next century and indirectly for even longer, given that those who absorbed his efforts would be absorbed in their turn. Similarly Charles Estienne's *Dictionarium Historicum, Geographicum, Poeticum...* (1553), became the source-book for a vast range of biographical and geographical dictionaries. Every lexicographer looks backwards. The language may have an infinite future; the lexicographers draw on a near-infinite past. Every early lexicographer tells the same story, either listing their sources or, like John Rider in 1589, assuring readers that they had read, as it were on their behalf, the 'learned workes of all the learnedst and best Dictionaries in England.' One may disapprove, but such is the system. The lexicologists DeWitt Starnes and Gertrude Noyes observe that 'As much as we may wish to overlook the fact […] (1) in this early period lexicography progressed by plagiarism; (2) the best lexicographer was often the most discriminating plagiarist; and (3) a good dictionary was its own justification, whatever the method of compilation.' Such truths are in the end a lexicographical donnèe and this absorption continurs, even if the lexica are produced not by individuals but by corporate or academic teams.

Most dictionaries, therefore, accept the tradition of 'many authors' plagiarism as to their headwords. But 'one author' plagiarism does exist, and it is most likely found in the accretion of citations, the 'usage quotations' that form the spine of the 'historical' dictionaries and provide vital information that unadorned headwords and definitions never can. After all, what is the ransacking of material, be it Shakespeare or hip-hop lyrics, but the theft of another person's original creativity. One is, of course, presenting such material for the best possible reasons – to illustrate language use and development – and it would take a truly wretched anorak to attempt, say, the recreation of

Ulysses from the 900 examples of slang I have plucked from Joyce's masterwork. But the truth is cut, cut and paste, and the more cutting and pasting you can achieve, the better the dictionary one makes. The superiority of the *OED* over all English language rivals lies in this enormous collection of cites. The proof and the ultimate underpinning of the pudding of headwords and their definitions and etymologies.

The first English-language citations appeared in Blount's *Glossographia* in 1656 (the lexicographers of Greece and Rome actually originated the practice) but for many dictionary-users it is Samuel Johnson who brings the art to fruition a century on. Johnson credits his primary influences as to headwords – his predecessors Bailey, Ainsworth and Phillips, and the etymologists Junius and Skinner – although he backed their word-lists (culled in their turn from centuries of earlier dictionaries) with his own reading as well as that of his amanuenses. As for citations Johnson turned to the literary canon, but with magnificent self-possession, was not above rewriting the masters. In all he amassed 116,000 quotations to cover 40,000 headwords. As he admitted in prefaces to the later editions, some quotations were compressed and others modified. This was an understatement. If Johnson didn't like a quote, he changed it. He would omit an opening phrase or amputate a conclusion and if a phase didn't convey the meaning he required, he had no scruples in rewriting it. Nothing was sacred. When Johnson disliked the style of an author, he did not hesitate to improve it. He was using earlier authors merely as a means of providing evidence of the proper usage of words: if in the end they failed to meet his standards, then the least he could do was provide his readers with the desired ideal.

But Johnson was an original. And he, or his helpers, read that which they quoted. By the time Sir James Murray and *his* team set about the OED in 1870s, a degree of 'creative borrowing' had appeared. Massive reading programmes underpinned Johnson's successor, but the Oxford Scriptorium took whatever was on offer. For a number of words, publication in an older dictionary provides the only, or at least earliest, use they could find. Aside from the oft-repeated notation '(J.)', indicating definitions plucked wholesale from Johnson himself, there are, inter alia, 1,306 citations from Henry Cockeram's work of 1623, 81 from Rider, 250 from John Kersey (1701), and more than 5,000 from Nathan Bailey. Thomas Harman's *Caveat* (1567), which reproduced the entire recorded cant lexicon of its time, seems, in its 235 citations, to have been taken up in its entirety. Only American's Noah Webster prefers to isolate himself from the trend. Like his peers, Webster drew on other researchers when amassing his word-list, but that most religious of lexicographers allowed but a single credit when assessing his work: to 'God Almighty'. The *OED*

is less self-denying: its current edition acknowledges Webster's various editions more than 5,000 times.

The nitty-gritty, however, comes not in reproducing another dictionary's definitions for citatory purposes, but its actual citations. At this point one enters, if such things matter, an opaque area best categorised as lexicographical ethics. Call it honesty, call it professional pride, call it downright absurdity in a world of genuine moral choices. Those who hear the pit-a-pat of angels tripping lightly upon pinheads are to be forgiven. So, the question. Here is a citation. I have found it by reading a specific 19th century work. However, when I look into Bartlett's *Americanisms* (1849), there it is. Checking Farmer and Henley's *Slang and Its Analogues* (1890-1902), reveals it again. And again when I look into the *OED*. And in Jonathan Lighter's *Historical Dictionary of American Slang*? Yes, same one. So what happened here? I can satisfy myself as to my honesty – but was it really necessary. Couldn't I have just checked one of the others – and done my cutting and pasting? Not to pretend that I don't – but I do give credit (initials after the cite) where it is due. But, and I am not trying to be smug here, this is not always so. The Second (1989) edition of the *OED* has gained a remarkably impressive selection of cites for mid and late 19th century American terms, usually culled from the small-town press. How are we to square this with the appearance, in 1951, of those very cites in Mitford Mathews' *Dictionary of Americanisms*. Is it a coincidence worthy of the *Fortean Times*. Or . . . shall we call it 'reasearch'.

In some sub-sets of the lexicon, 16th and 17th century criminal cant for example, one has no choice. Those who set it down in print are limited, mainly lexicographers themselves, each successor is forced back to their word-lists. Bereft of anything better, the dictionaries of 1531 or 1725 become the citations of 2000. But the search for new citations, and especially first use citations does have its charms, and provides as good a reason as any to forgo that easy 'research'. As one who has just stumbled upon a use of 'male chauvinist' that predates the current OED first use by fifteen years (and presented it to the word mavens of Clarendon Street with due pride) I can attest to this. But then we drudges – something has to liven our days.

In the end plagiarism, of whatever sort, is inescapable. Like the historian, who cannot conjure up a pristine 'new' version of history, but must add to what has gone before, the lexicographer cannot simply throw out a word-store that has accrued over centuries. As John Algeo has put it in the *International Encyclopedia of Lexicography*, 'a really new dictionary would be a dreadful piece of work, missing innumerable basic words and senses, replete with absurdities and unspeakable errors, studded with biases and interlarded with irrelevant provincialism.' And for those who ask, desperate for verification, 'Is it in the Dictionary?' it would be useless. And how many ways can there be of defining a given word? No reputable dictionary would simply reproduce material stolen from a predecessor, but it is inevitable, and generally accepted, that just as the language demands that the word-list must, while gradually expanding, remain relatively stable, so too must the definitions included in that list.

Let Oliver Wendell Holmes, writing in 1858, have the last word: 'When I feel inclined to read poetry I take down my Dictionary. The poetry of words is quite as beautiful as that of sentences. The author may arrange the gems effectively, but their shape and lustre have been given by the attrition of ages.'

[2001]