

## > **The Genre of the Etymological Dictionary**

ANATOLY LIBERMAN

### **1 Etymological dictionaries among other comprehensive dictionaries**

Etymological dictionaries are stepchildren of lexicography. In surveys, at best a few pages are devoted to them. Even lists of the etymological dictionaries of English, German, Dutch, and the Scandinavian languages, that is, of languages having a strong tradition of producing such reference works, were impossible to find until I compiled and published them (Lieberman 1998; 1999; 2005; the Dutch list still awaits publication). But years after I began this work I keep running into old, not necessarily worthless, works that fell between the cracks and wonder how many more I have missed (cf. the postscript to my 1999 paper added in proof and Lieberman 2009b). One does not have to look far for the reason complicating this search. The common habit of depending on the latest products, which, allegedly, contain more pieces of distilled wisdom than their predecessors, severed our ties with the past, and few experts consult the first editions of Kluge (1883, etc.) or any of the four editions of Wedgwood (1859-1865, etc.), let alone dictionaries by less distinguished authors. As a result, references to them are rare. Theory of lexicography and excellent instructions to lexicographers exist (the market seems to be always ready for new encyclopedias and voluminous ‘handbooks’), but no one except Yakov Malkiel (1975) has taken the trouble to analyze the practice of etymological lexicography or look at the multitude of etymological dictionaries written in the post-medieval period. Those who know his book may agree that despite its scope it is not a fully satisfactory guide to the subject, partly because of Malkiel’s penchant for baroque style and partly because he was preeminently an expert in Romance linguistics, which made his opinions about Germanic and Slavic lexicography less valuable. Nor can a bird’s eye view of any subject replace a series of more specialized works.

It is easy to see why etymological dictionaries have been pushed to the margin of theoretical lexicography. Ever since people became literate, they have been putting together glossaries and ‘lexicons.’ Travelers, merchants, statesmen, and officials had to communicate with foreigners, and in every epoch some language had the status of the most prestigious one, whether Egyptian, Hittite, Greek, Latin, French, or English. It is therefore no wonder that we have bilingual texts from the dawn of human civilization and thousands of medieval glosses. The collapse of the Tower of Babel provided language teachers and lexicographers with permanent employment. As time went on, culture gave an impetus to the

compilation of dictionaries of synonyms, homonyms, antonyms, slang, and so forth. By contrast, etymology, though excellent for lexicographic dessert, cannot pretend to be anyone's main course. To speak, read, and write well, we do not need information on word origins.

The study of language history is inseparable from etymology, but the public can thrive without knowing where words came from, and two factors keep this branch of scholarship afloat: the inertia of historical linguistics as an academic subject and humanity's natural curiosity. People love etiological tales ('just so stories'): they want to know how the big bang occurred, why the bat hunts at night, when and under what circumstances language originated, and, among other things, how sign (the form of any given word) and meaning are connected. This thirst for knowledge is almost instinctive (at least it is ineradicable), as shown by the popularity of word columns and countless books with titles like 'Why Do We Say So?' Etymological dictionaries purport to satisfy both professionals and the uninitiated, but, as regards their appeal, they cannot compete with explanatory, bilingual, and pronouncing dictionaries. To put it bluntly, they occupy the place they deserve, but without them the world would have been poorer; so may they live and multiply.

## 2 The reception of etymological dictionaries

In literary studies and art history, reception theory has been a major topic for decades. In lexicography, it hardly exists at all. Scandal once resulted in the appearance of books featuring and commenting on the main reviews of a dictionary (Sledd [and] Ebbitt [1962]; Morton [1994]). Of course, I mean *Webster's Third International...*, and how misspent those passions appear to us today! Dictionary wars have been documented. The reaction of the public to Samuel Johnson and the *OED* has been traced up to a point, but on the whole, as I said, reception of dictionaries by lay users and professionals is an almost nonexistent area. The authors of etymological dictionaries are even worse off than other lexicographers, for reviews of such dictionaries have never been collected or analyzed. Sometimes I wonder who reads them. Even the authors often disregard sensible suggestions while preparing later editions. Probably they have never seen the reviews.

I am speaking from experience. More than twenty years ago, I began work on a new etymological dictionary of English. My goal was to write entries in which the literature on the origin of words would be discussed as fully as possible, various conjectures sifted, and reasonable conclusions

drawn from the data. The models were many: Walde-Hofmann for Latin (1938-1954), Feist for Gothic (1939), Vasmer for Russian (1950-1958), and quite a few other etymological dictionaries (of Hittite, Classical Greek, French, Spanish, Old High German, Lithuanian, several Slavic languages, Old Icelandic, and Old Irish). An English dictionary of this type does not exist. Skeat (1882/1910), like his predecessors and followers, gave almost no references, so that someone who decides to study the etymology of an English word in depth starts practically from scratch. To what extent the project on which I embarked in the early eighties is feasible, given the resources at my disposal, is beyond the point in the present context, but the task I faced could not be clearer: it was necessary, for the first time ever, to collect the enormous literature on the origin and history of English words, summarize the findings, and offer convincing solutions.

To find the relevant articles not only in the most visible journals but also in countless fugitive periodicals (with minor exceptions, word columns and letters to the editor in newspapers remained untapped by my team of about a hundred volunteers and meagerly paid assistants) was a formidable task. I had no illusions about the completeness of the final product (one cannot read everything; besides, new articles and books appear every day), especially because etymology is based on a good knowledge of cognates. One should screen the literature in and on all the Indo-European languages (and occasionally on the languages of other families) in the hope of finding the sought-for answers outside English (for example, someone might have guessed the origin of German *gleiten*, and this would solve the etymology of Engl. *glide*, or perhaps a preliminary agreement has been reached on Dutch *big*, which would then shed light on its connection with Engl. *pig*; the importance of works on borrowings and on words belonging to the Indo-European stock needs no proof). Reviews were among the sources I studied with great care. All the publications used for the database have been copied, and more than 20,000 of them are kept in my office. At least a thousand of them are reviews.

The database, as well as the introductory ('showcase') volume of the dictionary, has now been published (Lieberman 2008 [dictionary] and 2009a [bibliography]). Since the reviews that ended up on my desk could be put to use only insofar as they contained discussion of words, some, however insightful, were not included in the bibliography, but I excerpted and preserved the rejects. My acquaintance with them (brief and long, devoted to minutiae and attacking general questions) justifies my statement that reviews of etymological dictionaries have not been used for any conclusions about the genre of the etymological dictionary and

exercise minimal influence on the authors. I hope to write a book on the etymological dictionaries of the Germanic languages and in addition to a survey and analysis of all of them, discuss their reception. Over the years, reviewers have been asking and often answering the same questions that interest me. Perhaps this chapter will even expedite the birth of reception theory in lexicography.

### 3 The readership and the market of etymological dictionaries

Above I said that etymology stays alive (or afloat) because it is the foundation of historical linguistics and because the public wants to know where words come from. Every dictionary has a certain user in view. Although the authors of etymological dictionaries cannot disregard this circumstance, they do not always know what to make of it, for their idea of their audience is dim. It is instructive to compare introductions to etymological dictionaries. In Western Europe, the earliest of them appeared in 1599 (Kilian; Dutch). Kilian's work was followed by similar dictionaries of German, English, French, and other languages. Front matter sometimes ran to more than a hundred pages and offered the author's views on the origin of language and the derivation of words (a tradition that was upheld by Wedgwood and Skeat among many others, who in this respect did not differ from Samuel Johnson and Webster). It was not deemed necessary to justify the production of such a book since the uses and benefits of etymology were taken for granted.

The first dictionaries were sometimes sold by subscription, and the lists of subscribers are long and impressive, from dukes down. Occasionally the first edition would be brought out by the author, who would break even or make a profit, so that the next edition would be undertaken by a commercial publisher. This is what happened to Minsheu, for example (1617; 1627). As late as the nineteen-eighties, in the days of Skeat and Kluge, etymological dictionaries still had a respectable market: every gentleman was likely to own a copy, and country squires read them like fiction. Nowadays publishers depend almost entirely on libraries. Outside the circle of historical linguists, most people have a lively but perfunctory interest in word origins, which seldom goes beyond *exotica*, slang and family names. Even if etymological dictionaries were less helpless in dealing with the origin of *cocktail*, *dodge*, *scalawag*, and their likes, what they say on the subject can be found in any other 'thick' dictionary. One would have thought that this state of affairs would have stopped the production of vapid etymological dictionaries, but the stream flows on unimpeded.

Books are published to be sold. Hence the tendency to coax the reader into purchasing etymological dictionaries by emphasizing their novelty (a new edition is supposed to be an improvement on the previous one by definition), accessibility (no knowledge of linguistics or any other ‘prerequisite’ is expected), and increased bulk (the more words are included, the better—also by definition). Every now and then an additional incentive is mentioned. Under the Nazis, Kluge’s classical work was advertised as ‘a German dictionary for the German people.’ The preface to a relatively recent (serious and scholarly) etymological dictionary of Icelandic celebrated the fact that it was the first etymological dictionary of Icelandic written in Icelandic. At nearly the same time an etymological dictionary of English brought out in the United States proudly announced that it was the first American work of its kind. This focus had predictable negative consequences. Presses churn out ever new dictionaries that recycle trivial information. Fortunately, the sources of academic subsidies have not yet dried up, and from time to time we witness the appearance of real etymological dictionaries, rather than their pompous digests.

The truth of the matter is that an etymological dictionary requires a prepared user. Since grammar is not considered to be ‘fun,’ our college graduates have trouble distinguishing nouns from adjectives and subjects from objects. (A recent handbook of linguistics for literary scholars provides its readers with the definitions of such terms as *vowel* and *consonant*.) The origin of slang is hard to discover but easy to explain. All the other cases are more complicated. No one without previous exposure to special courses can appreciate the methods of etymological analysis. The role of cognates (and the concept of a related form), their choice, the difference between a cognate and a borrowing, sound correspondences, the idea of a protoform, the periodization of language history (Old/Middle/early Modern English, archaic Latin, and so forth), their confusing nomenclature (Germanic, German; Baltic; Old Prussian, which has nothing to do with the language spoken in modern Prussia; Old Saxon, Anglo-Saxon, German Saxon dialects; Old and Middle German, which is sometimes ‘High’ and sometimes ‘Low’; Anglo-French and northern French as opposed to Parisian French; Middle Dutch contemporaneous with Old Frisian, the absence of ‘Old Dutch’; Old Norse: what is Norse?), the basic facts of history (the Scandinavian invasion, the Norman Conquest, the epoch of Humanism) are subjects most of which the so-called general reader rarely knows. Elmar Seebold supplied his revision of Kluge’s dictionary with a long list of terms, including *umlaut*, *ablaut*, *vridधि*, and many others, quite unlike *vowel* and *consonant*. Yet that same semi-mythical general reader who wants to learn the origin of a German

word will hardly agree to wander in the thicket of special terms and ponder their meaning. Nowadays we expect instant satisfaction or ‘money back.’ Etymology does provide satisfaction, but it is not instant, and there is not much money in it.

Reputable, especially academic, publishers have to choose between profit and excellence. In theory, they strive for both, but they cannot afford big losses: once they are out of business, there will be neither profit nor excellence. Good etymological dictionaries are doomed to attract only specialists. In principle, this conclusion does not spell disaster. Books on laryngeals in Indo-European, Verner’s Law, palatalization in Dutch dialects, Scandinavian accents, and even the style of Shakespeare’s sonnets are written for those who study accents, palatalization, and the rest. Our society is still rich enough to support the interest of a chosen few in such esoteric subjects. But a feeling prevails that dictionaries are different: allegedly, they must be ‘popular.’ This feeling may be justified in many cases, but not in etymology. Etymology is no less special than organic chemistry, and etymological lexicography has to resign itself to this fact. Since I work on an etymological dictionary of English, I will confine myself to the field I know best. The latest special etymological dictionary of English was published almost exactly a hundred years ago (Skeat 1910). It was not special enough (see section 5, below), but it lived up to the expectations of its readership. The dictionary was addressed to language historians and to those who had a good deal of Latin and Greek driven into them at school and after. Although Skeat never tired of berating his countrymen for their laziness, ignorance, and inability to understand what etymology is all about, he relied on their familiarity with the rudiments of grammar in its Latin guise (a luxury none of us can afford) and understanding that languages develop and change. His concise dictionary was just that: concise (the entries were shorter but not simplified in comparison with the full opus). Even if sometimes he despised his reader, he does not seem to have been worried by the idea that he was talking over his head. The next etymological dictionary of English, and the last written by a serious researcher (Weekley 1921), is a watered down version of Skeat and the *OED*, though Weekley had many original ideas, especially about words of French origin and words derived from names. In the English speaking world, specialists as the main target group were forgotten (not in theory but in practice), for they could not make such dictionaries profitable. Since roughly the First World War English etymological dictionaries have been written only for ‘everyman.’

The results of this attitude were disastrous. Despite the abandonment of the idea that *Prinzipien der Sprachgeschichte* is a synonym of *Introduction to Linguistics* and the reorientation of language sciences toward synchrony, etymological studies continued. The 20th century witnessed outstanding progress in the investigation of early Indo-European. Great etymological dictionaries of the living Romance, Slavic, and Baltic, as well as of several dead and reconstructed languages, were written, and numerous publications clarified the origin of both common and obscure words in and outside English. But the authors of new English etymological dictionaries paid minimal or no attention to them. In some cases they did not know those works (if I am not mistaken, the indefatigable Eric Partridge, the author of the widely used dictionary *Origins*, could not read German), in others they did not care to do sufficient research. Collecting books and articles on etymology is a time-consuming occupation (see above), and ‘everyone’ had no intention to ‘dive below’ or watch lexicographers perform such aquatic tricks. Recycling and repackaging the information in Skeat and the *OED* guaranteed safety (for even their outdated opinions were clever) and satisfied the public that could not distinguish between original work and a rehash of the classics. That is why not a single dictionary of English etymology contains more than one volume, while the great dictionary of French occupies a whole shelf, and in this respect von Wartburg, its author and editor, was not alone, though no other project acquired such gigantic dimensions (that is, if we stay with a dictionary of a single language rather than a group).

Surprisingly, with regard to etymology, two ‘thick’ English explanatory dictionaries went beyond what we find in Weekley and even Skeat. Since the days of Blount (1656) every comprehensive English dictionary has included information on word origins. As long as etymology remained guesswork and every conjecture seemed to be thought provoking, this custom made sense, but the appearance of the first edition of Skeat (1882), and before him of Wedgwood (1859-1865), though Wedgwood never wielded so much authority, dictionary makers could only copy from the recognized masters: inventing more sophisticated derivations became a dangerous enterprise. An etymologist or a consultant remains a familiar figure on the staff of some great dictionaries. However, this person’s duty is not to outdo Skeat and the *OED* but to replace their conjectures if those have been shown wrong (an unusual occurrence) and trace the origin of the most recent words. Even that task is hard to perform, and not too rarely we find the latest dubious solution replacing the old one only because it is brand-new and published in a prestigious journal by an illustrious author (as happened to *boy*, *girl*, and *filch*, among others; a typical example

of ‘haste makes waste’ or ‘don’t jump on the bandwagon’), along with the absolutely secure but irritating verdict ‘of unknown origin.’

The etymologies in Webster, even in the Collegiate digest, are not easy. They presuppose a user aware of the things mentioned at the beginning of this section, that is, someone who knows the difference between Old and Middle English, appreciates the role of cognates and protoforms, understands that English is a language of the Germanic group and of the Indo-European family, and has been taught that etymology can seldom be absolutely certain. But those who open Webster more often look up meanings and can skip the etymological introduction. The answer to the fateful question—has such a dictionary been written for the expert or for ‘everyman’?—depends on the etymologies to a minimal degree, while in a specialized etymological dictionary the information on the origin of words is all that matters. In any case, the two dictionaries, mentioned above, bravely offered detailed and highly professional etymologies. They are *The Century Dictionary* and Wyld’s *UED*. As far as I can judge, both are hardly ever used by modern etymologists, who rely only on Skeat and the *OED* (my opinion is based on the absence of references to them in the books and articles I read). Yet both have a great deal to offer, and some of their suggestions do not recur elsewhere. Depending on the word we investigate, they may be more useful than any post-Skeat and post-*OED* etymological dictionary of English.

Serious etymological dictionaries, such as are worthy of their name, should be written for specialists and in this respect share common ground with books on mechanical engineering, calculus, and a host of others. This probably means that they can be published only by presses depending on institutional support and that in the first year hardly more than two or three hundred copies will be sold. Skeat’s days are over: an etymological dictionary is no longer a status symbol and cannot rival an easel with an unfinished picture or a piano with the score of Brahms’s variations, even though no one around ever painted or played. After the appearance of a dictionary written for the expert, producing a simplified concise version is an excellent idea, but first the profession should be served. Otherwise there will be nothing to simplify or abridge and we will forever stay with books on ‘why do we say so?’ Nowadays the word *elitist* is a term of abuse. Let this semantic somersault remain on the conscience of those who coined it. We should not be bullied into the belief that the only possible variety of an etymology dictionary is the popular one.



#### 4 The stratification of vocabulary in etymological dictionaries. The words to be included

In what follows I will concentrate on dictionaries for the expert. Their genre also needs clarification. One of the main questions here is about the depth of etymologies. At the moment, the prevalent tendency in Indo-European etymological lexicography is to stress distant reconstruction. The thrust of the Leiden project is a good example of what can be expected in the future, and the Norwegian experience points in the same direction. This approach has its advantages and disadvantages. In the modern Indo-European languages, most words, if we exclude borrowings from Greek, Latin, and partly French, cannot be traced to hoary antiquity, so that the scope of the inherited element in their vocabulary is limited. In the Germanic, Slavic, and Celtic languages, hundreds of words, even the older ones, have likewise at best vague Indo-European connections despite their age. In the attempt to derive as many words as possible from the roots in Walde-Pokorny (1927-1932) and Pokorny (1959) or relegate them to the pre-Indo-European substrate compilers pay less attention to late medieval, early modern, and recent words. The *cocktail-dodge-scalawag* group is a distraction to them, and they prefer to ignore its existence or confess that they have nothing to say about its history.

Lexicographers never stop bothering about the number of words to be included in the dictionaries they edit. Considerations of size interest the authors of etymological dictionaries no less than their colleagues, but I am not sure that they have ever been debated. As pointed out, all one can find is an occasional blurb promising more words than ever. Since modern 'thick' explanatory dictionaries regularly feature etymologies (with the result that someone who wants to know the origin of *come*, *go*, *take*, *door*, *wall*, and other common words will find a brief but reliable answer in any non-etymological dictionary, whether on paper or online), it seems that at present the main effort should be directed toward the derivation of the words passed over or insufficiently explained. I see no virtue in writing a long entry on *brother* and *eight* in an etymological dictionary of English, as opposed to two uninformative lines on *scalawag* and omitting *dodge* altogether. Obviously, *brother* and *eight* cannot be excluded, but the level of our knowledge is such that the entries on them often contain only lists of cognates. Since knowing those cognates is not tantamount to understanding their origin, I would argue for giving them limited, perhaps even minimal space in prospective etymological dictionaries, unless the author has new ideas on how they were coined.

In my opinion, the tradition of writing Romance dictionaries has everything to recommend it. Wherever possible, their authors trace the word to Latin, and the rest is left out. The implication seems to be that anyone who is enlightened enough to understand an entry in a French, Italian, or Spanish etymological dictionary and wants to know more should turn to a dictionary of Latin. Nor is Latin our last destination, but in Walde-Hofmann the oldest reconstructable form in the spirit of Walde-Pokorny is also given. The message is: If you want to go as far as the present state of the etymological science allows us to go, use several dictionaries. Obviously, my proposal runs counter to the ideas of those who are mainly concerned with the Indo-European sources of our vocabulary. Yet my and their efforts are not at cross-purposes; they are rather complementary. I am only saying that even an ideal etymological dictionary cannot be all things to all people.

Etymologists venerate archaisms. It is their pleasure to walk among ruins, and their predilections should be treated with understanding and a measure of sympathy. A great favorite of English etymological dictionaries used to be the adjective *nesh* 'soft.' The word is regional, and the reason it has been favored over thousands of other local words is its ascertained old age (it was recorded in Old English, and its cognate turned up in Gothic, Old High German, and West Frisian) and its relative transparency (see *hnasqus\** in Feist). Very few people will look up *nesh* in Skeat, but those who will may not know Gothic *hnasqus\** or Old Engl. *hnesce* and will be grateful for finding an entry on it.

A more revealing example is the adjective *loom* 'moderate, gentle' (said of a breeze). This is also a local (northern English and Scots) word, and only the most detailed dictionaries of Modern English include it (for example, it will be found in the third edition of *Webster's International...*, but not in the latest edition of *The Shorter Oxford...*). According to Webster's dictionary, the origin of this adjective has not been discovered. It would be a waste of space to feature it in an etymological dictionary for the sole purpose of saying 'of unknown/uncertain/debatable origin.' But the history of *loom* (adj.) is not a blank. Frisian dialects have *luum* 'lazy, depressed,' *loom* 'thin, tired, lazy, etc.', and other similar forms. It is not clear how, if at all, they are related to Dutch *loom* 'slack, slow, etc.', Old High German *-luom* occurring in several compounds, German *lau-* in *lauwarm* 'tepid, lukewarm', the Germanic words for 'lame' (Engl. *lame*, German *lahm*, etc.) and Engl. *loom* 'appear indistinctly' (possibly of Low German or Dutch origin), but, in any case, the English adjective has been removed from its isolation, and now it is the etymologists' business to disentangle the knot. I knew neither

the English adjective nor its Frisian cognates until I read about them in Faltings (1996:106-107). Unfortunately, I missed the article while working on my database, and it does not appear in the published bibliography (Lieberman 2009).

This example shows that a consultant in the employ of a great dictionary has no chance of revising obscure etymologies. How can anyone confronted with the question about the origin of the English adjective *loom* find the elucidating passage in *NOWELE* 28/29? And who has enough time to begin searching for cognates in the multiple dialectal dictionaries of Frisian and Low German on the off-chance of running into something useful? Lexicographers are always ‘caught in the web of words’ and cannot afford spending long hours on what may become a wild goose chase, for they suspect that if the origin of a hard word is still unknown, there must be a good reason for it. (Don’t we remember Meillet’s unkind and unfair remark that all the good etymologies have been discovered, while new etymologies are usually bad? I think this was said at least a hundred years ago.) My team and I screened all the available philological and popular periodicals in two dozen languages for more than three centuries and ‘by chance’ unearthed countless important but forgotten publications whose titles frequently had no bearing on etymology; yet I managed to overlook Faltings’s article despite its promising title. Another lesson I can draw from this example concerns the choice of words. Even if a word is rare or local but if its origin is worthy of note (such are, to my mind, the English adjectives *nesh* and *loom*), it should be featured in an etymological dictionary. Including them is a luxury, but an etymological dictionary has been conceived as a feast for hungry minds, however Micawberian this statement may sound.

In deciding how many words to include in a dictionary, etymologists will be well-advised to show restraint. Not only the users’ expectations but the state of the art and common sense should guide their hand. Given the overabundance of competing presses and the easy access people (at least in developed countries) have to the Internet, an etymological dictionary has become a reference book for a limited, mainly professional readership (advanced students and their teachers). Those who want to learn the origin of *antelope*, *papaya*, *baritone*, *algebra*, *samovar*, and *Schadenfreude* will hardly buy or even open an *English* etymological dictionary, for their curiosity can be satisfied in an easier and cheaper way (just Google for *papaya*: etymology or *algebra*: etymology, or look it up in the pocket edition of *The American Heritage Dictionary of the English Language* and be informed: ‘*papaya* <Cariban [sic],’ ‘*algebra* <Ar[abic] *al-jabr*, *al-jabr* ‘the (science of) reuniting’; do many people want to go further?’). This does not mean that

borrowings should be ignored. *Bigot, ghetto, fiasco, rascal*, and others like them should be dealt with in detail, even though an English scholar hardly has enough expertise to risk an independent opinion about their history. Skeat included about 12,000 words in his dictionary, and many of them could have been dispensed with. At present, seven or eight thousand non-derived words will be quite enough to answer the main questions about the origin of English vocabulary.

Before concluding this section, I would like to explain to what extent I do what I preach. The English etymological dictionary on which I have been working for more than twenty years won't be comprehensive, like Skeat's or Weekley's, or the *ODEE*. At the very outset I realized that I would have little or nothing to contribute to explicating words like *brother* and *eight*. Not only the basic numerals and kin terms but also many other words with broad Indo-European connections, such as *hear* and *break*, have been the object of numerous profound articles, dissertations, and books. It would have been presumptuous to expect that I was able to offer innovative suggestions in this area, and I did not want to spend the rest of my life only writing summaries of other people's opinions. The same holds for loanwords from Romance languages. I can perhaps shed some light on the etymology of Germanic words without established cognates in the rest of Indo-European, but even here one has to tackle *bride, dwarf, God*, the notorious maritime vocabulary (*ship, sea, sail*, etc.), and the rest. So I decided to concentrate on the dregs of English etymology, the words lacking unquestionable cognates outside English. It is these words about which dictionaries usually say: 'Origin unknown.' Predictably, I ended up with *scalawag-cocktail-dodge*. My database is all-encompassing (whatever came to my mill was called grist), but, if my rough estimation is accurate, the dictionary will feature about 1,000 words like those three. Some of them have been around for centuries. Before looking at every word like *nesh* and *loom*, I cannot say which of them I will include. I will avoid volatile and exotic slang, but, other than that, slang will occupy a noticeable place in the final product.

So far, my experience has been entirely positive. As soon as about two-thirds of the bibliography had been assembled, it became clear that even conscientious researchers are unaware of some valuable publications to the field. Jacob Grimm read all there was to read, and so did (presumably) Benfey, Feist, and perhaps von Wartburg, but hardly anyone else (not Skeat, and certainly not Onions). The standard verdict 'origin unknown' about words like *cocktail* often does not reflect the state of the art. It is rather a comment on lexicographic practice: dictionary makers copy from one

another, and endless repetition produces the illusion of consensus. The origin of *cocktail* (to stay with the same example) was clarified decades ago, but this clarification has found its way into very few dictionaries. I assume that bibliographies do not belong to consultants' everyday reading, for the titles could not be more revealing ('The Origin of *Cocktail*'; five of them). The same is true of many other words. Even when no one could offer a fully convincing solution, I found that good suggestions abound. This happened in my investigation of the etymology of *dwarf* belonging to the Germanic protolanguage (in this case my mythological studies required visits to several foreign lands): a tentative suggestion in the first edition of Kluge's dictionary (under *Zwerg*) gave me a clue; the rest, as chess players say, was a matter of technique. Most of the words I will write about will probably remain to some extent obscure, but their origin will become partly 'known' (compare *loom*, above).

## 5 The depth and breadth of entries in an etymological dictionary

Entries in etymological dictionaries tend to be short (five and more per page; only the best Romance dictionaries are an exception to this rule). Skeat treated some words at greater length, but, in principle, he managed to say all he wanted in the concise version of his great work. The meaning and the pronunciation of a word in a living language can be discovered by turning to native speakers, whereas etymology depends on reconstruction and from the nature of the case is seldom 'final.' The main part of an etymological entry should be discussion, and this is what we find in Walde-Hofmann, Feist, *EWA*, and many other works. However, some dictionaries, including the earliest editions of Kluge, only state the opinions of their authors. Skeat sometimes explained why he disagreed with his predecessors or who inspired his solutions, but, as a rule, he avoided polemic and 'the history of the question.' In English studies, only such dictionaries exist to this day.

Skeat and Kluge were the first authors of reliable etymological dictionaries of English and German. In 1882 and 1884 the solid tradition at their disposal was a few decades old (this holds especially for Skeat's experience, even though the impact on him of early scholarship should not be underestimated). The 'pioneers' reticence had good reasons. Since the eighties of the 19th century, dozens of dictionaries and innumerable articles and books on word origins in the Indo-European languages have been written, and suggestions on where this or that word came from are countless. Some of them are too speculative, but most merit attention. I believe that a modern dictionary of a language with a rich tradition

of etymological research should contain a summary of the views on the origin of every word featured in its pages.

The breadth and depth of the discussion poses various questions, and the answers to them depend on the type of the dictionary and the nature of the material. Feist, the author of an incomparable Gothic etymological dictionary, worked with a closed corpus containing a limited number of items. Since the study of the Gothic language forms the foundation of Germanic and to a certain extent Indo-European philology, every recorded word of that language counts. This is a dictionary oriented toward reconstruction, and Feist showed the place Gothic vocabulary occupies in the entire panorama of Germanic and Indo-European. In performing such a formidable task, he could not always decide where to stop and gave superfluous references. In examining well-preserved dead languages (Sanskrit, Classical Greek, and Latin are the best examples), an etymologist encounters the problems similar to those confronting a student of Modern English, French, or Russian. The stock is huge, with some words belonging to slang and others having extremely low frequency. Yet the idea is that all of them should be given some space, for despite the bulk we face a closed corpus. The vocabulary of a living language is inexhaustible, especially if technical terms (for instance, the names of diseases and drugs, plants, animals, and mechanical gadgets), regional words, and the slang of all epochs are taken into account. I have touched on the problem of choice above. Every lexicographer selects what he or she finds indispensable. The important thing is not to fill the dictionary with what a smart 19th-century reviewer called obstructive rubbish (here I mean not the words of the *papaya* class but an excessive number of dispensable references), though one man's trash is another man's treasure, as evidenced by the dust heaps immortalized by Dickens.

In my work I call dictionaries like Skeat's and Kluge's dogmatic and those by Feist and Walde-Hofmann analytic. Until the middle of the 19th century all etymological dictionaries were analytic: that is, every entry in them contained a summary (polemical, oftentimes vituperative, or neutral) of what has been said about the origin of the word in question. Reliable criteria for tracing a word to its etymon did not exist. Perhaps a Hebrew look-alike would provide a clue, or the source might be Greek or Latin; when those respectable languages failed to yield desired forms, Old English, Old High German, or Dutch came to the rescue. For a long time etymology remained an exercise in imaginative, moderately intelligent guesswork, and every conjecture, however improbable, aroused interest and excitement. The advent of comparative linguistics and the discovery

of sound laws made the ‘prescientific’ period in the study of word origins obsolete. Skeat would sometimes criticize Wedgwood or quote Skinner (1671) with approval (Skinner made many good suggestions, some of which even the cautious James A. H. Murray accepted), but, other than that, he did not find it necessary to refer to his predecessors. Kluge and his French contemporaries were even more ‘dogmatic.’

Very soon it became clear that sound laws had their limitations. Onomatopoeic, symbolic, and jocular formations; blends, baby words, taboo and anagrams, inexplicable residual forms (*Restformen*) and hybrid forms (*Mischformen*); suspicious borrowings and substrate words, seemingly invulnerable to sound laws, and quite a few others challenged but did not abolish Neogrammarian algebra. Also, within the framework of that algebra solutions vary widely. Any entry in Feist or von Wartburg looks like a record of a military campaign: all scholars swear by sound laws, but their results are different. For this reason, the 20th century returned to the analytic format. The only philology still recycling dogmatic entries is English, so that the post-Skeat English etymological dictionaries are the least substantive in Indo-European. Presses advertise ever new books in which the information culled from the *OED* is presented as particularly ‘fascinating.’ But the *OED*, all its brilliance notwithstanding, is a historical rather than an etymological dictionary and cannot do for the English language what Vasmer did for Modern Russian or Jan de Vries did for Modern Dutch. My project was motivated by the wish to make a first notch in the dogmatic tradition of English etymology.

I think entries in analytic dictionaries should be of unequal length and breadth. In dealing with *brother* and *eight*, it is sufficient to give a succinct overview of the existing theories (those are numerous!) and a list of the main works in which the reader will find further references. My database contains close to a hundred citations for *God*. It does not mean that a hundred conjectures on the origin of this word have been offered (though there have been more than the two routinely repeated in our dictionaries). The entry should contain a summary of the type suggested above and the titles of the original works. Since, as a matter of principle, I look up every word and its cognates with which I deal in all the editions of all the dictionaries, I mention them in my text, just to alert the reader to the fact that nothing new can be found there in comparison with what has already been said. But when we approach *bird*, *boy*, *girl*, *lad*, *lass*, *cub*, *Cockney*, *ivy*, *oat*, *heifer*, *slang*, *witch*, *yet*, *ever*, and their likes (those are some of the words in English featured in Liberman 2008), that is, seemingly isolated words with unclear connections and of debatable structure (native or borrowed?)

simple or compound? ancient or late?), the overview has to be exhaustive. This is where even the opinions of the ‘prescientific’ etymologists matter, for to break the spell laid on those words by their capricious history, we need all the help there is, and in solving such puzzles modern linguists have few advantages over a resourceful scholar who lived three or two centuries ago.

An analytic etymological dictionary does not run the risk of degenerating into an annotated bibliography, because a specialist who has read everything on the origin of a hard word, thought of what has been said about it, and considered numerous proposals will undoubtedly draw conclusions that will be valuable to other researchers. Such an author will be able to reject obviously wrong connections, point out mistakes in earlier reasoning, wherever possible, combine bits and pieces of previous solutions, and weave them into a coherent whole. Some riddles will defy the most strenuous efforts, for it would be naïve to hope that any single person, even endowed with the talents of Jacob Grimm, Antoine Meillet, or Karl Brugmann, can puzzle out all the inscrutable etymologies. We often lack the required data to come to a persuasive result. Also, etymology is both a science and an art. A good deal in it depends on the knowledge of an obscure dialectal form, on an unusual association, and on serendipity. Every serious article on the origin of a difficult word begins with a glance at previous scholarship, and this allows others, regardless of the solution offered, to pick up where the author has left off. An analytic dictionary is called upon to gather several thousand such articles but in congested form. It may take decades to complete, but the effort is worth the trouble.

## 6 Conclusion

The author of an etymological dictionary needs a clear view of the work’s readership, of the vocabulary to be included, of the balance between the most ancient inherited words and those that have emerged in the full light of history, and of the state of the art and of the market. All those points sound trivial, but this impression is false. The methodology of etymological dictionaries and their reception have been discussed too rarely. The genre of the etymological dictionary has not yet been defined.

### > References

Blount, Thomas. (1656). *Glossographia: or, A Dictionary, Interpreting All Such Hard words...* London: Printed by Theo. Newcomb.



*The Century Dictionary: An Encyclopedic Lexicon of the English Language*. William D. Whitney, ed. New York: The Century Co.

EWA = Albert L. Lloyd, Rosemarie Lühr, and Otto Springer, *Etymologisches Wörterbuch des Althochdeutschen*. Göttingen, Zürich: Vandenhoeck & Ruprecht, (1988-).

Faltings, Volkert F. (1996). 'Bemerkungen und Nachträge zu Frank Heidermanns *Etymologischem Wörterbuch der germanischen Primäradjektive* aus der Sicht des Firesischen.' *NOWELE* 28/29, 103-18.

Feist, Sigmund. (1909). *Vergleichendes Wörterbuch der gotischen Sprache*. Halle a. S.: Max Niemeyer. 2nd ed., (1924; 3rd ed., 1939; 4th ed., by W.P. Lehmann, 1986 (The last two editions, Leiden: E.J. Brill).

Kilianus, Cornelius. (1599). *Etymologicum teutonicae linguae*. Antverpiae: Ex Officina Plantaniana, apud Joannem Mortum.

Kluge, Friedrich. (1883). *Etymologisches Wörterbuch der deutschen Sprache*. Straßburg: Trübner. (Kluge took care of the first ten editions of the dictionary until 1924. The latest edition to date is by Elmar Seebold, 2003.)

Liberman, Anatoly. (1998). 'An Annotated Survey of English Etymological Dictionaries and Glossaries.' *Dictionaries* 19, 21-96.

Liberman, Anatoly. (1999). 'An Annotated Survey of German Etymological Dictionaries and Glossaries.' *Dictionaries* 20, 49-89.

Liberman, Anatoly. (2005). 'Scandinavian Etymological Lexicography.' In Henrik Gottlieb, Jens Erik Mogensen, and Arne Zettersten, (eds.). *Symposium on Lexicography XI: Proceedings of the Eleventh International Symposium on Lexicography May 2-4, 2002, at the University of Copenhagen*. Lexicographica. Series Maior 115. Tübingen: Niemeyer, 351-67.

Liberman, Anatoly. (with the assistance of J. Lawrence Mitchell). (2008). *An Analytic Dictionary of English Etymology: An Introduction*. Minneapolis, London: University of Minnesota Press.

Liberman, Anatoly (with the assistance of Ari Hoptman and Nathan E. Carlson). (2009a). *A Bibliography of English Etymology*. Minneapolis, London: University of Minnesota Press.

Liberman, Anatoly. (2009b). 'English Etymological Dictionaries.' In A. P. Cowie, (ed.). *The Oxford History of English Lexicography*, vol. 2. Oxford: Oxford University Press, 269-289.

Malkiel, Yakov. (1975). *Etymological Dictionaries: A Tentative Typology*. Chicago: Chicago University Press.

Minsheu, John. (1617). *Ductor in Linguas. The Guide into the Tongues....* London: Published by the author. 2nd ed., London: John Haviland, 1627.

ODDE = C. T. Onions, ed., with the assistance of G. W. S. Friedrichsen and R. W. Burchfield, *The Oxford Dictionary of English Etymology*. Oxford: Clarendon Press, (1966).

OED = James A. H. Murray et al., (eds.). *The Oxford English Dictionary*. 2nd ed., by J. A. Simpson and E. S. C. Weiner. Oxford University Press, (1992). 1992-, online.

Pokorny, Julius. (1959). *Indogermanisches etymologisches Wörterbuch*. Bern and München: Francke.

Skeat, Walter W. (1882). *An Etymological Dictionary of the English Language*. Oxford: Clarendon Press. 2nd ed., 1884; 3rd ed., 1897; 4th ed., 1910.

- Skinner, Stephen. (1671). *Etymologicum Linguae Anglicanae*.... London: Typis T. Roycroft.
- Sledd, James [and] Wilma R. Ebbitt. [1962]. *Dictionaries and That Dictionary: A Casebook on the Aims of Lexicographers and the Targets of Reviewers*. Chicago: Scott, Foresman.
- UED = *The Universal Dictionary of the English Language*. Henry C. Wyld, ed. London: Routledge & Kegan Paul Limited, (1932).
- Vasmer, Max. (1950-1958). *Russisches etymologisches Wörterbuch*. Heidelberg: Carl Winter.
- Vries, Jan de. (1971). *Nederlands etymologisch woordenboek*. Leiden, etc.: E. J. Brill (three later editions).
- Walde-Hofmann = Walde, Alois. (1938-1954). *Lateinisches etymologisches Wörterbuch*. 3rd ed. by Johann B. Hofmann.
- Walde-Pokorny = Alois Walde, *Vergleichendes Wörterbuch der indogermanischen Sprachen*. Julius Pokorny, (ed.). Berlin: Walter de Gruyter, (1927-1932).
- Wartburg, Walther von. *Französisches etymologisches Wörterbuch*.... Leipzig, Berlin: B. G. Teubner, etc; Basel: Zbinden, (1934-1998).
- Wedgwood, Hensleigh. (1859-1865). *A Dictionary of English Etymology*.... London: Trubner. 2nd ed., 1872; 3rd ed., 1878; 4th ed., 1888.
- Weekley Ernest. (1921). *An Etymological Dictionary of Modern English*. London: John Murray.