The EFL Dictionary Pioneers
and their Legacies

A. P. Cowie

After reading Modern Languages at Oxford and training as an EFL teacher in London, Anthony Paul (Tony) Cowie taught English in Western Nigeria, later becoming an English Language specialist with the British Council. In 1964, he was appointed Lecturer (later Senior Lecturer) in the School of English, University of Leeds. He has since been involved in six major ELT dictionary projects. From 1980-83, and as part of that continuing involvement, he directed at Leeds an OUP-funded Lexical Research Unit. In 1992, he was appointed Reader in Lexicography. He was a founding Executive Board member of the European Association for Lexicography (EURALEX), and continues to serve there and on the Editorial Board of the New OED. He is Editor of the International Journal of Lexicography and author of English Dictionaries for Foreign Learners – a History (Oxford 1999). He organized the first of three International Symposia on Phraseology, and spoke at all (Leeds 1994, Moscow 1996, Stuttgart 1998). His research in the field has led to the establishment of a major bibliography on the EURALEX website http://www.ims.uni-stuttgart.de/euralex/bibweb and the publication of Phraseology: Theory, Analysis and Applications (Oxford 1998). A.P.Cowie@btinternet.com

[This paper is an abridged version of the presentation made at the inauguration of the Israel Association for Lexicography (ISRALEX) at Seminar Levinsky, Tel Aviv on 3 February 2000. The text is available also on http://www.kdictionaries.com/newsletter/kdn8-1.html]

1. Introduction

It should come as no surprise that, in recent years, ELT dictionaries of all kinds have become a vitally important focus of applied linguistic research, innovative design and substantial investment. Learners’ dictionaries are a crucial part of the response to a worldwide demand for English that is constantly expanding. Within the broader field of EFL dictionary development, the monolingual learner’s dictionary (MLD) for advanced students occupies a special place. The use of the foreign language as the language of definition in the MLD — with the actual choice of defining words being carefully controlled — is an acknowledgement that the typical user has a relatively firm footing in the semantic structure of the L2. The broad scope of the listed vocabulary reflects the global importance of English as the language of science, commerce and mass communication. And, not least, the attention given to grammatical words and patterns, and to example sentences, is a clear indication that MLDS are designed to meet the needs of writers as well as readers. The way in which this information is presented (making it, as it were, more “learner-friendly”) has been enhanced, in recent years, by a growing body of research into the use which students actually make of their dictionaries (Atkins 1998, Cowie 1999). And the authenticity of the grammatical claims made about English, and of the examples selected, has been improved beyond recognition by the use, since the early 1980s, of large-scale computer-stored corpora of English, the best known of which are the British National Corpus and the Bank of English.
Because of the marked improvement in the resources available to EFL lexicographers, it is tempting to assume that their products have undergone fundamental changes, and that they now have little in common with the very earliest MLDS – those of the 1930s and 1940s. But as recent research has revealed, the ‘founding fathers’ of the MLD – Harold Palmer, Michael West and A.S. Hornby – had even at that early stage added to the established features of mother-tongue English dictionaries a new set of elements that were inspired by the needs of non-native learners (Cowie 1998b, 1999). With the passage of time, those new elements in turn ‘acquired the status of convention, as the monolingual learner’s dictionary developed into a distinct genre’ (Rundell 1998). It is important to bear in mind, then, that despite the considerable advances of the 1980s and 1990s, especially in the quality of information available to lexicographers and in the way this is presented, those pivotal elements and basic needs have changed very little. So it is worthwhile to go back to those early years, to ask what those original features were, and to enquire into the sources from which they sprang.

The key figures who come into focus are three expatriate Englishmen, teaching and conducting innovative programmes of research in Japan and what was then British India. Harold E. Palmer and his assistant (and later successor) A.S. Hornby, were based in Tokyo, at the Institute for Research in English Teaching (IRET); Michael West served in Bengal. It is a curious fact, by the way, that though these programmes of lexical research led almost inevitably to the MLD – we could in fact, without exaggeration, claim that they gave birth to the learner’s dictionary – the intended goal of all this effort was the development of classroom syllabuses and simplified readers. Palmer, for instance, was not aware until the mid-1930s that ‘a learner’s dictionary’ (his term) would be the most complete and natural end-point of the lexical research he had started up (Palmer 1934).

Here, I shall look at various aspects of this research programme and trace its connections with the earliest achievements in EFL dictionary design. The period covered is from about 1927 to 1942. I begin with the so-called ‘vocabulary control’ movement. Then I move on to consider Michael West and the idea – which was not taken up by Palmer and Hornby – of using a controlled vocabulary for defining. Afterwards, I shall consider the development of phraseological research, and assess its subsequent impact on pedagogical lexicography. In this project, A.S. Hornby played a key role. Finally I shall try to show how the research into collocations and idioms is firmly tied up with the design of dictionary examples.

2. The vocabulary control movement

Harold Palmer had been interested in vocabulary control (or vocabulary limitation) since 1903, when he had run a language school at Verviers, in Eastern Belgium. His interest in vocabulary control is explained by a desire to ease the learning burden of the foreign learner by pinpointing those relatively few words which carried the main weight of everyday communication. This interest was closely linked to the preparation of simplified readers. In fact, Michael West, researching in India, was largely inspired by the desire to produce schemes of simplified readers for schoolchildren (Howatt 1981). The purpose of preparing the limited vocabularies was chiefly educational, then, but the scientific soundness of any word-list could later be tested by using it to simplify other unabridged texts. Words which occurred seldom or never in those texts would probably not be kept, while words which occurred frequently but were not already part of the list might be considered for inclusion. West, as we shall see, applied similar principles when testing successive versions of his limited defining vocabulary for the New Method English Dictionary (NMED) of 1935.

It is not difficult to see now why the word-lists produced in the early 1930s by Palmer, or by Palmer and Hornby working together, should evolve into designs of dictionary entries suitable for foreign learners (Cowie 1998b, 1999). When Palmer was asked by IRET, in 1927, to compile a controlled vocabulary for middle-grade Japanese schools, he was already aware that drawing up a word-list was a more complex affair than producing an alphabetical inventory of spelling-forms, based on frequency of occurrence. He knew, for instance, that the form act could embrace different meanings (‘act’, as in a play, ‘act’, in the sense “pretend”, etc) as well as different parts of speech (‘act’, noun or ‘act’, verb). To run all these differences together, and count simply one form, act, would be to gloss over distinctions that were crucial to the learner (and also to the dictionary-maker). Taking careful account of the meanings and grammatical functions of words, Palmer, West and later Hornby succeeded in producing, from 1930 onwards, what I have called ‘structured lexicons’ (Cowie 1999).

We can see this approach at work in Thousand-Word English (TWE), a word-list begun by Hornby, then refined with Palmer’s help, and finally published in 1937. First, the entries in the list were words, or ‘lexemes’, as they are in most dictionaries. They were not inflected forms, like acts, acting, acted, for instance. Those forms would be arranged inside the entry (in this case for ACT), where the reader would also find a cluster of derivatives (e.g. the nouns ‘actor’, ‘actress’). To see how this works out in more detail in TWE, consider the entry for DRAW, at (1). In this entry, irregular inflected forms are picked out in italic, and basic meanings, at 1 and 2, are conveyed by ‘a picture’, ‘a line’, which are actually collocates of ‘draw’. The derivative ‘drawing’ is placed at the end of the entry.

(1) DRAW [drəʊ], v.
drew [drʊ], pret.
drawn [droyzn], past ppl.
(1. e.g., a picture)
(2. e.g., a line)
drawing [droyzin], n.
(Palmer and Hornby 1937)

Harold Palmer drew on the word-list of TWE when he published his own MLD, A Grammar of English Words (GEW), in 1938. The entry structures of the two works also share certain features, including the positioning of meanings and derivatives. But the dictionary has examples and idioms – these do not appear in the word-list – and it also has some interesting features of arrangement designed to help the user grasp connections of meaning and form across the entry. Notice in the SOFT entry at (2), below, how the definition at soften contains the object nouns ‘leather’, ‘one’s voice’, ‘a person’s heart’. These enable the user to link that word to meanings 1, 3 and 4 of soft, but not meaning 2. These arrangements are of great help when writing, because placing the verb soften in the same entry as the adjective soft helps the writer to avoid repeating the adjective in the same sentence (like this: ‘If it isn’t soft, soften it!’). However, if you are reading, ‘nesting’ words such as soften, softly or softness in the entry for SOFT probably makes it more difficult to locate the individual derivatives. For the reader, they should arguably be placed in separate entries.

(2) SOFT
soft [soft], softer [softer], softest [softest], adj.
1. = not hard
a soft bed.
soft leather [wood, etc.].
The ground is soft after the rain.
Which of the two chairs is softer?
soft to the touch. ...
2. = smooth
a soft hand.
soft hair.
a soft skin.
as soft as silk.
soft to the feel [touch], ...
3. Said of the voice and other sounds =
low, not harsh
a soft voice [sound, etc.].
soft music. ...
4. Said of the character
a soft heart.
the softer side of his nature.
A soften ...
soften leather [one’s voice, a person’s heart,
etc.].
(Palmer 1938)

This excerpt from GEW illustrates only a few of the many connections between IRET research and the earliest learners’ dictionaries. But perhaps it is enough to show that the direction taken by that research encouraged the development of dictionaries that strongly favoured production (encoding). First, extreme vocabulary limitation would give special prominence to structural words (e.g. the, may, in) and heavy-duty verbs (e.g. make, send, bring) and these are, of course, the basic building-blocks of sentence construction. Then, as we have just seen, the Palmer-Hornby approach to lexical analysis gave us an entry structure in which derivatives were clustered around their roots, also with potential benefits for encoding. Of course, within the vocabulary-control project there was also scope for prioritizing the needs of the reader. This could be achieved by adapting the word-lists produced in the course of research to compose controlled defining vocabularies for dictionaries. In fact, there was only one attempt made, in those early days, to develop and utilize a limited vocabulary for defining. It appeared in NMED (1935), which was first and foremost a dictionary for the reader.

3. Michael West’s limited defining vocabulary
The New Method English Dictionary, jointly compiled by Michael West and J.G. Endicott, was the first MLD to be published. The dictionary contained definitions based on a ‘minimum adequate definition vocabulary’. Also in 1935, West published Definition Vocabulary, an account of how the defining vocabulary had been systematically chosen, checked and revised. The research involved compiling a preliminary version of the dictionary, in which a defining vocabulary of 1799 words – eventually to be reduced to 1490 – was used to define 23,898 vocabulary items (West 1935: 34-41).

The definition vocabulary devised by West was to prove enormously influential – it was the basis of the controlled vocabulary used in the first Longman MLD, of 1978 – and in his 1935 essay West put his finger on several of the problems that would later face lexicographers wishing to devise defining vocabularies of their own. The discussion was in fact remarkable for the range of issues that he raised. West identified several of the characteristic weaknesses of definitions in mother-tongue dictionaries – the fondness for defining the known (say, pencil) in terms of the unknown (‘instrument?’, ‘tapering’?), and the tendency to resort to ‘scatter-gun’ techniques, whereby ‘one fires off a number of near or approximate synonyms in the hope that one or other will hit the mark and be understood’, as in the example at (3) (1935: 8):

(3) sinuate tortuous, wavy, winding

One important question that concerned West was how far one could depend on the learner to use prefixes and suffixes as building-bricks. He included some of the commonest prefixes and suffixes (e.g. dis-, in-, -able, -en) in the defining vocabulary, and in the definitions he allowed these to be attached to various words – provided their meanings were regular. So the deverbal suffix -able can be added to drink, eat, read, etc, on the assumption that the user will infer the
meanings of drinkable, eatable and readable (cf. West 1935: 16). In this way great economies can be made.

One of the lessons that West was quick to learn—and that others have not been slow to profit from—was that to arrive at natural and precise definitions of very many words, he had to include in his definition vocabulary a number of very general (‘genus’) words, including behaviour, belief, engine, insect, instrument, metal, noun, quality, relation, science, skill, solid, surface, vegetable. These do not necessarily occur very frequently, but their importance is easily seen if we try to define onion or parsnip without using vegetable, or bee or fly without using insect.

4. Research into collocations and idioms

Harold Palmer had set up a programme of research into phraseology at IRET in 1927—at the same time as it was decided to compile a limited word-list for middle-school children. The result was the first, large-scale analysis of English phraseology to be undertaken with the needs of the foreign learner in mind. The project was directed by Palmer, but much of the actual collecting and classification was carried out by A.S. Hornby. The first detailed findings were published as the Second Interim Report on English Collocations, in 1933. The importance of the Interim Report cannot be overstated. It consisted of a meticulous classification of word-combinations in English, but it also showed how much of everyday speech and writing is in fact made up of ‘fixed phrases’, and it helped pave the way for the strong growth of interest in phraseology in the 1980s and 1990s (Cowie 1998a, 1999).

The expressions classified in the Interim Report were familiar word-combinations (called ‘collocations’ by Palmer and Hornby) that could function as elements in simple sentences. These could be broken down initially into verb-collocations (to toe the line), noun-collocations (a tidy amount), adjective-collocations (as pleased as Punch), and so on, but much finer sub-categories could be recognized within those broad divisions. One sub-category, shown at (4), was the verb-collocation ‘VERB x SPECIFIC NOUN (x PREP x N)’ (no. 31211). This included:

(4) To catch a cold
To entertain a belief
To give notice (x of x N)
To hold one’s tongue
To keep good [bad, etc.] company

The Palmer-Hornby approach was not perfect. For example, the term ‘collocation’ was applied not only to the very large groups of ‘word-like’ combinations—which the Report actually treated—but to proverbs, slogans and catchphrases as well. Nowadays, most phraseologists would restrict the term collocation to word-like combinations such as to catch a cold or a tidy amount. But in doing this they would also be saying that each two-word combination consists of one word used in a normal, familiar, sense (cold, amount), and another word (catch, tidy) whose special meaning is confined to that context and a few similar ones (cf. to catch a chill, a tidy sum). And, of course, they would be implying that these are not idioms—that is, not fixed phrases that are difficult to explain in terms of all of the individual words (Cowie 1998a). Now, neither Palmer nor Hornby recognized this important distinction.

The most obvious practical effect was that, in Palmer’s GEW and the earliest editions of the Advanced Learner’s Dictionary, idioms were not always given the special prominence that they deserve. In the list at (4), the phrase to hold one’s tongue is both fixed and ‘unmotivated’, and is therefore an idiom, while to catch a cold can be internally modified, as in to catch a fever, chill, and is therefore a collocation. Ideally, the former should appear in bold print, but not the latter.

All the same, the Palmer-Hornby approach had a number of enduring strengths. Remember that the Interim Report was a very detailed grammatical classification. This impressed many dictionary-makers and explains the emphasis given in several British phraseological dictionaries to the grammatical treatment of idioms and collocations. Once the notion had caught on of classifying collocations and idioms according to form and structure (see again the heading of the examples at (4)), the natural next step was to provide a more detailed description, particularly by indicating whether idioms could be used in the passive and other ‘transformations’. We find this finer detail in both volumes of the Oxford Dictionary of Current Idiomatic English (1983/1993).

5. Phraseology and the design of dictionary examples

Another benefit of this early interest in phraseology was that it undoubtedly influenced the design of illustrative examples in learners’ dictionaries. Palmer, Hornby and West were all aware that examples would help to show the learner what words (in their various uses) mean. Palmer, though, and later Hornby, were particularly interested in examples which showed the lexical and grammatical contexts in which words typically occurred. In a paper on vocabulary lay-out published in 1936, Palmer illustrated this point by referring to the adjective used to:

(5) to be used to something or somebody
to get used to something or somebody
to be used to doing something
to get used to doing something
(Palmer 1936)

Now, we can argue that these are strictly not examples at all, if by examples we mean instances of performance, whether real (that is, taken from a corpus) or imitated. They are simplifications (there is no grammatical subject) and abstractions (‘something’ stands for a whole range of possible noun phrases). I call these ‘minimal lexicalized
patterns’ (Cowie 1995, 1996), and their value has long been recognized in French and Italian monolingual dictionaries as a basis for imitation and expansion when writing. Palmer referred to them as ‘skeleton-type examples’ and he and Hornby were responsible for introducing them into the MLD (Cowie 1998a).

Hornby, very interestingly, developed for the Idiomatic and Syntactic English Dictionary (ISED), of 1942, later to become the Advanced Learner’s Dictionary, a kind of skeleton clause example which is less abstract than those I have just shown. All the same, it was sufficiently simplified to provide a good model for imitation and expansion. Here, at (6), are some examples of this kind from ISED:

(6) to cut steps in a rock
to cut a figure in stone
to man a ship
to manage a horse

You can see that these are subjectless clauses, that the verb is in the infinitive, and that modification of object nouns (a rock, a ship, etc) is cut to the minimum. And notice how close these examples are to the verb-collocations we looked at earlier (cf. to catch a cold, to man a ship). Some of the ISED examples are indeed part of phraseology.

6. Conclusion
We have had to wait until the last few years of the twentieth century and the first of this to gain a reliable, rounded picture of the early history of the learner’s dictionary and a true sense of how fully it drew on the linguistic research carried out by a small group of Englishmen working at a remote distance from Europe in the 1920s and 1930s. It is only fitting that much of the opening-up of a neglected episode in lexicographical history has taken place in Japan, or has drawn extensively on Japanese sources (Imura 1997, Cowie 1999. Smith 1999a. b). It is an astonishing fact, too, that Palmer, West and Hornby were not aware until about the mid-1930s that the true and natural goal of their research would be lexicographical. In the end, all the key innovative features of the new dictionaries sprang from this programme – West’s defining vocabulary, the verb-patterns of Palmer and Hornby, the skillfully designed examples, and not least the collocations and idioms, an area in which Hornby made such an immensely important contribution. And once these corner-stones had been put in place, they acquired the status of convention, giving the MLD a unique lexicographic character. This is the enduring foundation, upon which all the subsequent developments have been built.

References

A. The early monolingual learners’ dictionaries


B. Other references


Note
(1) The degree of interest is reflected in the intensity of the competition (cf. Herbst 1996). Dictionaries from four major publishers at present share the field: the Cambridge International Dictionary of English (1995), the Collins Cobuild English Language Dictionary (2/e, 1995), the Longman Dictionary of Contemporary English (3/e, 1995), and the Oxford Advanced Learner’s Dictionary (6/e, 2000). The latter is the newest edition of the work initially compiled by Hornby and his associates.
The ASIALEX Second Biennial Conference, Seoul, 2001

ASIAN BILINGUALISM AND THE DICTIONARY

The Asian Association for Lexicography (ASIALEX) was inaugurated during the Dictionaries in Asia conference at Hong Kong University of Science and Technology in 1997. Its first conference was hosted at Guangdong University of Foreign Studies, Guangzhou, in 1999, by the first Asialex president and president of the university, Professor Huang Jianhun, and was attended by two hundred scholars from Asia and Europe. The current president, Professor Sangsup Lee, chairs the organizing committee of the second biennial conference, due on 8-10 August 2001, at the Center for Linguistic Informatics Research, which he directs, at Yonsei University.

The general theme of the conference is ‘Asian Bilingualism and the Dictionary’. This subject has special aspects in the Asian context which have not received sufficient consideration so far, and whose significance is growing very fast in this age of accelerated cross-cultural informative contact between the peoples of Asia and the rest of the world. Other topics related to the dictionary are also welcome. Following is a list of major topics proposed by the organizing committee:

- Corpus Linguistics, Lexicology, Lexicography
- Computational Lexicography
- Lexicographical Processes and Methods
- Language Education and the Dictionary
- Terminography
- Lexicographical Problems in Asian Bilingualism

Renowned scholars and lexicographers will speak at plenary sessions. The languages used will be English and Korean. The deadlines regarding papers are as follows:

- submission of abstracts: 31 October 2000
- notification of acceptance: 30 November 2000
- submission of finished papers: 30 April 2001

The Korean government has cordially agreed to subsidize the conference costs. The registration fee is $45, and early registration by 15 March 2001 is $38. The fee is inclusive of a copy of the proceedings, 3 lunches, stationery items, refreshments, etc. Optional cost for the conference dinner and excursion is 30,000 won (approximately $27 at the current exchange rate).

On-campus accommodation is offered at

- Sangnam Guesthouse: 65,000 won single, 80,000 won double
- Allen Hall Guesthouse: 50,000 won single, 68,000 won double

Since these guest facilities are limited in number, early reservation is strongly recommended. Deposit in the amount equivalent to a night’s accommodation and the registration fee is required, either by an international money order or a bank draft payable to the Center for Linguistic Informatics Research, or by deposit to Account Number 126-64861-02-001 (ASIALEX), Havit Bank Yonsei University Branch, 134 Shinchon-dong, Seodamoon-Gu, Seoul 120-749 Korea.

Abstracts and enquiries may be addressed by e-mail to asialex@lex.yonsei.ac.kr, or by mail to: Center for Linguistic Informatics Research, Yonsei University, Seoul 120-749 Korea.
Tel: +82-2-361-3511, 4197 (+82-2-2123-3511, 4197 after 1 September 2000).