PEDAGOGIC LEXICOGRAPHY
FROM A PEDAGOGIC VIEWPOINT

Presentation at the ASIALEX Conference
Kyoto, Japan
Monday, August 22, 2011, 14.00-14.30
By Ari (Lionel) Kernerman
Kernerman Publishing
www.kernermanpublishing.com

This presentation is about looking at pedagogic lexicography from a pedagogic viewpoint.

Most papers about learners' dictionaries are written from a lexicographic viewpoint, which is a theoretical viewpoint widely espoused by both academics and compilers. This paper is written from a pedagogic viewpoint, which is the practical viewpoint of the language learner and the language teacher.

The aim of this paper is to dispel a few very popular myths regarding pedagogic lexicography. Regrettably, these myths, though broadly accepted as valid by lexicographers, and not generally questioned, constitute erroneously, in my view, some of the basic tenets of contemporary pedagogic lexicography, and are among the principles on which most modern learners' dictionaries are based. Four of these myths are discussed here in detail. Each myth is explained and negated, and alternatives are suggested to replace them.

The first topic discussed is the use of word corpora in pedagogic lexicography. The second topic is the use of defining vocabularies in pedagogic lexicography. The third topic concerns the popular use of monolingual learners' dictionaries by foreign language learners and users. And the fourth topic is the use of bilingual learners' dictionaries by language learners.

1 THE IMPLEMENTATION OF WORD CORPORA AS THE MAIN SOURCE FOR DERIVING EXAMPLE SENTENCES IN COMPILING LEARNERS' DICTIONARIES

Two of the most important features of word banks are that they can serve as data bases for providing example sentences or phrases that are both authentic and typical, and that they can provide information regarding the relevant frequency of the use of words and of their senses.

Both these features of corpora are a major contribution to modern linguistics, as well as to general purpose and special purpose dictionaries. But that does not automatically make them appropriate for learners' dictionaries.

Contrary to what is commonly claimed, sentences derived from corpora are no more authentic than those made up by lexicographers who are native speakers of the language in which they are writing, nor are the sentences more typical. Moreover, the subject matter of corpora-derived sentences is apt to be less relevant than specifically written examples. And the fact that the selected sentence reflects the most common use of that headword may not necessarily be of didactic relevance. On the contrary, it might even be didactically undesirable.

Regarding authenticity, the fact that sentences derived from corpora have been extracted from recognized sources does not make them more authentic than the sentences that I am saying to you right now, which are *not* documented. So this whole idea about being authentic is misleading. Why are the sentences that I am uttering at this very moment, or the sentences that my lexicographers write every day, less authentic or less typical than those in a word corpus? Sentences that are made up by human lexicographers are just as authentic as those gleaned from newspapers or other printed or spoken material – perhaps even more so, because they are intentionally and thoughtfully composed. Moreover, they are likely to be even more common and more typical, reflecting popular use even more.

But the main value of made-up sentences is their didactic value. Made-up sentences are intended not only to explain, but to teach the meaning and the use of a particular word or sense. Good pedagogic lexicographers will see to it that their example sentences:

- do not contain vocabulary that is liable to be more difficult than the word being explained,
- do not contain syntax, verb tenses or grammatical forms that the person looking up that word is liable to be unfamiliar with,
- preferably contain collocations, and

- preferably are also of an explanatory nature, thereby reinforcing what is explained in the definition.

Source material for corpora (books, periodicals, etc.) is apt to be both linguistically and contextually inappropriate for language learners. Besides which, the didactic goal of learners' dictionaries is not to provide, first and foremost, a faithful picture of every-day English, as may be the case with general-purpose dictionaries, but rather, to give easily learnable examples that can be readily understood, used and remembered. Too many corpora-based examples refer to events in British history, politics or institutions, and give confusing or useless information for the learner of English as a foreign language who is primarily interested in learning English for communication.

Human dictionary writers can provide the learner/user with examples that, though they may not be the most frequent, are useful for language-learning purposes. Corpora often reveal that what we consider to be the basic sense of a word is not necessarily its most frequently used sense, which might even be an idiom, a verbal phrase or an expression. But the language teacher's objective is to teach the basic meanings of words first, and then, when that is done, to teach their various meanings. Language learners should not learn by *imitating* what native speakers say before they *learn to understand* the uses of words.

When using word corpora, the editors of learners' dictionaries should be aware of their drawbacks, as it is necessary for dictionary writers to invest considerable time and labor adapting and modifying these corpora-derived sentences to users' needs. *This raises doubts as to whether the required amount of human intervention and the resultant editing and rewriting justify the use of word corpora in the first place.* The editors at Merriam-Webster must have realized this from the start, when they *opted out* of basing the first American advanced-level ELT dictionary, their Advanced Learner's English Dictionary, on a word corpus. But even so, the context of the sentences is life in the USA rather than English as an international language.

So the question I pose to this audience is:

Does the degree of intervention required in order to select good example sentences and to adapt them to the needs of language learners justify basing learners' dictionaries on corpora at all?

My recommendation is to employ only dictionary writers who are native speakers and are experienced foreign language teachers, in order to create sentences that are of the desired length, are contextually, pedagogically, linguistically and culturally appropriate, that will introduce

worthwhile collocations, if possible, and that will even reinforce the meaning – all this in less time, with less effort, and at a smaller cost, than is the case when using corpora.

2 THE USE OF DEFINING VOCABULARIES IN WRITING DEFINITIONS

Defining vocabularies are lists containing words that it is *assumed* that users of a particular dictionary would be familiar with. The compilers of that dictionary strive to use only words from their selected list in writing the definitions. What caused the creation of word lists in the first place was the search for better ways to teach foreign languages.

Pedagogic lexicography cannot be separated from language teaching. Before 1935 there were no pedagogic dictionaries. Until then, dictionaries were considered books for deriving information on the sources and meanings of words, not for teaching languages. The creation of the learner's dictionary was the creation of a new learning tool, which was also the creation of a new teaching tool.

It is interesting to note that the first defining vocabulary preceded the first learner's dictionary. Ogden published his 2,000-word Basic English word-list in 1932, while in 1935 Michael West and James Endicott published the first learners' dictionary – The New Method Dictionary, containing about 7,500 entries. Both of these events constituted a monumental step forward. More than any other publisher, Longman embraced the concept of using only a controlled defining vocabulary in their dictionaries, while later, Macmillan used what they called a restricted defining vocabulary.

Language teachers had long understood that when using difficult words in definitions - words that users might not know - users would have to cross reference for these unfamiliar words. Cross referencing in order to understand a definition is inefficient, in that it is time consuming and frustrating. What, then, could be more user friendly than employing a basic word list that contains only words the user is expected to know? It sounds simple. But this theory contains several flaws.

One is that the so-called word list is not a list of words, but rather, it is a list of the most frequent meaning or use of each of the words in the list. How misleading to cause dictionary users to believe that being familiar with the most frequent meaning of a word means knowing how to use that word, which may have a dozen or more other meanings! Besides, the user may already happen to be

familiar with one of the other meanings of that word, which could cause confusion when reading the definition.

Another flaw in the word-list theory is that 3,000 words, or whatever, may be the minimum amount needed by a lexicographer in order to explain the meanings of 50,000 headwords, but 3,000 is actually a lot of words to know when entering the advanced stage of language learning. On one hand you might even say that if you already know these words you don't need the word-list. And if you don't already know them the word-list won't help you. Also, if I'm looking up a word in the 2,000-word range I am liable not to understand an explanation of that word that uses words in the 2,000-3,000-word range.

The solution is not to use word lists. Instead, dictionaries should be compiled by experienced language teachers who are familiar with what dictionary users are likely to know when looking up a particular word.

3 THE POPULAR USE OF MONOLINGUAL LEARNERS' DICTIONARIES IN FOREIGN LANGUAGE LEARNING

The most important event in lexicography during the past 75 years was the creation of pedagogic or learners' dictionaries. As I said above, the Longman Group started to publish English learners' dictionaries as early as 1935. As far as I know, Oxford started in 1948, Collins in 1974, Heinemann and Nelson both in 1979, Chambers and Harrap in 1980, Macmillan in 1984, and Cambridge in 1995. So that, while the British Empire was disintegrating politically, some British publishers were reconquering the world (including the USA), very much unnoticed, lexicographically - without a navy, and without a shot being fired or a drop of blood being spilled! British domination in the teaching of English-as-a-foreign-language was really consolidated with the appearance of the advanced-level learners' dictionaries. Oxford published its first advanced learner's dictionary in 1948. Thirty years passed before Longman published their Dictionary of Contemporary English in 1978. Collins came out with their innovative COBUILD English Language Dictionary in 1987. Cambridge University Press published their first advanced learner's dictionary in 1995. And Macmillan published their advanced learner's dictionary in 2002. The Big Five ruled the world.

There were others, like Encarta and the BBC, that tried to join the bandwagon of advanced learners' dictionaries, but didn't quite make it. And only recently, just three years ago, after three of

the Big Five had already appeared in American versions, its first American member joined the club – Merriam-Webster! Congratulations!

These are great books, all of them - each in itself, and together as a group. Each of these advanced learners' dictionaries contains between 50,000 and 75,000 headwords. Some of them have smaller, derived editions with half that number of headwords, like Oxford Student's, Longman Active Study, Cambridge Student's and COBUILD Learner's dictionaries. Each one is on a very high standard - pedagogically, linguistically and lexicologically.

Ladies and Gentlemen, the Big Six are all diamonds - they are forever. No electronic device, on-line or off-line, will dislodge these beautifully printed, visually attractive, and tangibly gratifying books. Even though digital data bases provide an alternative, with more versatile dimensions, our printed dictionaries will never become superfluous. They are works of art, and for us lexicographers they are among mankind's great intellectual achievements and aesthetic creations. Like Beethoven's symphonies and Rembrandt's paintings, they will never die. But - and here's the but - each has its own limitations and shortcomings.

But before I go into this, let me deviate for a moment.

When I started teaching English as a Foreign Language in Israel, over fifty years ago, teachers of EFL were required to teach only in L2. For me this was not a problem, because English was my mother tongue. So I taught English-in-English, using what was called The Direct Method. True, the grammatically-based American audio-visual method became the standard teaching method in the 1950s, replacing the British vocabulary-based Direct Method. But teachers remained wary of translation, and rejected the old-fashioned grammar-translation school of teaching. So this is how I would teach:

I would write on the blackboard any new word that arose during the lesson, and next to it, an explanation – in English, of course. I might add that I considered myself to be a good teacher, so I was sure that my explanations were very clear. I would then ask the class that question – the one that every teacher asks – "Is that clear?", or "Did everyone understand?", which is the most stupid question a teacher can ask a class, because naturally no one replies. Of course, I was kidding myself, because it left no way of knowing who understood what.

A few years later I tried a different method. After writing the explanation on the blackboard, I would remain standing with my back to the class, waiting, until some kid in the back row *whispered* out the meaning in Hebrew or Arabic, and then there was a general relaxation of tension. It was

then that I realized that what was needed in language teaching was translation. Because what was missing for language learners was to know the equivalent in their own mother tongue. And indeed, teaching methodologists in the 21st century are justifying controlled use of the mother tongue in foreign language teaching.

So now, after that little anecdote, what is wrong with the six wonderful advanced learners' dictionaries? They are all monolingual.

But how can you translate such gigantic dictionaries into a second language? You can't, and you don't have to. You have to translate only the headword, giving a brief translation or explanation for each sense in the user's mother tongue. This is exactly what I did, calling the method, for want of a better term, semi-bilingual. I will briefly outline here four of the main features of semi-bilingual learners' dictionaries.

One – The teacher can sing and dance in an attempt to explain, but there are always students in the class who might not have grasped the correct meaning of that word or sense when it is explained only in L2. One can never know, and there is no way of knowing, whether or not every single student in the class understood the meaning. By providing the L1 equivalent, the correct meaning becomes known by everyone.

Two – The translation prevents misunderstanding. To misunderstand means to understand – but mistakenly. And when you have misunderstood a word, you don't realize it. We all acquired the meanings of most of the words we know in our mother tongue just by picking them up, assigning to them, as children, our own subjective meanings. How many of these words did we misunderstand at first, only to have to correct this misunderstanding later in life, and perhaps even going through life without being aware that ours is not the correct meaning?

Three – After guaranteeing correct comprehension, the translation reinforces the meaning.

Four – A translation promotes confidence. It dispels frustration and uncertainty. Once you have the translation there is no longer any uncertainty, because now you *know* that you know. How does the process of learning work? How do we remember things? We hang every new bit of information onto an already existing bit of information. By hooking the new word onto its translation it is more likely we will remember it.

My conclusion is that all language learning dictionaries should have brief L1 equivalents for all senses of the headwords, making them semi-bilingual.

4 THE USE OF BILINGUAL LEARNERS' DICTIONARIES IN FOREIGN LANGUAGE LEARNING

If translation is so good, then why not translate the whole dictionary into the learner's native tongue? This question is asked chiefly by publishers in East Asia, in countries where there are large populations, so that translating a 1,500-page dictionary is technically and financially viable. So some of the Big Five were entirely translated into Japanese, Chinese, and even Korean – lock, stock and barrel. This is called, not semi-bilingualization, but full bilingualization.

I regret to say that translating everything, including the part-of-speech, the definition, the examples, and even notes and other learning aids, into the user's L1, is self-defeating. Because once all of this information is provided in translation, the user will not bother to read the corresponding text in English, since it's so much easier to read it in Chinese or Japanese or Korean. But what's wrong with not reading the information provided in an entry, if I understand the meaning of the word I have looked up?

By reading the definitions in English, users learn how to *explain* in English, and by reading the examples in English, they are learning to *think* in English. In no way is it possible to acquire, in the classroom, the skill of explaining things in English, better than by reading a large number of definitions in English over a long period of time. And if the definitions, the examples, etc., are *not* translated, then foreign language learners have no alternative, if they really want to learn the language, but to read the information in L2 – thus practicing the skills of thinking and explaining in English.

SOME OTHER POINTS

I have discussed here only four of the points that I consider to be drawbacks in contemporary pedagogic lexicography. Here, briefly, are a few more:

– There should always be an introduction, preface or foreword that is written for the users, not only one for teachers, lexicographers or reviewers.

– Learners' dictionaries should be culturally oriented to their intended users, most of whom do not reside in the U.K. or the U.S.A. They should abstain from including references that are of a political, religious or social nature.

– Learners' dictionaries should not be encyclopedic. The average language learner has no interest in having to read more information than what is needed to understand lexical meaning, so there should not be more information than what is necessary.

– There should be a set of basic standards for the system of presentation in learners' dictionaries, so that dictionary users can easily go from one dictionary to another, without having to learn how to use each dictionary separately.

– Basic information, that every dictionary user already knows, and is unlikely to ever look up, should be presented more concisely, and not take up so much valuable space.

– The IPA should be replaced with a system that does not require learning additional foreign symbols.