**We don't say it because we don't say it
- a brief look at collocations in English**

Martin Eayrs

_____

Most teachers of English will by now be familiar with the work of Michael Lewis on *The Lexical Approach*, an innovative new look at the language teaching/learning approach. Lewis seems at times to have set out to shake the very foundations of what EFL/ESL teachers take for granted. Nevertheless, what he has to say is undeniably well informed and well researched, and it is certainly worth taking a closer look.

So what is the Lexical Approach ? As Lewis uses it, the word 'lexis' is not just another trendy word for vocabulary, rather it defines a new concept for language teachers, in which 'lexical items' are the 'chunks' of which language is formed. There are different kinds of such chunks, and a well-balanced language course would need to take account of these different types. In this respect not only do spoken and written language differ greatly in their grammars and lexical content but also, within these two areas, there is great variation between, for example, literary, scientific, academic or business language - each of which has different lexical inventories.

These chunks may consist of individual words or groups of words. The two most useful types Lewis cites as being of interest to language teachers are "word-partnerships" (fixed collocations) and "institutionalised utterances". Lewis claims that if a university-educated native speaker may have as many as fifty thousand individual words in his vocabulary he may also have up to twice as many chunks involving combinations of these which occur as pre-formed units.

Collocations are important in many kinds (not all) of written texts and institutionalised utterances are the building blocks of much of 'conversational' language. Lewis claims that it is the failure to recognise, teach and practice these chunks that causes the classic phenomenon of the student who seems to know a lot of English but is helpless when faced with any real 'purpose' outside the classroom.

Traditional grammar practice is based on a 'building-brick' theory in which the learner builds up a grammar deductively from given bits. The lexical approach suggests that multi-word lexical items provide a basis for inductive learning which will lead to mastery of a grammatical system. Lewis talks in terms of 'grammaticalised lexis' rather than 'lexicalised grammar'.

Some other features of the lexical approach that one might mention here include:

1 Collocations are more important than single items if students are to have communicative power

2 In terms of consciousness raising students need to be made aware of the importance of chunks and 'chunking'

3 Error becomes intrinsic to the learning process, being evaluated and used as an aid

to learning

4   Greater importance is given to successful language (in outside-the-classroom use) than to formally accurate language.

5   The language learning metaphor is seen as holistic rather than atomistic; more like growing a plant than constructing a building brick by brick.

6   The importance of grammar is recognised, especially as a receptive skill requiring cognitive involvement, but it is downgraded, grammar becoming subordinated to lexis. Lewis points out that where there is vocabulary without grammar you may have impaired communication but if you have grammar without vocabulary there can be no communication at all.

7   The traditional emphasis on grammar at the sentence level, mainly built around verb tenses, is replaced by concentrating on grammar both above (text & discourse features) and below (word and phrase grammar) the sentence level.

8   The traditional PPP paradigm (Present - Practice - Produce) is rejected as an inadequate description of both language and learning. In its place we are offered a cyclical Observe - Hypothesize - Experiment paradigm.

9   In student activities priority is given to task and process activities.

Clearly this is a limited overview that can do little justice to a whole theory. Teachers who wish to follow it up can read more in Michael Lewis' book [1].

mf ...

---

1   Lewis M, *The Lexical Approach*, LTP Publications, Hove, England, 1994.

Idioms

To define what an IDIOM is (and isn't) we need to first understand the concept of SEMANTIC PROPORTION, defined here as "the reproduction of the same difference of sounds and the same difference of meanings within a different framework".

Meaningful units have semantic proportion, that is they are capable of showing the same formal contrast in two different contexts.

What does this mean ? Look at the following two pairs of sentences

I found the *box*   in the garden   Give me a *box*
         coin   *coin*

Here both the semantic contrast of <u>box</u> - <u>coin</u> and the phonological contrast of /bɔks/ - /kɔın/ are reproduced in two different contexts. This can be done with any meaningful unit.

Compare   I found the b- *-ox*   I saw a f- *-ox*
                              -in   *-in*

Here the contrast is not the same in each context. Although there is a similar contrast of sound (in /ɔks/ and /ın/) there is no similar contrast of 'meaning'. Therefore *-ox* and *-in* are not meaningful units. We disregard here the homophones 'ox' - the animal - and 'inn'.

And to perform this test, both frames must be completely different, so in

The milk is *boi-*   -ling   The water *boi-* -led
               coo-   *coo-*

boi- & *coo-* are not meaningful units because of the /l/ sound which is common to both.

A meaningful unit carries its meaning with it. Put at its simplest, it's a bit of sound trailing its own meaning. In the examples given above, *-ox* and *-in* don't carry any meaning, but *box* and *bin* do. They can be contrasted for sound and meaning and, meeting the criteria for semantic proportion, are therefore meaningful.

Consider now   *Mon*-day /*Tues*-day

where the elements *Mon* and *Tues* cannot be reproduced in any other context to give the same contrast of sound and meaning,

and   *bil*-berry/*straw*-berry

where again we can reproduce *bil*- and *straw*- to give the same contrast of sound, but not of meaning.

Now look at   *dis*-mount/*re*-mount
                                   dis-embark/*re*-embark

where we do have the same contrast of sound and meaning (*dis*- having the idea of 'off' and *re*-

'back on'). So *dis-* and *re-* are meaningful units here (but not, note, in a pair like *dis-* appoint/*re-*appoint where there is no such contrast).

Test for Semantic Proportion

Consider the word *leg* in these sentences:

> I broke my *leg* when skiing.
> John is pulling Bill's *leg* about his hairstyle.

contrasted with

> I broke my *arm* when skiing.
> John is pulling Bill's *arm* about his hairstyle.

If we test each element for semantic proportion we can see that

> *John*
> Tense   ...... all have meaning
> Aspect
> *Bill*

What doesn't pass is (NP) . . . . *pull* + Tense + (NP) *'s leg*

and this kind of thing is what we call an idiom. The meaning of an idiom cannot be deduced from its components, as the parts themselves have no meaning.

Our definition of an idiom here is: "a group of words with a set meaning that can't be calculated by adding up the separate meanings of the parts"

Idioms are both semantically & syntactically fixed or frozen.

> *John pulled Bill s leg*   is ambiguous (theoretically at least).
> *John pulled Bill's left leg*   alters meaning, only one possible interpretation.

Pieces of the idiom cannot be interfered with:

> *He is driving me round the bend*
> * *He is driving me round the left hand bend*

but sometimes a humorous variation can lead to a new idiom, in which the old one is transparent: e.g. 'round the bend' can become 'round the twist', 'to take the Micky' can become 'to take the Michael' or 'to extract the Michael'. All this is a little arbitrary and dependent on native speaker intuition and acceptance!

Consider now: *To hit the nail on the head*
                              To kill two birds with one stone

These do depend (to a certain extent) on the meaning of some of the elements. We could say for example

*To **strike** the nail on the head*
*To **shoot** two **rabbits** with one **bullet***

where the meaning can still be extracted. We can call these cases of limited transparency DEAD METAPHORS and syntactically these are much freer than idioms.

Idioms and collocations

We must be careful to distinguish between IDIOMS and COLLOCATIONS.

We can say: (c) The man earnestly wanted them to leave (3)
        The man earnestly wished them to leave (3)

but look at   (b) The man badly wanted them to leave (3)
       * The man badly wished them to leave (7)

The reason that the last sentence is unacceptable is that it breaks the rules of collocation; the words simply do not go together. There is a 'co-occurrence restriction'.

A further example: *indelible ink* is a strong collocation, but transparent
         red herring is an idiom, and opaque

Collocations are "phrases which are regularly repeated, come readily to the mind of the native speaker and are relatively fixed".

Why we do accept 'The man badly wanted them to leave' but balk at * 'The man badly wished them to leave' ? If 'badly' is replaced with 'earnestly', both sentences are normal. Why do builders not produce a building or authors not invent a novel ? After all, they do invent stories and plots. For no reason, as far as dictionary definitions of words are concerned. We don't say it because .... we don't say it.

Collocations are arbitrary, and generally speaking unpredictable. How can the non-native speaker know that we *make an estimate* but we don't * *make an estimation*; that we *commit treason* but we don't **commit treachery*; that we can both *commit fraud* and *perpetrate fraud* but although one can *commit suicide one* can not * perpetrate *suicide* ?

Native speakers are often at a loss too with some less commonly used expressions. Try these [2] on your native speaker friends - they may not come up with them automatically although once given the answer they will almost certainly say "Oh, yes, of course !"

> The jury .... an acquittal (*brought in*)
> to be .... by Caesarian section (*delivered*)
> he .... a cartwheel (*turned*)
> to .... a circuit-breaker (*trip*)
> to ....copyright (*infringe* or apply *for, grant, hold, claim, secure*)

There is often a difference between US and UK collocational usage. For example, UK speakers generally *have a bath* while US generally *take a bath*. While most speakers of English *make decisions*, UK speakers also *take* them.

---

2   After Benson M, Benson E & Ilson R, *The Combinatory Dictionary of English,* John Benjamins, Amsterdam, 1993.

The focus of the verb is relevant too: think of the word *copyright*. The patents office may *grant* or *register copyright*, but the author or the publisher *holds* or *secures* it.

Other determining factors

Other factors can be involved in determining word proximity and co-occurrence restrictions. Consider two sentences stressed as follows:

a)   I **don't** find her   **all** that   at**trac**tive

b)   I don't **find** her   all **that**   at**trac**tive

Now look at the collocation possibilities with the following paradigmatic variations

a)   I **don't** find her   **all** that   at**trac**tive
                              very
                              par**tic**ularly
                              es**pec**ially
                              * extr**em**ely

b)   I don't **find** her   all **that**   at**trac**tive
                              very
                              par**tic**ularly
                              es**pec**ially
                              extr**eme**ly

where a) is more or less a bald observation whereas b) seems to contain an element of rejection of a suggestion (that she is attractive) [3]. The word 'extremely' collocates with 'attractive' in one stress/intonation group (with one pragmatic context) but not in another, giving us us a fascinating glimpse of the variety and complexity of factors in play.

mf ...

---

3   Remotely, a native speaker may use the a) version with 'extremely' as a complaint about unfulfilled expectations - e.g. he had requested an extremely attractive model from an escort agency and is explaining why he is unwilling to accept her services on the grounds that she is described in the book as 'extremely attractive' and he does not agree with this description. The connotations are in any case different and the pragmatic factors informing this interpretation merely emphasize the complexity of factors in force).

The range and variety of collocations is enormous. Not all persons will agree with every judgment of acceptability but the important thing is that such judgments can be made. It is our experience of expressions repeated over and over in given circumstances that makes for collocation (in addition to providing us with the regularities of our grammar), and it would be remarkable indeed if that experience were uniform all over the English-speaking world.

Collocations can be grammatical or lexical (and, as we have seen, conditioned by other factors). For the remainder of the workshop examples of each were given according to the following classification [4]:

Grammatical examples

1.   Article binding
2.   Fixed phrases (often a step away from clichŽ)
3.   Word + word class/category.
4.   Miscellaneous complex phrases.
5.   Restructured phrases.

Lexical Examples

6   Noun + adjective
7.   Preposition and nouns.
8.   Verb + Noun
9.   Noun + verb
10   Adverb + Adj
11   Verb + adverb
12.   Unit word + noun

Finally, a number of practice exercises was done by the group in order to demonstrate and apply this classificatory system.

<div align="right">

Martin Eayrs
August 1996

</div>

---

[4]   after Benson M, Benson E & Ilson R,1993, *Ibid.*