Michael A. Lewis
Hove (UK)

Towards a Lexical View of Language – a Challenge for Teachers

Notre compréhension de la nature du langage et de la manière dont notre cerveau stocke et conserve l'information a connu des changements récents qui rendent également nécessaires des changements dans l'enseignement des langues. La division traditionnelle entre grammaire et vocabulaire, en particulier, n'est plus pertinente: une part importante de chaque langue est en effet constituée de phrases préfabriquées – qui offrent des possibilités plus ou moins grandes de variation; ces items lexicaux préfabriqués facilitent les processus de traitement chez le lecteur ou l'auditeur qui peuvent dès lors se concentrer sur l'information nouvelle contenue dans le message.
C'est pourquoi le vocabulaire apparait aujourd'hui bien plus important que la grammaire. Et les enseignants devraient avant tout s'efforcer d'aider les apprenants à découvrir les multiples possibilités de collocation *de la langue (les "chunks") plutôt que de gaspiller le temps des apprenants en leur infligeant trop d'exercices de grammaire. (réd.)*

Introduction

Do you think you would teach languages the way you do if you had not been trained as a language teacher? How closely do you think the activities you use in the classroom resemble those of learners who learn a language 'naturally' by being immersed in situations in which they need to use the language. I suspect many of your classroom activities focus on vocabulary ('new words') or grammar, particularly the 'tenses', while I suspect 'natural' learners are preoccupied with only one thing, the apprehension and creation of meaning. The purpose of this article is to argue that modern research in corpus linguistics strongly suggests classroom practice needs to move away from vocabulary and grammar and towards lexis and a new way of looking at text through lexical eyes precisely because that is how meaning is created. First, however, some background.

Why do you teach languages they way you do? Habit? Because it is how you were taught at school yourself? Just following the coursebook? Do you give your learners explicit grammar rules? Encourage them to record "new words" in vocabulary lists? Do transformation exercises *('Put the following into the passive'*)? Do you work with the 'tidy' language of typical language learning materials (or do you prefer to use real, naturally occurring English from, for example, magazines or the internet)? Reflective teachers base their teaching on their more or less explicit beliefs about the nature of language and the nature of learning. Any new understanding of either should prompt changes in what happens in the classroom. This article will consider the implications of recent changes in our understanding of the nature of language and the way our brains store and retrieve language. It will suggest that all of the activities mentioned above may actually make learning more difficult, and that these activities need to be replaced by other, more efficient, ones. This will only happen, however, if teachers understand **why** any such changes are desirable.

It is only comparatively recently that the phenomenon that is language (as opposed to languages) has been studied. The earliest studies were often of Latin which is highly inflected. One consequence has been that in Europe the study of languages has always placed a disproportionate value on what has traditionally been called "grammar", concentrating on the forms of the verb. Until the last 25 years the study of language was essentially intuitive, although certain pseudo-scientific claims were made. In truth, it is only since the advent of computer corpora of naturally occurring language that properly scientific study has been possible.

We need to note too the unhelpful influence, particularly in the US tradition, of Chomsky who claimed that a language was "the possible sentences of the language". This profoundly misguided definition served to re-emphasise the importance of the grammar of individual sentences. Computer studies show that many things which **could** occur do not in fact do so or are extremely unlikely (*'We've had our downs and ups.'; 'Good morning, Gentlemen and Ladies'; 'Hello, I*

haven't seen you for seven and a half months'; ' I like weak tea but powerful coffee.' Any comprehensive account of a language must explain both what is possible and what, although perhaps possible, is highly unlikely.

Until 30 years ago it was generally agreed that language could be divided into 'grammar' and 'vocabulary'; this analysis, with which most teachers will be very familiar, results in a "slot and filler" approach to grammar practice. Once again, recent computer studies of naturally occurring language demonstrate that this approach is deeply flawed, and makes language learning unnecessarily difficult (and, one might add, for many less able students, particularly intimidating).

It is worth mentioning that most corpus linguistic research has been done on English, and all references in this article are to English; it is evident, however, that the characteristics of language referred to are typical of all languages. Some years ago, I presented many of these ideas in Washington to an audience of teachers of, among other languages Chinese, Farsi, Finnish, Korean, Arabic, Hebrew and other language which are thought of as 'far' from English. All the teachers, irrespective of their own subject, recognised the ideas, some even claiming their language was 'more lexical' than English. Although I refer to a few English-specific examples, the principles are, I believe, universally applicable.

Some research findings

So what have studies of computer corpora of naturally occurring language shown? Here are some of the key insights:

1. Few patterns or 'rules' are totally fixed; almost any statement we can make about linguistic patterning is about tendency or probability: there are few certainties. John Sinclair, one of the fathers of corpus linguistics puts this very clearly:
   grammatical generalisations do not rest on a rigid foundation, but are the accumulation of the patterns of hundreds of individual words and phrases. [*Corpus, Concordance, Collocation, p100*]
2. Vocabulary and grammar are not separate categories, but are inextricably linked. Separating them creates confusion not clarity. Sinclair again:
   The evidence now becoming available [c1990, it is now confirmed, MAL] casts grave doubts on the wisdom of postulating separate domains of lexis and syntax. [*op cit p 104*]
3. Most language consists of strings of pre-fabricated phrases (lexical items) which exhibit more or less possibility of variation - *on the other hand* does not permit the (grammatically acceptable opposite) *on this hand*, but requires the fixed item *on the one hand*, which must occupy an earlier position in the discourse; *a week on Thursday* can be varied as *a week on [day]*, but *a [month/year] on Thursday* (again grammatically acceptable) seem most unlikely.
4. If we think in traditional 'vocabulary' and 'grammar' terms, it is the word which determines the patterning around it – the 'grammar'- not the pattern into which a word is slotted. In short, real language operates in exactly the opposite way from the traditional (now mostly discarded) slot-and-filler approach which was in general use 20 or so years ago.
5. Patterns which are typical of one kind of English (genres), may be rare or not occur at all in other genres. This is much more subtle than the simple difference between spoken and written English. The most comprehensive grammar of any language yet published – The *Longman Grammar of Spoken and Written English* (LGSWE) compares 4 genres: conversation, news, fiction and academic writing. There are many surprises: questions are comparatively rare in any form of written English; the ten most common (full) verbs are ubiquitous in conversation but extremely rare in academic writing and the patterns in which they do occur in writing are mostly radically different from those typical of the use of the same common words in spoken genres, indeed, the typical patterns of many if not most individual words are highly genre-specific. The overriding message is that truly general 'rules' are very rare but correspondingly much more language consists of lexis - multi-word mini-patterns - than anyone ever suspected. John Sinclair perceived the most radical implication of this nearly twenty years ago:
   The overwhelming nature of this evidence leads us to elevate the principle of idiom from being a rather minor feature, compared with grammar, to being at least as important as grammar in the explanation of how meaning arises in text. *[op.cit. p112]*

The Principle of Idiom

We need to look in detail at the idea of 'the principle of idiom', sometimes now called 'idiomaticity', which has almost nothing to do with the traditional 'idioms' in language teaching. The traditional term was used for items which were thought to be fixed, (usually) colourful, non-literal expressions (famous examples are, of course, *It's raining cats and dogs* and *He kicked the bucket the other day*, neither of which, incidentally, have I ever heard used by a native speaker except in discussions of idiomaticity!). These expressions were seen as typical of speech, but rare to non-existent in serious academic writing. Idioms in this sense were left to advanced level courses and learners were often advised to avoid using them as, slightly

mis-used, the effect could often be comical or ridiculous.
One other group of fixed phrases was recognized – clichés – but the term is pejorative, and again, learners were discouraged from using them. Indeed, George Orwell, no less, had this to say:

> This invasion of one's mind by ready-made phrases (*lay the foundations, acquire a radical transformation*) can only be prevented if one is constantly on guard against them, and every such phrase anaesthetizes a portion of one's brain.

Interestingly, as I note in *Teaching Collocation*, Orwell himself cannot avoid lexical items such as *bitter winter, carried on as best they could, the outside world* even in a work with such unusual subject matter as *Animal Farm*.
Far from anaesthetizing the brain, the use of prefabricated lexical items facilitates processing by the listener/reader, allowing them to concentrate more easily on the new information contained in the message.
As we noted, Sinclair claimed 'the principle of idiom is **at least** as important as grammar'; I suggest that we now understand that it is almost certainly **more** important than grammar in the creation of meaning. Now we have reached a challenging point – language is essentially about the creation and exchange of meaning; we do not speak in order to make examples of the present perfect; we never form a sentence in our head and then 'transform' it into the passive, or turn it into so-called 'reported speech'. Language produced in such classroom activities scarcely merits being described as 'language' at all, indeed, some linguists describe it – accurately – as 'language-like behaviour'. Teachers who use such activities in their classrooms need to consider carefully if they can justify them as helpful to their learners. I do not believe they are; indeed, I think they make learning artificially difficult and actually constitute a barrier to efficient acquisition.

Truly general 'rules' are very rare but correspondingly much more language consists of lexis - multi-word mini-patterns - than anyone ever suspected.

Traditional idioms
We may note *en passant* that, far from being fixed, traditional idioms rarely occur in their dictionary form; native speakers use the 'base' form and vary it slightly or even in almost unrecognizable (of course they must remain recognizable) ways. Here are some examples from my own corpus:

> When Diana, Princess of Wales, walked into a room all heaven broke loose. [Daily Telegraph, 29/8/98]
>
> An awful lot of blood has flowed under the bridge… [Robin Cook, then Foreign Secretary, explaining a change in British Government policy on Kosovo, TV News 49/4/99]
>
> Sound like the wag the dog syndrome to me. [Alex Thompson, interviewing on Channel 4 News 16/12/98]
>
> He's a man of a couple of medium-sized ideas. [Political commentator on the election of George W Bush 12/00]

Apart from being witty and amusing, these examples are also very revealing in several ways. Firstly, they give the lie to Orwell's assertion; 'bending' a fixed phrase helps the speaker/writer create meaning, but this depends on both speaker/writer and listener/reader knowing the underlying, supposedly 'fixed', expression. Secondly, we see that there is no clear distinction to be drawn between traditional idioms – *all hell broke loose, the tail is wagging the dog* – and collocations – *a big idea* would not traditionally be thought of as an 'idiom' but this is precisely what computer corpora have revealed – a huge amount of all the language we use is produced from our memories either as fully-fixed prefabricated items, or as slight variations on (almost)-fixed items and this kind of language is all-pervasive, in speech, fiction, and serious academic writing.

Pedagogical implications
What are the pedagogical implications of all this? In short, teachers need to pay much more attention to 'vocabulary' – we will consider what that term means in more detail below - and much less to traditional grammar; as my erstwhile colleague Jimmie Hill observes:

> Spending a lot of class time on traditional EFL grammar condemns learners to remaining on the intermediate plateau. [*Teaching Collocation, p 68*]

If teachers are to make the necessary change of emphasis we need to move the idea of 'vocabulary' away from individual 'new words' by developing strategies which encourage teachers and then learners to look at texts through new, lexically aware, eyes. It is wholly unhelpful to ask learners 'Are there any words you don't understand'. Firstly, this focuses on what is unknown and is correspondingly intimidating for learners, particularly those who are less able. Secondly, and perhaps more importantly, it is difficult if not impossible to acquire anything which is completely unfamiliar; what is most likely to be usefully acquired is language which is already 'on the edge' of the learners' knowledge – perhaps items they recognize but have never used actively, or words where they know one or two simple collocations of a word which has a large collocational field. Peter Skehan makes the point thus:

> It is proposed here that very often the pedagogic challenge is not to focus on the brand new, but instead to make accessible the relatively new. [*A Cognitive Approach to Language Learning, p 139*]

One obvious strategy is to explore

other collocational possibilities, encouraging learners to extend their ability to actual use half-known words. (This idea is extensively developed in George Woolard's new series *Keywords for Fluency*.) One simple way of 'getting into' the text is to ask learners to underline all the nouns they can find in a text or a selected part of the text. They then search for any other words which they think form part of the 'chunk' which contains the noun. Depending on the age and level of the learners, teachers may wish to offer varying degrees of guidance – first search for any adjective immediately in front of the noun, then any verb which comes in front of the noun (verbs after nouns are much less significant, as we shall see shortly), then any small words – articles, prepositions - which are part of the chunk. Try this activity yourself; at first you may see relatively little but with practice you will start to see more chunks, and the chunks you find will be bigger. (It is, perhaps, counter-intuitive for teachers who have been trained that breaking things down makes them easier to learn, but I suggest breaking down is actually frequently a source of later problems. If you learn two separate words and need to put them together to make a phrase, this is obviously more difficult than learning a phrase which you then break down; clearly, too, breaking down closely resembles the way in which we all learned our first language.) The objective, then, is to encourage learners to see, record (and perhaps translate, thereby addressing the familiar 'Don't translate word-by-word' problem) the **largest** chunks they can find based on partly known key words, the exact opposite of focusing on individual new words.

Most importantly, with practice you will start to keep grammatical words such as articles and prepositions with the central meaning-carrying noun; you will have (started to) discover word-grammar, the central element of language which is the exact opposite of the traditional vocabulary/ grammar dichotomy.

It must be emphasized that the Lexical Approach (extensively outlined and discussed in *The Lexical Approach* and *Implementing the Lexical Approach*) does not distain or dismiss 'grammar', it simply invites teachers to revise their concept of grammar; transformation exercises are dismissed as nonsense, traditional rules – often misleading over-generalisations - and practices of the so-called tenses are seen as very small parts of a balanced syllabus; word-grammar is given a new, high priority and one immensely important new grammatical feature which we have not met so far is given the importance it deserves – a feature which is wholly unrecognized in conventional analyses and which hardly features in any coursebooks.

If teachers of English reflect, they will recognise that traditional vocabulary teaching concentrated largely on nouns, while grammar teaching concentrated on the structure of the verb phrase. Corpus studies show that the second most common word in all genres of English is the apparently insignificant word 'of' (it is about 2% of **all** English text) It turns out this is central to the building of noun phrases which are often the grammatical subject of English sentences (hence the importance of nouns which **precede** the main verb, they are central to noun-phrase word-grammar): *One **of** the principal causes **of** the Second World War; The death **of** President Kennedy; New members **of** staff.* It is impossible to write English well without the ability to use *of*-expressions but it has never formed more than a minute part of traditional syllabuses (*a pint of milk, a bar of chocolate*!). (Woolard's *Keyword* series has extensive practice of *of*-expressions) So important in the structure of English is *of*, that the monumental LGSWE mentioned earlier takes some 40 pages to discuss *of*-patterns. Teachers of languages other than English may care to ask themselves how much of their classtime is devoted to teaching the structure of the noun-phrase; it is highly likely that it will hardly feature and, as with English, the verb-phrase and individual nouns will tend to predominate.

So, modern linguistic research shows beyond doubt that language is much more lexical – built of fixed or partly-fixed prefabricated chunks – than we ever suspected; that many supposedly useful and/or typical patterns beloved in traditional grammars are rare or non-existent in naturally occurring language; that there are patterns central to the good use of English which were not recognized until these recent studies. The evidence is conclusive. It leaves one huge question – are teachers going to change what they teach and how they approach texts? It is only teachers who can implement the necessary changes in the classroom. Will they – you – rise to the challenge?

References

BIBER, D*et al*. (1999): *Longman Grammar of Spoken and Written English*, Longman.

HILL, J. (2000): *Revising priorities: from grammatical failure to collocational success*, in: *Teaching Collocation.*

LEWIS, M. (1993): *The Lexical Approach*, LTP.

LEWIS, M. (1997): *Implementing the Lexical Approach*, LTP.

LEWIS, M. (2000): *Teaching Collocation*, LTP.

SINCLAIR, J. (1991): *Corpus, Concordance, Collocation*, Oxford University Press.

SKEHAN, P. (1998): *A Cognitive Approach to Language Learning*, Oxford University Press.

WOOLARD, G. (2004, 2005): *Key Words for Fluency, (Pre-intermediate, Intermediate and Upper Intermediate)*, Thomson ELT.

LTP books are now published by Thomson ELT.

Michael A. Lewis

is the author of some 30 ELT books for both students and teachers. He has lectured in over 30 countries. He is now semi-retired and devotes most of his time to opera not ELT!