The Lexical SyIlabus

A new approach to language teaching

Dave Willis

COLLINS E.L.T
London and Glasgow

Contents

Collins ELT
8 Grafton Street
LONDON W1X 3LA

COBUILD is a trademark of
William Collins sons & co Ltd

First published 1990

10987654321

Introduction

There is general agreement nowadays that we learn a language best by using it to do things, to achieve outcomes. Communicative activities involving games playing and problem solving have become a more and more important part of the language teacher's stock in trade over the last fifteen years or so. Some writers (see, for example, Maley and Duff, 1978) display great ingenuity in devising such activities and there is a wealth of supplementary material which exploits these activities. Yet in spite of this virtually all coursebooks rely on a linguistic syllabus which 'presents' the learner with a series of linguistic items.

It seems that communication is good fun and well worthwhile for a bit of variety, but that the serious business of language learning needs to have a firm grammatical basis resting on the assumption that the grammar of the language can be broken down into a series of patterns and reconstructed in a way accessible to the learner. Even coursebooks based on a notional-functional syllabus specification, which take units of meaning as syllabus items, still rest on a methodology which 'presents' learners with a series of patterns. The notionalfunctional syllabus is communicative in that it tried to specify the syllabus in terms of meaning, in terms of what was to be communicated. But the methodology which realises the notional-functional syllabus is little different from the methodology which realises the structural syllabus which it seeks to replace. Both depend on a three part cycle of presentation, practice and production.

My dissatisfaction with this methodology has a theoretical basis but it is strongly reinforced by experience in the classroom. The theoretical base draws on the work of people like Prabhu (1987) and Rutherford (1987) both of whom point to the glaring inadequacy of pedagogical grammars. They argue that we cannot begin to offer anything like an adequate description of the language on which to base a pedagogical grammar. Given this, our only recourse is to depend on the innate ability of learners to recreate for themselves the grammar on the basis of the language to which they are exposed.

The conclusion is similar to that drawn by interlanguage theorists like Corder (1967) and Selinker (1972) and classroom resear~hers like Ellis (1984). Teachers and researchers have been aware for many years that 'input' does not equal 'intake', that what teachers claim to be teaching bears only a tenuous relationship to what learners are actually learning. But in spite of this, coursebook writers continue to act on the assumption that language can be broken down into a series of patterns which can then be presented to learners and assimilated by them in a predictable sequence. It does not seem to worry people a great deal that this assumption flies in the face of our experience as teachers.

My experience in the classroom, like that of all teachers I suppose, has seen both failures and successes. On the one hand I found that students often failed to learn what I thought I was teaching them. On the other hand most of them showed an ability to transcend the limited language which I had so carefully presented to them. It was clear to me that my efforts to present the grammar of the language met with very limited success, yet in spite of this mv students' English was improving. It is encouraging to know that so much learning is taking place in the classroom. It is sobering to realise just how little control the

teacher has over what is being learned. The obvious conclusion to be drawn from this is that students learn a great deal directly from exposure to language through reading and listening, without the need for the teacher to impose a description on what is learnt.

One of the most plaintive cries in any staffroom goes along the lines of "I've taught them that so many times, and they still get it wrong." There is overwhelming evidence from my experience as a teacher that teachers have little control over what is actually learnt and reproduced in spontaneous language use. How many times, for example, do we teach the distinction between the present simple and the present continuous before students begin consistently to get it 'right'? It usually takes them a long, long time. Could this be because it is not the controlled presentation which does the trick but rather constant exposure over a period of time? Could it be that students learn in the controlled environment of the language classroom not because language is presented to them, but because they are constantly exposed to language? And if this is the case should we not be looking to methodologies which maximise meaningful exposure to and use of language?

Taking meaningful exposure as a starting point it is possible to develop an approach to language teaching which takes advantage of the learner's natural tendency to make sense of language and to learn for himself. In order to take full advantage of this approach, however, two other things must be done. First a methodology must be defined which encourages the learner's ability to learn. Teachers need to encourage learners to look critically at language and to recognise the need to develop and refine their language code in order to achieve their communicative aims. Secondly we need to look carefully at the kind of language to which learners are exposed. Random exposure is of little value. Exposure must be organised.

What should be aimed at, is exposure that is organised in three ways. First the language that learners are expected to understand and produce should be graded in some way so that learners do not face such difficulties and complexities at an early stage that they become demotivated. Secondly the language they are to be exposed to should be carefully selected so that they are given not random exposure, but exposure to the commonest patterns and meanings in the language - the patterns and meanings they are most likely to meet when they begin to use language outside the classroom. Thirdly there should be some way of itemising the language syllabus so that it should be possible not simply to expose students to language, but also to highlight important features of their language experience, and to point to what language we might reasonably expect them to have learned from their experience.

The first of these problems is relatively easy to surmount. It is not too difficult to design tasks which involve a meaningful use of language but which can still be handled by learners who have relatively little control of the language - the kind of learner who is often referred to, somewhat unfortunately in my opinion,as a remedial or false beginner. Tasks which have a clear outcome and involve the exchange of highly specific information can be made accessible to false beginners. As I have said, such tasks have been used as supplementary material for many years.

The second and third factors were, until recently, more problematic. When

my wife Jane and myself were asked by Collins ELT in 1983 to begin to research and write a series of coursebooks, the Collins COBUlLD *English Course*, we began to ask ourselves a number of related questions. How could we identify the commonest patterns and meanings in English and how could we highlight these for students? Obviously many of them are covered in most elementary courses. The verb be and its forms and most of its uses would obviously come high on any list as would prepositions of place. But other equally common forms such as the passive voice and modal verbs are traditionally left until much later. Also, we discovered as we became more involved in the research that a number of important words and patterns are often omitted altogether. Words like problem, solution, idea, argument and thing are commonly used with a noun clause introduced by that to structure discourse. It is difficult to get very far in speech or writing without them.

And what about items which seemed to take up far too much time in elementary courses, items like the present continuous used to talk about what is happening here and now? Apart from a traditional belief that certain patterns are 'difficult', there seems to be little objective reasoning behind the selection and ordering of items. We were soon to find evidence that a syllabus based on these established values was likely to be highly uneconomical.

But how could we go beyond the traditional approach to itemising and organising a syllabus? Given the range of language experience which is bound to come from exposure to a series of tasks which are graded for difficulty but not otherwise linguistically graded, how would one choose which elements of language to highlight? How would one decide which items to specify as part of an efficient learning programme? Perhaps the most convincing attempt in the field so far was the Council of Europe Threshold and Waystage Syllabus. But this was ultimately a very subjective piece of work. It took as its basis the intuitions of scholars and teachers. It did not rest on an analysis of actual language use.

In the mid-1980s a number of things began to come together. After years of teaching English as a foreign language, a period of work as a teacher and teacher trainer in the second language environment of Singapore had forced me to look more closely at methodological issues, particularly the relationship between accuracy and fluency (Willis and Willis 1987). This helped to formalise a communicative approach to ELT and to identify some of its important components. The writing of the *Collins* COBUILD *English Course* provided us with the opportunity to put these methodological insights to work.

The coursebooks were to be a part of the COBUILD research project in lexical development, a major computing and publishing venture involving cooperation between Collins and the English Language Research Department at Birmingham University.

The first part of this project had involved the assembly on computer and subsequent analysis of a 7.3 million word corpus (later extended to over 20 million words) of spoken and written English. It was proposed by John Sinclair, Professor of Modern English Language at Birmingham and Editor-in-chief of the COBUILD project, that this computational analysis should provide the basis for a new coursebook syllabus, a lexical syllabus. Sinclair advanced a number of arguments in favour of the lexical syllabus, but the underlying argument was to do with utility and with the power of the most frequent words of English.

vi The Lexical Syllabus

The 700 most frequent words of English account for around 70% of all English text. That is to say around 70% of the English we speak and hear, read and write is made up of the 700 commonest words in the language. The most frequent 1,500 words account for around 76% of text and the most frequent 2,500 for 80%. Given this, we decided that word frequency would determine the contents of our course. Level I would aim to cover the most frequent 700 words together with their common patterns and uses. Level 2 would recycle these words and go on to cover the next 800 to bring us up to the 1,500 level, and Level 3 would recycle those 1,500 and add a further 1,000. We would of course inevitably cover many other words in the texts to which students were exposed, but we would highlight first the most frequent 700, then 1,500 and finally 2,500 words in the language.

In one way this took us back to the pioneering work in the analysis of lexis of scholars like West and Thorndike in the 30s and 40s. But the computer would be able to afford a much more thorough and efficient analysis than had been possible in those days. The database assembled at Birmingham would provide us with detailed information about the commonest words and patterns in English and the meanings and use of those words and patterns. At first we had doubts about the practicality of the lexical syllabus. But the more we worked with the information supplied by the COBUILD research team the more we became convinced that the syllabus which emerged was highly practical, entirely realistic and vastly more efficient than anything we had worked with before.

I have already pointed to words like problem, solution and idea which are omitted from most language courses, even though they play a vital function in structuring the way we speak and write. A particularly striking example is the word way, the third commonest noun in English after time and people. The word way in its commonest meaning has a complex grammar. It is associated with patterns like:

. . . different ways *of cooking* fish.
A pushchair is a handy way *to take* a young child shopping.

What emerges very strongly once one looks at natural language, is the way the commonest words in the language occur with the commonest patterns. In this case the word way occurs with of and the ring form of the verb and also with the to infinitive. It is also extremely common with a defining relative clause:

I don't like *the way* he talks.

The lexical syllabus does not identify simply the commonest words of the language. Inevitably it focuses on the commonest patterns too. Most important of all it focuses on these patterns in their most natural environment. Because of this, the lexical syllabus not only subsumes a structural syllabus, it also indicates how the 'structures' which make up that syllabus should be exemplified. It does this by emphasising the importance of natural language.

As we began work on the course design, therefore, a number of basic principles were agreed:

- The methodology employed would be based entirely on activities involving real

language use.

- Learners would be exposed almost entirely to authentic native speaker language. They would not be taught through the medium of'TEFLese'- a language designed to illustrate the workings of a simplified grammatical system and bearing a beguiling but ultimately quite false similarity to real English.
- Spoken material recorded specially for use in the course would not be scripted and rehearsed. It would be spontaneous speech and would therefore contain many linguistic features normally idealised out of language teaching material.
- We would not 'present' learners with language but would encourage them to analyse for themselves the language to which they were exposed and thus to learn from their own experience of language. We would not say to learners 'I, the teacher, will exemplify for you the important features of English, and you, the learner, will thereby build up a description of the language in the way that I have determined'. We would say instead 'You, the learner, already have valuable experience of the language. We will help you to examine that experience and learn from it'.

In effect what we planned to do was create a learners' corpus and encourage learners to examine that corpus and generalise from it. I have already referred to the COBUILD corpus of 20 million words. By studying this corpus in great detail, lexicographers were able to make valid and useful generalisations about the meanings and uses of the words in the language. For Level 1 of our course we intended to create a corpus which would contextualise the 700 most frequent words of English and their meanings and uses. We would then highlight those words with their meanings and uses to provide learners with valuable exposure and experience. We would then devise exercises to encourage learners to analyse that experience of language and to learn from it. Levels 2 and 3 would go on to do the same at the 1,500 and 2,500 word frequency levels.

We set about designing tasks for use in the classroom. Some of these were based on written and some on spoken texts. All of the spoken tasks designed to be performed by learners were carried out by native speakers and recorded. This gave us a bank of texts, both spoken and written which we could use to provide learners with balanced exposure to the language. The balance was determined by the original COBUILD research. We identified from that research the important features of language we wished to illustrate and then constructed our corpus by selecting texts which would indeed illustrate those features of language. This was a long and time-consuming process. All the texts we used had to be closely analysed and many of them were rejected on the grounds that they did not afford us economical coverage. What we finished with was a small corpus of language which presented the learner with a microcosm of the 20 million COBUILD corpus. In becoming familiar with this corpus, the learner would become familiar with the language as a whole since the corpus contained all the important features of the words which make up 80% of language use.

The lexical syllabus, therefore, affords the learner a coherent learning opportunity. It does not dictate what will be learned and in what order. It offers the learner experience of a tiny but balanced corpus of natural language from which it is possible to make generalisations about the language as a whole. It then provides the learner with the stimulus to examine that mini-corpus in order to make those productive generalisations.

The process of syllabus design involves itemising language to identify what is to be learned. Communicative methodology involves exposure to natural language use to enable learners to apply their innate faculties to recreate language systems. There is an obvious contradiction between the two. An approach which itemises language seems to imply that items can be learned discretely, and that the language can be built up from an accretion of these items. Communicative methodology is holistic in that it relies on the ability of learners to abstract from the language to which they are exposed, in order to recreate a picture of the target language.

The lexical syllabus attempts to reconcile these contradictions. It does itemise language. It itemises language minutely, resting on a large body of reseach into natural language. On the basis of this research it makes realistic and economical statements about what is to be learned. But the methodology associated with the lexical syllabus does not depend on itemisation. It allows learners to experience language items in natural contexts and to learn from their experience. It relies crucially on the concept of the learners' corpus. It is the concept of the learners' corpus which reconciles the contradiction between syllabus specification and methodology. Once we had come to this realisation the concept of the learners' corpus was simple. The processes by which we came to this concept, and the procedures which realised it are far from simple. It is those processes and procedures which are described in this book.