Draft

A lexical syllabus for language learning

John McH. Sinclair and Antoinette Renouf

1 The EFL syllabus 1.1

Introduction

An EFL syllabus is a set of headings indicating items which have been selected, by a language planner or materials writer, to be covered in a particular part of the curriculum or in a course series. Its content is usually identified in terms of language elements and linguistic or behavioural skills. Sometimes there is a methodology built into it, although syllabus and methodology are in principle distinct. (See Section 1.5 for further discussion of the role of methodology.)

The syllabus may be a simple list, or it may have a more complex structure. The list may be prioritized according to some notion of importance or usefulness; or it may be graded according to some notion of difficulty; or hierarchically ordered.

The traditional view of a syllabus gives it an independence from any particular course that follows it. It is not negotiated in the knowledge of the precise needs and expectations of a given individual or group of students. At times when linguistic description is settled and unchallenged, it is reasonable to consider abstractly what areas of grammar and vocabulary should be covered in one or more years, and regard the matter as settled.

Syllabuses, then, are usually presented as independent statements; they may show a family connection with a prevalent theoretical approach, because the terms they use indicate the orientation of those who write them. But it seems that nowadays the syllabus is in fact influenced by other considerations in the teaching spectrum and is less independent. It should be noted that throughout this paper we are using the word 'syllabus' to mean an official, explicit, public statement intended to control the teaching activity, and not the variety of unofficial, hidden, incidental syllabuses which are adduced from time to time.

Most syllabuses, but not all, are expressed principally in linguistic terms, and there are many different approaches to language that can be used as a basis.

An English language syllabus generally used to be organized structurally, in that the briefest statement of it was a list of grammatical points - verb tenses, comparison of adjectives, etc.

A closer look showed that there were also some secondary organizing features, in particular the introduction of vocabulary words. From this inspection one could appreciate the connection between the syllabus and the prevalent theory of language at the time. A summary statement of that theory would be: language consists of a set of rules for the combination of words into well-formed and meaningful sentences. A small number of frequent words are used to indicate the structural frameworks and these have no independent content. The frameworks provide places for the selection of content words chosen from a large lexicon. Pride of place is given to the grammar, and the vocabulary is clearly secondary.

In recent years, the specification of syllabus has changed. There is now a large group of notional, functional and communicative syllabuses which reflects a different theory of language, deriving not so much from traditional linguistics as from theories of discourse based on speech act philosophy. Language is viewed as a list of potential acts, and language behaviour is a succession of such acts. Syllabus headings look like a selection of verbs with pronounced illocutionary force, or nouns formed from them, like 'inquiry', 'comparison'.

There is no comprehensive theory of language in these terms available as yet, so such syllabuses rest on shaky ground. The partial descriptions of discourse that exist (see Coulthard 1985 for a review of these) suppose a hierarchy of functions rather than a succession of them, and offer structural frameworks that serve to organize the individual acts. In the new syllabuses, the structural frameworks are largely ignored, and no criteria are provided for distinguishing functions and other features which appear to overlap. There is no claim to provide a list of functions which is comprehensive in a given area.

1.2 Vocabulary in syllabuses

The measurement of progress in a language often includes an assessment of the number of words a learner knows. One of the clearest examples of this is a graded reader scheme, which produces word lists at several levels. The school syllabuses in several countries are similarly organized; on occasions the target word" lists are the product of substantial research (Gougenheim *et al. 1956).*

However, in recent years there has been little interest taken in the

DRAFT

lexical approach, and little apparent control exercised in published materials (McCarthy 1984). Different books which offer themselves as covering similar ground show widely differing treatment of vocabulary.

For example, an analysis of nine major EFL courses (Renouf 1984) shows that in the first book of each series, the number of different word forms introduced ranges from 1,156 to 3,963, which is a wide variation. Also the average number of times a word form recurs ranges from six to seventeen: this means that the pattern of reinforcement ranges widely too.

It is not clear what is signified by the presence of a word in the published word list of a coursebook. Many words which occur several times in the body of the book are not acknowledged at all, whilst official teaching words sometimes receive very little reinforcement, with some occurring only twice in the entire volume.

This suggests that the approach taken to the vocabulary has generally not been systematic and that there has been little coordination in establishing targets. The vocabulary is regarded merely as the means of exemplifying other features of the language. It serves all the other syllabuses, or syllabus strands. Therefore it is not normally organized in and for itself, and receives only partial attention.

An example is the word *give,* which is always included in courses, but with a restricted and different range of uses in each. As expected, in all courses it tends to represent the archetypal di-transitive verb, particularly in the forms *give* and *gave,* and in contexts such as:

/ give Tom this *book!*
/ he gave her some sugar/
/ here's the card. You gave it to me on Monday/
/ Mary is giving Arthur a cup of teal
/ what shall we give him?/
/ you've got my phone number. Give me a ring!
/ give my love to JiII and Carole/

Perhaps less obviously, in most courses the forms *give* and *giving* also play a significant role in the meta-Language, in instructions to the learner as to how to proceed, such as:

/ give more answers like this/
/ give facts to describe: a lake, a city, a river/
/ give another word or phrase to replace the following!
/ and say when things will be ready, giving time limits/
/ write giving details of likely length of stay/

In nine major course series, one third of the total instances of *giving* are in fact devoted to this use.

In functional-notional courses, there is a common metalinguistic use of *give* and *giving* for categorizing speech functions. In this case, the di-transitivity is not fully realized. Typical examples are:

/ (ask for and) give information about *x*/
/ give (instructions and) advice/
/ (ask for and) give directions/
/ giving and receiving instructions/
/ giving directions/
/ giving advice/

In the same nine course series, every second instance of the word form *giving* is used in this way.

Other uses also occur in language courses, and altogether it is clear that each course presents a different profile of the word *give.*

1.3 Mixed syllabuses

Statements of syllabus can be mixed, so that a grammatical list and a lexical list may together constitute a syllabus. There will be no indication of how these are to be coordinated, so it must be assumed that they will be focused upon in separate sessions.

Of course, it is almost impossible to teach grammar without in passing teaching some vocabulary. Vocabulary fleshes out the structures, introduces variety and promotes practice of the structure in question. The vocabulary is not the organizing force, but many teachers feel that this kind of teaching covers an adequate vocabulary.

In the same way, a class devoted to expanding its vocabulary will not be able to avoid syntax, assuming that it is not exclusively committing lists to memory. And for the same reason, the structures will not be controlled, being those that arise in passing. It is exceptionally difficult to teach an organized syllabus of both grammar and lexis at the same time.

In the activity of text explication, there is often a good balance between grammatical and lexical focus, but this does not constitute an organized syllabus in either field. The points come up as they occur in text, and that is the only organization.

The newer notional-functional syllabuses, of which the *Threshold Level* is a notable example (van Ek 1977), are not usually a simple list of headings, but are related to grammatical and lexical features. Typically, each notion or function is presented with a range of linguistic expressions that are held to be appropriate realizations. The latter constitute a partial syllabus, but once again the vocabulary is not organized for its own sake and, in many cases, the words are principally grammatical.

DRAFT

There is another type of syllabus which has some popularity. Language proficiency is expressed in terms of levels of proficiency in reading, writing, speaking and listening - the four skills. Skill-based syllabuses may be free-standing, or may be mixed with others, or partially coordinated. For example, a reading skills syllabus may be coordinated with a graded vocabulary, or a speaking skills syllabus with a range of speech functions. *Reading and Thinking in English* (Widdowson 1979-80) is based on a series of syllabuses like these.

Since the mid-1970s, there has been a growth of interest in the task-based syllabus. This is not expressed in linguistic terms at all; syllabus items refer to activities in the world, like 'interpreting a timetable', or 'changing a wheel'. A judicious selection of tasks will provide a varied learning environment, and the language engendered will be quite natural. This type of syllabus was in operation on a large-scale teaching project in Jeddah (Harper (ed.) forthcoming).

A task-based syllabus is not normally mixed or coordinated with any other, because it is held that, properly designed, such a syllabus will cover a sufficient range of vocabulary, grammar, notions, functions and skills. In relation to language, then, a task-based syllabus is a contradiction in terms. Obviously there is no linguistic theory which corresponds to a non-linguistic syllabus. Some check on the basic assumption could be provided *post hoc* by field research, involving the recording of a series of actual classes constituting a course, and the analysis of the language which occurred. This would be extremely laborious, however, and one such exercise would be no real guide to what might happen in the next.

In most modern course books, there is evidence of an attempt to coordinate several parallel threads of syllabus (see, for example, the *Cambridge English Course,* Swan and Waiter 1984-5). Sometimes the structure is stated explicitly and elaborately. Typically the books contain a recurrent series of activities which imply a syllabus that mixes skills, structures, lexis, notions, functions and tasks. The variety is often bewildering, and the actual coordination minimal.

But the books give expression to a point of view which is probably held fairly generally - that no one method of organization is adequate for a balanced and comprehensive course. Language has many facets, and corresponds partially to many different patterns of organization. True, it can be represented substantially as a set of structures and a list of words; true also that it can be seen as performing a variety of functions. Or the learner can be monitored through the skills or through a set of tasks. Some teaching from each of these points of view is necessary to make the teaching effective and efficient.

1.4 Eclecticism

Some language teachers favour an eclectic approach to the planning of work, which can be a different position from adopting a detailed pre-set mixed syllabus. Instead, they prefer performance targets to be set, and to have a fairly free hand in how the targets are achieved. They thus take on themselves the responsibility of devising a reliable syllabus.

Most language teachers, however, do not have the choice, but are obliged to use a textbook and nothing else. This state of affairs is dictated often by economics, sometimes by politics and religion, sometimes by educational tradition or bureaucracy. In our experience, there is generally very little resentment from the teachers, because the language they have to teach is quite tidily presented through a textbook, particularly if it has been specially composed for them.

Although there are exceptions, particularly in the UK, there is for language teachers in state schools and private organizations increasingly no distinction between syllabus, methodology and course book. All are blended in an officially blessed publication from which it is imprudent or illegal to deviate.

1.5 A methodology is not a syllabus

The profession of English Language Teaching in recent years has seen the rise of methodology to a dominating position. The content of language teaching - the specification of what has to be taught _ has been relegated to a secondary role. The assumption seems to be that plenty of the right sort of activity will provide a sufficient framework for language learning to take place. The exact nature of the content, the sequence of events and the pattern of coverage will not be specified fully.

In extreme cases, such an approach to methodology denies the relevance of a content syllabus - in task-based learning, for example, or in the communicative approach. Such confidence in method renders syllabuses unnecessary.

It also implies a lowering of confidence in the reliability and usefulness of independent syllabuses. Lists of structures and vocabulary words are reminiscent of teaching methods which are not currently in fashion; notions and functions are not exhaustively specified and do not of themselves constitute a comprehensive syllabus.

This point of view is in line with a shift of interest away from language data which has characterized the profession recently. The British Council's 50th Anniversary Conference in 1984 was entitled

DRAFT

'Progress in English Studies'. So little progress was reported, however, that the resultant conference record has a different title: *English in the World* (Quirk and Widdowson 1985).

1.6 A coursebook is not a syllabus

A coursebook is essentially a set of instructions concerning operations in the classroom. Whether or not it contains one or more syllabus statements, or refers to an external syllabus, the bulk of it is an elaboration of only one of many ways in which coverage of the syllabus may be achieved.

There is a tendency at present for syllabuses to be incorporated into coursebooks - both in the large number of national textbook projects, and in the fairly free area of international publishing. The danger is that the syllabus could be confused with other aspects of the teaching-learning apparatus, and be little more than an appendix.

For a syllabus to have an important role in education, it should either pre-exist or be devised independently of other elements like course materials, methodology, and assessment. It should be as independent of linguistic or pedagogical theory as possible, and the theoretical background should be seen primarily as a vehicle for the clear expression of the syllabus. A syllabus which is negotiated in advance of being taught shares many features with an independent one.

A syllabus which is dependent on a particular course book is a degenerate syllabus, not very much different from the table of contents. It might even have been composed after the materials rather than before.

2. The lexical syllabus

2.1 A word list is not a syllabus

A simple list of words is not nearly explicit enough to constitute a syllabus. In order to construct an adequate syllabus, it is necessary 'to decide, in addition to which words we want to include in our syllabus, such things as what it is about a word that we want to teach, and what counts as a word.

2.2 What is a word? - word and word form

Syllabus designers and coursebook writers as a whole have conducted relatively little empirical research into the nature of lexis, and consequently the concept of 'word' remains blurred. The conventional view is an inclusive one: that the term 'word' denotes a unit of language comprising a base form, such as *give,* and an 'associated' set of inflexions, such as *gives, giving, gave, given.* Sometimes derivations will be included, e.g. *gift.* This concept of 'word' is also an established one in computational linguistics, where all forms, including the base form, can be subsumed under the term 'lemma'.

Where a word list accompanies a particular language course, it will typically consist of abstractions derived from the base form of a word group - *go* will stand for *go, goes, going, went, gone,* and so on. Unless there is an indication to the contrary, the implication will be that all forms of the word go are covered in the material. However, this is not necessarily true, nor is it necessarily desirable. It is not actually the case that all forms of words in a given word list are shown in use in a language course, even taking into account the various levels of progress. *Hold* and *holding* may appear, but not *holds; stand* and *stood* may be exemplified, but not *stands.* There is no evidence that such omissions are based on principle, and in any case, the principles involved would not be lexical.

From a lexical point of view, it is not always desirable to imply that there is an identity between the forms of a word. Textual evidence shows that an inconsistent relationship holds between such elements. Sometimes all forms of the word - for example, *get, gets, getting, got,* and the singular and plural forms of many nouns - share a similar range of meanings and usage patterns, and it is justifiable to indicate this. But often, particularly with the commoner words of the language, the individual word forms are so different from each other in their primary meanings and central patterns of behaviour (including the pragmatic and stylistic dimensions), that they are essentially different 'words', and really warrant separate treatment in a language course.

The morphological pair *certain* and *certainly* is one case in point. Consider the following contrast in their uses, listed in order of importance as they are shown in the 7.3 million word Birmingham Corpus which forms part of the Birmingham Collection of English Text (Renouf 1984):

certain
Function 1. (60% of occurrences) Determiner, as in:
/ a certain number of students/ in certain circles/
Function 2. (18% of occurrences) Adjective, as in:
/ I'm not awfully certain about. . ./ We've got to make certain!
Function 3. (11 % of occurrences) Adjective, in phrase 'A + *certain* + noun', as in:

DRAFT

/…has a certain classy ring! there is a certain evil in all lying!/

certainly

Function 1. (98% of occurrences) Adverb, as in:

/it will certainly be interesting! He will almost certainly launch into a little lecture . . ./

There is one area of overlap between the two, where *certain* appears in contexts like:

/there is certain to be water *here/*

which is arguably paraphrasable by 'water will certainly be here', or 'there is certainly water here'. But this use of *certain* occurs rarely.

Other pairs of this kind include *easy* versus *easily; near* versus *nearly; real* versus *really; particular* versus *particularly; vain* versus *vainly;* also *west* versus *western; use* versus *used; one* versus *ones; detach* versus *detached.*

2.3 Which words? - criteria for lexical selection

Whilst the question of lexical selection has passed many course writers by, there have been attempts made through the years by a number of individual linguists to establish criteria for creating lexical inventories for teaching purposes. These will not be gone into here, though we acknowledge the efforts of such people as Ogden (1930),
Thorndike and Lorge (1944) and West (1953), and the existence of selectional criteria identified in terms of 'disponibilite' (Gougenheim *et al.* 1956), 'familiarity' (Richards *et al.* 1956/1974), 'coreness' (Carter this volume), and in various terms by others.

All these people have been concerned with the problem of identifying the lexical items which should be introduced into an all purpose programme for teaching English for general purposes. The needs of a specific group of learners are usually easier to identity.

It seems reasonable to us, in the absence of any specific guidelines, to propose that, for any learner of English, the main focus of study should be on:

a) the commonest word forms in the language;
b) their central patterns of usage;
c) the combinations which they typically form.

In the Birmingham Corpus the list of top-ranking word forms looks as follows. It is in fact not particularly controversial, confirming largely the intuition of the language teacher about which words should be in any course, and it continues to be familiar for the first 800 or so items.

First 200 word forms in the Birmingham Corpus, ranked in order of *frequency* of *occurrence:*

1 the
2 of
3 and
4 to
5 a
6 in
7 that
8 I
9 it
10 was
11 is
12 he
13 for
14 you
15 on
16 with
17 as
18 be
19 had
20 but
21 they
22 at
23 his
24 have
25 not
26 this
27 are
28 or
29 by
30 we
31 she
32 from
33 one
34 all
35 there
36 her
37 were
38 which
39 an
40 so
41 what
42 their
43 if
44 would
45 about
46 no
47 said
48 up
49 when
50 been
51 out
52 them
53 do
54 my
55 more
56 who
57 me
58 like
59 very
60 can
61 has
62 him
63 some
64 into
65 then
66 now
67 think
68 well
69 know
70 time
71 could
72 people
73 its
74 other
75 only
76 it's
77 will
78 than
79 yes
80 just
81 because
82 two
83 over
84 don't
85 get
86 see
87 any
88 much
89 these
90 way
91 how
92 down
93 even
94 first
95 did
96 back
97 got
98 our
99 new
100 go
101 most
102 where
103 after
104 your
105 say
106 man
107 er
108 little
109 too
110 many
111 good
112 going
113 through
114 years
115 before
116 own
117 us
118 may
119 those
120 right
121 come
122 work
123 made
124 never
125 things
126 such
127 make
128 still
129 something
130 being
131 also
132 that's
133 should
134 really
135 here
136 long
137 I'm
138 old
139 world
140 thing
141 must
142 day
143 children
144 oh
145 off
146 quite
147 same
148 take
149 again
150 life
151 another
152 came
153 course
154 between
155 might
156 thought
157 want
158 says
159 went
160 put
161 last
162 great
163 always
164 away
165 look
166 mean
167 men
168 each
169 three
170 why
171 didn't
172 though
173 fact 174 Mr
175 once
176 find
177 house
178 rather
179 few
180 both
181 kind
182 while
183 year
184 every
185 under
186 place
187 home
188 does
189 sort
190 perhaps
191 against
192 far
193 left
194 around
195 nothing
196 without
197 end
198 part
199 looked
200 used

DRAFT

Certain of these word forms will not warrant separate treatment, but can be subsumed under their base form or full form in a teaching list. Others, according to the criterion of textual frequency, will also be eliminated from the final list. Among the top 650 items, *clothes* will appear, but not *clothe; building,* but not *build; roughly,* but not *rough; simply,* but not *simple; suddenly,* but not *sudden.*

A few word forms appear here which are perhaps less expected - *back, own, life* and *great,* for example. But a look at their use in text explains their prominent status in the corpus: they are in frequent daily use, but probably largely at a subliminal level. Let us try to account for the prominence of *back* and *own:*

back

Function I. Adverb, as in:

/ are you getting the bus back?/ I've just come back from *Cl* afterwards we go back to sleep/

Function 2. Noun/Headword, as in:

/ Brody put his left hand behind his back/ will you three sitting at the back please move round/

Function 3. Adverb, as in:

/ into the car without looking back/ she moved back a little/ I turned back to Mary/

own

Function I. Adjective or Phrasal Element, as in:

/ I can say that my own childhood was unhappy/ a . . . unit of my own/ both phrases will do though I prefer my own/

Function 2. Noun/Headword, as in:

/ Oh, I'm not on my own, then/ he didn't like sitting on his own and reading about it/ he had developed it on his own/

These are surely everyday uses which need to be reflected to some extent in a course.

To base a selection of words on a study of native-speaker usage is not, however, to imply that there is an identity between the worlds of the learner and the native speaker. There are already signs that specialized corpora will be established, to serve the needs of the major English language learning communities.

The statistics of word occurrence are vindicated when usage is examined. But it would be difficult to construct a motivating course based entirely on the 200 words listed above, and ludicrous, say, to try to start with the top fifty. Hardly any text of any length, spoken or written, will be found with such an impoverished vocabulary. The list must be extended to include some lower frequency items.

In this way, the materials writer will have s!}me flexibility, and there will be a reasonable range of topics which can be covered, and a chance that the work will be lively and interesting.

The additional list will probably include, among other things, words relating to domestic reality, such as days of the week and kinship terms, and other common lexical sets; also further words to refer to physical sensations and personal emotions, and to use in making evaluations. These additions should be monitored carefully, so that the final word list contains items of maximal utility and power. A balance has to be achieved, however, between natural usage and utility.

As said, the introduction of whole lexical sets is not justified by the criterion of frequent use, and evidence shows why it is that some set members feature more centrally in the language. This is sometimes due to facts in the real world - for example, that Sunday is the most cited day of the week; it sometimes reflects the extent of metaphorical usage, as in the case of certain colour terms. Set membership is also only one of the roles played by many common words. In the case of *black* and *white,* for instance, their sociological reference is actually far more common today than their physical one. All this needs to be taken into account in the selection of such words, and the spurious tidiness of language that one achieves by listing words and phrases that are like each other should be viewed critically.

2.4 What to teach? - central patterns of usage

Moving on from the selection of a word list to the idea of basing a syllabus on normal mainstream usage, the everyday core of the language, the question is how to establish what this is. There has not until very recently been any way of reliably doing so.

Now there are a small but growing number of large, computer-held banks of text, such as the Birmingham Collection of English Text, that can provide evidence of typical language use. The retrieval systems, unlike human beings, miss nothing if properly instructed no usage can be overlooked because it is too ordinary or too familiar. The statistical evidence is helpful, too, because it distinguishes the commoner patterns of usage, which occur very frequently indeed, from the less common usage, which occurs very infrequently.

The human being, contrary to popular belief, is not well organized for isolating consciously what is central and typical in the language; anything unusual is sharply perceived, but the humdrum everyday events are appreciated subliminally. Let us take as an example the word *see.* Textual evidence shows us that the first and second most frequent uses of *see* are those found in the familiar phrases *you see* and *I see.* The first is an indication of interactive concern in spoken discourse:

DRAFT

/ - 1 thought they were all away, you see - /
/ well, you see, 1 have to like her, you see, because she's invited . . ./
/ Yes 1 know but, you see, computers might rule the world one day/
/ You see, my wife's nervous when she's left alone/

and the second is a response with a wide range of meaning, to some extent controlled by its intonation, and falling within the semantic area of 'understand':

/ '. . . and one as you leave the country, you see.' 'I see, and how long is that?'/
/ '. . . It's Saturday to Sunday,' 'Oh, 1 see.' 'cause you don't count your first night . . .'/
/ . . . Oh, I see. Well, that's as good a reason as any for. . ./

Most native speakers, despite accepting the truth of the above, would probably say, as we did, that it is the *seeing* through one's eyes which is surely the major use/meaning associated with *see.* Lyons (1977 p.247) explains this kind of disposition by reference to the concept of 'salience', psychological or biological.

It is most important that the evidence of very long texts is not dismissed without careful thought. Language text is the record of linguistic choices, but it is not necessarily a completely comprehensive and reliable record of the way the choices are made. No claim is made here that frequency of occurrence is the only relevant factor. Equally, no description of usage should be innocent of frequency information.

In view of its prominence in natural text, the interactive function in discourse of word forms like *see* above tends to be under-represented in coursebooks. In nine courses, instances of *you see* and *I see* together account for only 10 per cent of all occurrences of *see,* as opposed to 53 per cent in the Birmingham Corpus.

The conventional view of the words in a language is that they either have lexical meaning or are confined to syntactic functions in the sentence. Hence usages which are discoursal or pragmatic, which carry out functions to do with the larger patterns in texts, are often missed. For example, the humble and exceedingly frequent word *to* has a discourse function which is important and valuable to a user. It often occurs at the beginning of a move or sentence, and indicates that the comment which it introduces is an evaluation of the main part of the utterance. Some of these uses have become familiar phrases, like 'to be honest', 'to cut a long story short', and the contrast between this function and the use of *to* as if it was 'in order to', can be seen in the ambiguity of the following (constructed) example:

To be fair, Tom divided the sweets evenly

In one of the meanings, the person who is fair is the person speaking or writing, and the phrase is a judgement on Tom.

Delexical verbs

A major feature of the language is not specifically taught in current coursebooks. It is the phenomenon known as 'delexicality', the tendency of certain commoner transitive verbs to carry particular nouns or adjectives which can in most cases themselves be transitive verbs. In general, the more frequent a word is, the less independent meaning it has, because it is likely to be acting in conjunction with other words, making useful structures or contributing to familiar idiomatic phrases.

A clear example of this is *give,* which is most commonly used di-transitively in conjunction with certain nouns, particularly *look, information,* and *advice.* In Section 1.2, we showed that this delexical use of *give* does occur in language courses, but primarily in the rubric of the text, which is apparently not part of the teaching programme.

Rave has a range of delexical uses, and combines with various classes of abstract nouns. *Look* is the single most common collocate. The nouns are commonly modified, as for example in *a good look, minor doubts, a deep longing, a heart to heart talk, a strange feeling, legitimate expeaations.*

In contrast, the delexical contexts for *have* in most coursebooks are somewhat more concrete. The frequent occurrence of *look* here reflects natural usage, but otherwise there seems to be an obsession with having a drink, a bath, or a shower, in that order. An exception to this is the reference in some books to 'words which *have* the same or different *meaning,* and so on. But again this occurs only in the rubric of the text.

Textual evidence now shows us the extent to which the phenomenon of delexicality occurs. The primary function of *make,* for example, is to carry nouns like *decision/s, discoveries, arrangements,* thereby offering the alternative phraseology 'make your own decisions' to 'decide on something'; 'make her travel arrangements' to 'arrange her travel,' and so on. Which of the two formulations to choose is obviously a strategic matter in text creation, but the delexical option is firmly there. Since it complicates the syntax, it must provide an overriding facility to justify the complexity. Other verbs which operate in this way are *take,* with e.g. *care of, note, aaion;* and to a lesser extent *put,* with e.g. *question.* Aisenstadt (1981) has noted the importance of delexical verb combinations for vocabulary teaching and learning.

DRAFT

2.5 What to teach? - typical word combinations

In the two previous sections, we have made reference to the fact that words combine, or collocate, with each other in certain characteristic ways. With the benefit of a corpus of real text, we can now be clearer about what these preferences are, and be more systematic in presenting them to the learner of English.

The combinations are often lexical collocations, where two words occur next to each other, e.g. *happy marriage* or *accidental death.* Sometimes the combinations have grammatical restrictions, and the words are regularly found in a particular syntactic pattern, such as 'accede to X's demands', where a different pattern, such as 'X's demands were not acceded to' is unlikely to Occur.

Even common grammatical words have collocational patterns. *Each,* for instance, occurs significantly with units of time - *hour, day, week,* etc. Similarly, *of* collocates frequently, in its left-hand context, with *kind, part,* and *sort.*

Common grammatical words also combine with each other, often in discontinuous frameworks, such as *a... of* In turn, they attract particular lexical words, so that the *a... of* framework typically encloses the following: *lot, kind, number, sort, couple, matter, bit, series, piece, member,* in that order of frequency.

In these ways, the essential patterns of distribution and combination in modem English will be included in the lexical syllabus. It is not possible here to present a full account of the field, but it is clearly one of the growth points of research which will feed into language syllabuses for many years to come.

3 Implications of a lexical syllabus

3.1 Vocabulary

The approach to a lexical syllabus which is taken above highlights the common uses of the common words. The common words are very common indeed, and mastery of them is rewarding in practice. Typically they are found each to have a few very common uses and a number of minor ones that can be given a low priority in the selection of items to be taught.

The is approximately 4 per cent of all text; *and* and *of* make up another 4 per cent. The top ten words in the list printed earlier in this paper account for about 17 per cent. The little words that make up the structural framework of the language and that are the recurrent elements of phrases are found to dominate the frequency lists. Only *time, people, new, know, man* and *little* bring any great semantic content into the top hundred. After that, familiar lexical items come in much faster.

Almost paradoxically, the lexical syllabus does not encourage the piecemeal acquisition of a large vocabulary, especially initially. Instead, it concentrates on making full use of the words that the learner already has, at any particular stage. It teaches that there is far more general utility in the recombination of known elements than in the addition of less easily usable items. The more delicate discrimination of meanings which is accessible to someone who commands a large vocabulary is postponed - at least from the business of the first few years of English.

It is important to recognize that this is not so much a point of view about how to teach a language as a statement about the nature of modem English. Other languages may be different; English makes excessive use, e.g. through phrasal verbs, of its most frequent words, and so they are well worth learning.

3.2 Relation to other syllabuses - implementation

In the construction of a balanced and comprehensive course, the designer will no doubt keep a tally of structures, notions and functions, as well as vocabulary. But in the presentation of materials based on a lexical syllabus, it is not strictly necessary to draw attention to these check lists. If the analysis of the words and phrases has been done correctly, then all the relevant grammar, etc. should appear in a proper proportion. Verb tenses, for example, which are often the main organizing feature of a course, are combinations of some of the commonest words in the language.

This is different from attention to combinations of the four skills, and the use of tasks to practise effective communication, about which the lexical syllabus is neutral. It is an independent syllabus, unrelated by any principles to any methodology. It may suggest that certain types of teaching practice can readily adapt to it, but that is not a principled connection, just a similarity in philosophical approach.

Whenever it occurs that the learning process would be improved by introduction of, say, a grammatical table, that does not disturb or interfere with a lexical syllabus - it merely sheds light from a different angle.

3.3 Efficiency

One big advantage of a lexical syllabus is that it only offers to the learner things worth learning. Variations are not built in, as they are

DRAFT

in conventional presentations of language structure, but are introduced when they are necessary. So instead of building up phrases, the learner will be gradually breaking them down, sensing the variability. Needless variation can lead to ungrammatical expression, as we know from the common phenomenon of over-correction.

On the other hand, there is no suggestion of the stifling of creativity; not only is it possible to teach to a lexical syllabus just as creatively as to any other, it is likely that sensitivity to the rule margins, where so much linguistic creativity lies, will be greater than in conventional presentations, which are about as sensitive as a fruit machine, and espouse a similar theory of message structure. It sometimes appears to teachers that a syllabus based on observation of language is somehow backward-looking, giving licence only to what has occurred. The lexical syllabus is just a much more detailed inventory of the possibilities of the language.

3.4 Utility

The argument for utility in a lexical syllabus is not confined to the exploitation of common words. What are exposed are the *uses* of those words, and prominent among those are devices, signals, and strategies in discourse, both spoken and written. The emphasis shifts from constructing messages to delivering them, and delivering them to maximum effect, and to achieving communicative goals.

The description of discourse has brought out clearly that language text is simultaneously organized on at least two different dimensions, or 'planes'. One of these is used for constructing and elaborating messages, and another is concerned with commenting on, labelling, evaluating and generally negotiating the messages interpersonally. Because language teaching has not until recently recognized the importance of the purposes, intentions, objectives, etc. of language users, the second of these planes has been largely neglected. If the teacher stops the talk and the class waits for a student contribution, the student is unlikely to learn how to get an opportunity to speak, or to recognize the structures and vocabulary that will express his intentions.

3.5 Statement

One form which a lexical syllabus of the kind discussed could take is shown in a sample below (an abridged version of the original). The syllabus in this case is several hundred pages long and forms the linguistic specification for a new English course (D. and J. Willis forthcoming).

Lexical syllabus extract for the word 'by'

No. of occurrences in the Corpus: 21,916

CAT 1: **used to indicate the person or thing that performs or causes the action mentioned; usually preceded by a verb in passive voice** (prep) [53% of occs]
Henry was surprised by *the* plopping sounds in the water/the daily business is still announced by *the* procession of the speaker/Another survey carried out by *the* University of Florida! He was brought up by *an* aunt! He had been poisoned by *a* mushroom/ he found himself touched by *a* bittersweetness/worker-elected directors have been accused by *their* former colleagues/ Carlson was interviewed by *a* major television station/ they are protected by armour/ the first atom bomb was manufactured by famous men with bogus names/ it was bought with his own money by *his* own *cook!* The affection with which it is regarded by *its* old pupils is evidence of its success/ an investment of 12 million pounds by .Courtaulds/attacks on EEC ministers by *a* commission member/ I see this change in position by Reagan/

CAT 1.1: **used to show who is the author or artist of a particular work** (prep) [4% of occs]
Three books by *a* great and original Australian writer/ An article by J. B. Priestley in the New Statesman/ he bought great numbers of paintings by Hook, Millais, Orchardson/

CAT 2: **used (with the present participle of a verb) to show that you perform a particular action and to indicate that something happens as a result of this action** (prep) [11 % of occs]
Holmes became WBC champion by beating Norton/ They were making a meagre living by selling artefacts to the tourists/ You win by being older/ dared to take the law into his own hands by evicting a tenant! as a bird changes direction by dipping one wing and lifting the other/ is an artificial way of making the child learn by doin~ you'd be amazed the places you can get into by just looking confident!

CAT 2.1: **used to indicate the means used in order to achieve something, or to introduce the circumstances which lead to something happening** [9.5% of total occs in sample cones]
The contest was settled by *a* practical test/ The rabbit escapes once by *a* last-second change of course/ There will be a lift to go *up* and

DRAFT

down by/ approached by *a* most imposing flight of steps/she was plucked to safety by *a* helicopter/ which arrived by chancel washing our dinner dishes by hand/

CAT 3: **beside something and close to it** (prep) [3% of occs] She lingered by *the* door/ I sat by *her* bed/ We sipped tea by *a* hot stove in his living room/ the pilot with his arm full and the red can by *his* fee1/' Ralph was kneeling by *the* remains of the fire/ We moved down by *the* river/

Phrases and miscellaneous:
i) *by myself, himself,* etc [1.5% of occs], meaning 'alone'.
ii) Used in reference to times and dates [1.5% of occs].
iii) Used of standards, rules, etc. [1.5% of occs], e.g./; by British standards/

Notes:
i) In CAT 1, 48 per cent of occurrences occur in the passive; 5 per cent occur after nouns with transitive verbal counterparts (see last three examples).
ii) Some passive forms also Qccur in CAT 2.1.

Immediate collocates

RIGHT HAND
the - 5130 occs
a - 1321 occs
his - 327 occs
an - 293 occs
their - 233 occs

LEFT HAND
and - 380 occs
up - 222 occs

Acknowledgements

The authors would like to acknowledge their debt to colleagues in English Language Research at Birmingham University, in particular to members of the COBUILD project teams. Much of the evidence reported here has arisen from the project, which is largely financed by grants from William Collins Sons and Co. Ltd., publishers. Without such support, this paper could not have been written.

Points for further development

1. Sinclair and Renouf take as their point of departure the notion of a syllabus as an explicit statement, not dependent on methodology, controlling what is to be taught in a language course. Many current approaches to syllabuses lack rigour and proper control over lexical content, even though they may appear to be highly organized in other respects. Consider the syllabus you are working to in the light of their remarks.

2. Most published courses, according to Sinclair and Renouf, are inconsistent with regard to how lexis is presented. They work with word lists of base-forms of words, but then use inflected forms of those bases in a way that often does not correspond to natural language use. What is needed is a careful consideration not only of what words occur in natural use, but in what forms and in what patterns they typically occur. Sinclair and Renouf clearly feel that course books let us down in this respect; is there any evidence that the coursebooks you use take cognizance of natural language use?

3. The lexical syllabus should be devised independently of considerations of course materials, methodology and assessment. But it is not enough simply to construct a word list. First we have to be clear what we mean by *words* and *word forms;* it may not be desirable to include all the possible forms of a word in the syllabus. Textual evidence shows that word forms (e.g. *certain/certainly)* often behave quite independently of each other in terms of meaning and usage patterns; this fact seems rarely to be recognized and yet affects a vast number of word forms. The problem would seem to be how to break the traditional expectations that all forms of a word are equally important and that all will behave in the same way; consider restrictions such as the following:

He owns a factory in London.
He is the owner of a factory in London. He runs a factory in London.
'"He is the runner of a factory in London.

Should vocabulary teaching explicitly deal with such areas?

4. The basic principles for word selection are commonsense ones of frequency and centrality of patterns of usage, along with observations of the typical combinations that words form. An interesting activity is to make a list of about twenty words by sticking a pin at random in a dictionary. When, if ever, might you want to teach these words? What criteria are you using for accepting or rejecting them?

DRAFT

5. Frequency is useful, but language teaching cannot, of course, proceed with only the 200 or so most common words, most of which will be semantically 'empty' function words. Words relating to everyday domestic reality should be added, words for classroom procedures, etc. What would be the 20 or 30 most useful words in addition to the top 200 for your teaching situation?

6. In deciding what are the central patterns of usage of items, computers are often more reliable than native-speaker intuition, and often produce evidence that goes counter to expectations. Such evidence should not be ignored in the preparation of syllabuses. Should teachers be demanding more access to computer evidence? Can dictionaries and coursebooks afford to ignore such evidence?

7. The kind of evidence computer corpora can give us is illustrated by the case of delexical verbs *(have, make, take,* etc.), which do heavy duty in the formation of idioms and other multi-word combinations in English. How are these verbs dealt with in the materials you use?

8. The lexical syllabus does not foster massive vocabulary acquisition in the initial stages, but encourages learners to make full and extended use of words they already have by recombining elements. Compare this with the stance of the post-war structuralists such as Fries discussed in Chapter 3.

9. In the lexical syllabus, such things as lists of structures and notions and functions would be secondary, and would come out of the implementation of the lexical syllabus rather than constrain it. A course book based on this syllabus would be radically different from conventional ones and would almost certainly meet resistance at first; the residue of the structuralist position is still very potent (see McCarthy 1984).

DRAFT