CHAPTER 3: The lexical research and the COBUILD project

The corpus
I have suggested that the word may be a better unit of syllabus design than the structure. This is partly because word is very often prior to structure in that it is word meaning which determines which structures are grammatical and which are not. A description of language which takes the word as its starting point offers more powerful generalisations and is more accessible to learners than a structural description. A lexical description of language, therefore, should offer a powerful basis for syllabus specification.

I would like now to look at the research programme which went into the production of the Collins COBUILD English Language Dictionary. This programme was to produce a new lexical description of language which would eventually provide us with the basis for a new approach to syllabus design.

The basic aim of the COBUILD project was to develop:

- a new, thorough-going description of the English language, and one which was not based on the introspection of its authors, but which recorded their observations of linguistic behaviour as revealed in naturally occurring text. (Renouf 1987)

The first stage in the project was to gather together a corpus of language on a computer ready for analysis. Since a corpus represents a sample of the language under study it is obviously important to obtain as representative a sample as possible. Our aim was to identify those aspects of the English language which were relevant to the needs of the international user. We therefore defined these for ourselves as follows:

- written and spoken modes
- broadly general rather than technical language
- current usage, from 1960, and preferably very recent
- ‘naturally occurring’ text, not drama
- prose, including fiction and excluding poetry
- adult language, 16 years or over
- ‘standard English’, no regional dialects
- predominantly British English, with some American and other varieties.
  (Renouf 1987)

Renouf goes on to explain how texts were selected to give the right coverage and gives a broad description of the corpus. A complete list of texts can be found in the corpus acknowledgements in the Collins COBUILD English Language Dictionary.

The next important feature of the corpus is simply its size. Obviously, the larger the corpus the more likely it is to be representative of the language as a
whole, or of that part of the language researchers are aiming to describe. This need for size has to be balanced against the aims of the study and also against rapidly diminishing returns to scale beyond a certain point. By mid-1983 a Main Corpus of 7.3 million words had been built up by the COBUILD team, and this was large enough for a study of the commonly occurring words of English. The most frequent 700 words of English all occur at least 650 times in the Main Corpus. All of the 2,500 most frequent words of English, which were eventually to form the basis of the lexical syllabus for the Collins COBUILD English Course, occurred at least 120 times in the corpus. For the study of less frequent words, those occurring less than fifty times in the Main Corpus, a Reserve Corpus of a further 13 million words was added by the end of 1985. In producing the dictionary this Reserve Corpus played a vital role as about ninety per cent of the word forms in the Main Corpus occurred with a frequency of 50 or less.

The first stage in processing the corpus was to run a computer programme to produce concordances for each of the words under study. Let us look at the word way. This comes after time and people as the third most common noun in English, with around 7,000 occurrences in the Main Corpus. Here is a sample from the concordances:

- fluctuated. It is not, I sus\pound;ing on; fewer still had premises in any assertively un-urban that we affected a way of dressing quite unsuited to Unive\pound;ed to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep\pound; to the flexible response, with its way of dressing quite unsuited to Unive\pound;uted to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep\pound; to the flexible response, with its way of dressing quite unsuited to Unive\pound;uted to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep\pound; to the flexible response, with its way of dressing quite unsuited to Unive\pound;uted to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep\pound; to the flexible response, with its way of dressing quite unsuited to Unive\pound;uted to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep\pound; to the flexible response, with its way of dressing quite unsuited to Unive\pound;uted to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep\pound; to the flexible response, with its way of dressing quite unsuited to Unive\pound;uted to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep\pound; to the flexible response, with its way of dressing quite unsuited to Unive\pound;uted to the idea very gradually. The best way a hanged, and a manned craft was the best way that is born of deprivation. The spa\pound; to review in my mind various animal way of quieting him when he was obstrep;
among the horizontal ledges, the way is easy—a few inches of horizontal wire cutters and work the hook all the way across. The only way in which a telegram can be delivered is if we are thrown. Such pebbles that came his way is? as-y-a few inches of horizontal ground, the path to Camusfearna as if I had the impression that he was in no way taxing his powers, and could greatl...
The Lexical Syllabus

prompt the researcher and to hold information in a form suitable for computer input to the dictionary database. The outcome of this process, then, was a database which recorded all the relevant information about way to be incorporated in the final dictionary entry.

From concordances to database and dictionary entry

The database for way lists the main semantic fields covered by the word. It runs to over 40 typewritten pages, but can be summarised as follows:

1 method, means: It's a useful way of raising revenue.
   The cheapest way is to hire a van.

   behaviour: Play soccer Jack Charlton's way.

2 manner, style: He smiles in a superior way.

3 what happens: That's the way it goes.

   what is the case: We were so pleased with the way things were going.

4 degree, extent, respect: She's very kind and sweet in lots of ways.

5 location, movement, direction, space: In no way am I a politically effective person.

   A man asked me the way to St Paul's.

   Get out of the way.

6 distance, extent: I flew the rest of the way to Danang.

   It was downhill all the way after that.

7 time: National revolt was still a long way off.

8 miscellaneous: You're way below the standard required. (= a considerable distance)

   The AEU, in a classic balls-up, voted both ways. (= one of a number of choices)

In addition to these semantic areas, a number of discourse uses are listed, such as:

by the way: By the way, that visit of Muller's is strictly secret. (used to add something to what you are saying)

by way of: Well, that's really by way of introduction. (used to explain the function of something you are about to say, for example whether it is intended as an introduction an example, an apology etc.)

The computer input slips used to build up the description of each of these uses of way, drew attention to a number of relevant features associated with each example. An expanded entry for:

The cheapest way is to hire a van.

for example, would read:

way 1.01

DEF (definition): Used to refer to something that must be done, or a series of things that must be done in order to achieve something. [Used to answer the question 'how?' Also closely connected to the prep 'by'.]
This description relates the use of **way** to a particular meaning, gives us a syntactic environment showing that this use of **way** is associated with the infinitive with to, and adds the comment that this particular use is anaphoric - that the example refers back to 'ways of moving house'.

If we look at the syntactic environments associated with this use of **way**, (meaning means or method) we find:

- N + INF-TO
  - as in the example given,
  - N + PREP/OF + -ING
    - as in:
      - The different ways of cooking fish.
      - N + WHICH:XP
        - as in:
          - They kill animals in a way which would disturb the ordinary town boy.
          - PREP/IN + DET/THIS + N
            - as in:
              - In this way the energy in the pile is controlled.

If we look at the pragmatics of this meaning of **way** we find that it is used anaphorically as in the example given above; cataphorically as in:

- You can qualify for a pension in two ways.

for instructions, as in:

- Do it this way.

In this way we begin to build up a picture of the word **way** - not only of its meaning, but also of the syntactic patterns with which it is associated and its use in discourse.

Also annotated in the database are common phrases with way which are found so frequently that they function almost like lexical items in their own right:

- By the way this visit of Muller's is strictly secret. . . . by way of introduction.
- Little in the way of strategic thinking was needed.
- He's not on board. No two ways about it.
- In a way these officers were prisoners themselves;

Eventually all these insights are incorporated in a dictionary entry. The entry for an important word like **way** in the *Collins COBUILD English Language Dictionary* runs to two full pages and, if we include fixed phrases, runs to 36 categories, some of them subdivided. Typical sections from the entry are:
The Lexical Syllabus

way /wel/, ways 1 If you refer to a way of doing something or a way to do if you are referring to how you do it, for example the series of things that you do in order to achieve it, or the course of action that you take. EG different ways of cooking fish... A pushchair is a handy way to take a young child shopping... You can qualify for a pension in two ways ways in which the present service could be improved... In what way can I help you?.. She had decided on this course as the only way out of a hopeless situation.

25 You say by the way 25.1 when you add something to what you are saying, especially a question or piece of information that you have just thought of. By the way this visit of Moller's is strictly secret... My father's dead by the way.

25.2 to indicate that a comment or remark is not directly relevant to the main topic of the discussion. E.G That point is quite by the way.

26 You use by way of when you are explaining the purpose of something that you have said or are about to say, for example whether it is intended as an introduction, example, apology, etc. E.G I'm going to sketch in a bit of the background by way of introduction.

Entry for the word form 'WAY'

Total no. of occs. in corpus: 6,791

Category 1: REFERS TO THE MANNER OF A PERSON'S BEHAVIOUR OR ACTION, ESPECIALLY A CERTAIN ATTITUDE, STYLE OR FEELING THAT SOMEONE HAS. [Approx 44% occs.]

we take a look at the way IN WHICH computers are revolutionizing our/the disrespectful way IN WHICH these flighty females carry out their / studies of the way IN WHICH today's continents fit together / in much the same way THAT we dispose of Kleenex or beer cans / people related to each other in a way THAT I had never seen before / Vorster said in his heavy way: "Now are there any questions about the bill?" / the Spanish chroniclers did it their OWN way / about to change in a radical way / it'll be convenient because of the way we're going to work at this / the SAME way / we're not going to deal with it in the ordinary way / thinking of abstractions in quite a different way from the way we think of them / highly contemptuous of the American way OF LIFE / those defects being her virtues, her faith, her way OF LIFE / life isn't the way it ought to be / that's the way I feel about it anyway / I am old fashioned in this way I

Category 1.1: THE MEANS OR METHOD BY WHICH SOMETHING IS DONE, OR HOW IT OCCURS. [approx 10% occs.]

the most effective way OF countering the Soviet air threat / there's one other way OF getting hydro-electric power in areas / this process was a round about way OF achieving something that could have been done / an artificial way OF making the child learn by doing / we have no way OF knowing whether the kinds of men represented / family duties and responsibilities were a way IN WHICH sharing was institutionalized/I believe this is the only way THAT an ordinary person can inspire others/he had wished that there had been some way he could exchange words with this man /

Category 2: USED WITH REFERENCE TO DIRECTION, EXTENT, DISTANCE AND TRAVEL GENERALLY; ALSO REFERS TO A ROAD, PATH, ROUTE OR JOURNEY SOMEWHERE. [Approx 26% occs.]

From data sheet to language course

The information in the database was edited down, then, to provide a dictionary entry of the kind exemplified above. It was also edited down to produce a 'data sheet', one of 700 which provided the raw material for the lexical syllabus on which CCEC Level 1 was to be based, and which would be recycled through Levels 2 and 3:
I drove the wrong way ROUND a roundabout/British girl secretaries who work their way ROUND the States/it's got pretty embroidery all the way ROUND the bottom/they walked down the stairs and on the way OUT I heard him say his first words / having addressed the troop ship on the way OUT to South Africa / it does not matter WHICH way the vectors go / what sort of ship it is and WHICH way the thing is going / more of transit flying, and of fighting their way THROUGH defences / hold a tray aloft as he weaves his way THROUGH the crowd in his new role of barman / and then barge your way THROUGH and shout at the back of the queue / quite a LONG way past them was the greenhouse with the vine / it certainly was a LONG way FROM Cape Town / she keeps gaining on me all the way DOWN the long hill to the bottom of the lane / after giving him minute directions of the way. They could see the whole ridge of Wirral Hill / he led the way over the rocks / a time when we got lost - right out Dennington way. But we found our path eventually / with no fear this time of losing the way/ended up as far out of their way as Pleiku, fifty miles to the south / Phrases and misc:

I IN/OUT OF THE WAY (5 occs. in sample):
he had kept out of DeMille's way / we constantly get in the way of and interfere with /

ii BY THE WAY (4 occs. in sample):
By the way, Castle, you might get me the name of his dentist / By the way, do you keep cats here?/

Notes:

i WAY is a sub-technical word with all the expected features e.g. its dummy role in some of the examples in Category 1. e.g. bird must have flown in a direct and purposeful way /

ii It was too difficult to subdivide Category 1 further in the time available, although there seems to be some basis for doing that.

iii WAYS: this word functions as the plural of WAY, and generally follows the same usage except that almost all instances of use fall into Category 1 and Category 1.1 e.g. some babies become so set in their ways during this period/this will add another to the many ways in which the rich can buy youth / I examine various ways in which the ills of this society can be tackled / the old ways are the best ways /

iv WAYS is also characterised by the frequency of occurrence for particular left hand collocates. The most common of these are MANY, OTHER, SEPARATE, SOME and other expressions of quantity.

e.g. we improved the paper in a number of OTHER ways / there are MANY such ways in which we--chave as if we were two people / in MANY ways it was a bad bargain / we look at various ways in which over the years Britain has / China is in MANY ways a developing country too /

Further information on right-hand collocates.

<table>
<thead>
<tr>
<th>WAY</th>
<th>WAYS</th>
</tr>
</thead>
<tbody>
<tr>
<td>OF</td>
<td>637 occs.</td>
</tr>
<tr>
<td>TO</td>
<td>637</td>
</tr>
<tr>
<td>IN</td>
<td>272</td>
</tr>
<tr>
<td>THAT</td>
<td>255</td>
</tr>
<tr>
<td>AND</td>
<td>188</td>
</tr>
<tr>
<td>THE</td>
<td>180</td>
</tr>
<tr>
<td>I</td>
<td>175</td>
</tr>
<tr>
<td>OUT</td>
<td>125</td>
</tr>
<tr>
<td>THEY</td>
<td>120</td>
</tr>
<tr>
<td>HE</td>
<td>111</td>
</tr>
<tr>
<td>AS</td>
<td>108</td>
</tr>
<tr>
<td>IT</td>
<td>107</td>
</tr>
<tr>
<td>THROUGH</td>
<td>96</td>
</tr>
</tbody>
</table>
Each of the two main categories of meaning for the word *way* is the focus of an exercise in CCEC Level 1:

### 78 Ways of saying numbers

#### 78a

**a** How do you say telephone numbers in your language?

**b** Look at the numbers on the right. What are they? What about 1989 for example? Could it be a telephone number, or a date, or car V number? How would you say it if it was a date? One thousand nine hundred and eighty-nine? One nine eight nine...? Discuss with your partner how you could say the numbers. How many different ways can you find and what do they each mean?

- Tell the class

### 78c

Bridget and David talked about the same numbers. Did they think of the same things as you? Write down the things David and Bridget thought of.

<table>
<thead>
<tr>
<th></th>
<th>22</th>
<th>0</th>
<th>1989</th>
<th>3.14</th>
<th>748</th>
<th>22756</th>
<th>10.12</th>
<th>021 337 0452</th>
</tr>
</thead>
</table>

In addition to the uses of *way* in the rubric for this activity:

- Ways of saying numbers.
- How many different ways can you find...

A recording of native speakers doing the task contrasts the American way of saying dates with the British way. Inevitably the word way will feature a good deal in the exchanges in the classroom between teacher and learners, and among learners.

- How many ways did you think of?
- Yes that's another way.
- We got three ways.
- etc.

This use is highlighted again in a summary of the useful words and phrases from this unit, Unit 6:
The Lexical Research and the COBUILD Project

a) way

There are different ways of writing 'colour'- the American way (color) or the English way (colour).

How many ways are there of saying this number?

Practice these ways of agreeing and disagreeing.

I like the way he sings.

Do it this way. Look.

Unit 9 takes as its theme finding the way:

122 Landmarks

When people ask us the way and we give them directions we usually use landmarks. We say things like this.

- It's just past the hospital.
- It's opposite St Joseph's school.
- It's near the Post Office.
- It's behind the supermarket.

Look at these landmarks. Do you know what they are?

Here again it is likely that in addition to the forms occurring in the coursebook and its accompanying recordings, the word way will feature in classroom discussion.

Two other examples which occur later in Level 1 are picked up in a review section:

I may be able to stop off on my way to the USA.

He went all the way back.

These are sentences which have been contextualised earlier in Level 1 and are later highlighted.

Unit 7 looks in detail at the uses of the word to and category 7 draws attention to the pattern: N + INF - TO

| Grammar words | 8 from _ to_
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>to</td>
<td>It was reduced from £2s to £5.</td>
</tr>
<tr>
<td>1 where</td>
<td>Our lesson lasts from _ to _</td>
</tr>
<tr>
<td>2 who (with give, offer, present etc.)</td>
<td>I gave it to David.</td>
</tr>
<tr>
<td>3 listen or speak to someone/something</td>
<td>I went to see my sister.</td>
</tr>
<tr>
<td>4 listen to Bridget</td>
<td>I've come to Liverpool to stay with my parents.</td>
</tr>
<tr>
<td>5 ask, want, plan etc.</td>
<td>We asked people to write about.</td>
</tr>
<tr>
<td>6 it (see it 2, section 8 8)</td>
<td>It's for people who want to take better photographs.</td>
</tr>
<tr>
<td>7 after better photographs</td>
<td>It was nice to see you</td>
</tr>
<tr>
<td>8 after possible to phone your partner?</td>
<td>When is it possible to phone your partner?</td>
</tr>
<tr>
<td>9 after place, way, thing etc.</td>
<td>What's the best way to travel?</td>
</tr>
<tr>
<td>10 past</td>
<td>London is a good place to live.</td>
</tr>
</tbody>
</table>

101 Which categories do these sentences belong to?

a) We only have to do seven.
b) Say these words to your partner.
c) Work in groups to do these puzzles.
d) A man dressed as Napoleon went to see a psychiatrist.
e) We're going to see a film after class.
f) The cheapest thing to do is take a bus.
g) Bridget works from Monday to Friday.
h) I'd like to come back here.
i) The psychiatrist asked her to sit down.
j) Read these phrases to your partner.
k) It's difficult to see the tree.
l) He wants to go to Britain to learn more English.

Compare the examples in each category with examples in the Grammar Book.
The Lexical Syllabus

This use is given again in a grammar reference section at the back of the Level 1 coursebook.

Level 2 reviews the uses of way which are highlighted in Level 1 and goes on to show the pattern:

N + PREP/OF + -ING

in Unit 6 which is entitled ‘What's the best way of travelling to Paris?’ A ‘Wordpower’ exercise in the same unit offers another summary of the meanings of way:

<table>
<thead>
<tr>
<th>way</th>
<th>Look up way in the Lexicon. Which meanings does way have in these examples?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a)</td>
<td>This word can be used in many different ways</td>
</tr>
<tr>
<td>b)</td>
<td>I like the way he sings that son.</td>
</tr>
<tr>
<td>c)</td>
<td>After the class, or on the way home …</td>
</tr>
<tr>
<td>d)</td>
<td>The cheapest way would be to go by bus.</td>
</tr>
<tr>
<td>e)</td>
<td>Er, sorry, is that in the way?</td>
</tr>
<tr>
<td>f)</td>
<td>It’s interesting the way computers have changed our lives.</td>
</tr>
<tr>
<td>g)</td>
<td>I can remember thinking that way about teachers.</td>
</tr>
<tr>
<td>h)</td>
<td>The American way of life is very different.</td>
</tr>
<tr>
<td>i)</td>
<td>I can go back the way I came.</td>
</tr>
</tbody>
</table>

(NB the above exercise was accompanied by cartoon pictures to illustrate some of the examples.)

Way is also one of the words selected for inclusion in a lexicon at the back of the Level 2 coursebook:

<table>
<thead>
<tr>
<th>way 6</th>
</tr>
</thead>
</table>

1 Way refers to the manner In which a person or thing behaves or acts, or the certain style someone or something has, or feeling or attitude of a person. EG Just look at the way he eats! It's horrible!

2 Way refers to the means or method by which something is done, or how it happens.

EG The best way of getting to Paris is by train and boat. (64)

3 Used with reference to a direction, distance, route, road, path or journey somewhere.

EG Which way do I go? ‘Turn right at the shops, and go all the way down that road.’

By the time learners reach Level 3 they have had ample opportunity to become familiar with the main meanings and patterns associated with the word way. In Level 3 the word occurs 87 times. There is also a grammar section in
Level 3 which reviews the pattern:
N + PREP/OF + -ING

144 Grammar

· Of + ing
Some words are very commonly followed by of + ing. Look at these examples and make a list of words followed by of:

1 Another way of doing it is to work abroad. (140)
2 I think it's more a question of specializing in the country in which you work. (140)
3 Their first memory of singing together was during their days as Boy Scouts. (13)
4 His prayer had been answered and he gave up the idea of committing suicide. (36)
5 I always had this fear of falling downstairs. (34)
6 This would have the twofold effect of getting the job done cheaply and making it safe for the local people to cross the river. (97)
7 He took every opportunity of visiting the zoo. (91)
8 So the thought of competing with a three year old is quite difficult. (106)
9 . . . how to reduce the risk of falling a victim to violent crime. (150)
10 The POW Group also accuse the government of refusing to provide water as a deliberate policy. (163)
11 It would have to keep right on going if he was to have any chance of winning it now. (229)
12 And then he hits on this crazy plan of jumping overboard... (243)

This is another way of talking about ideas and actions. You could rewrite sentence 4 like this:
. . . . he gave up the idea that he would commit suicide.

Can you rewrite sentence 5 in the same way?
Sentence 1 can be rewritten like this:

Another way to do it is to work abroad.

What about sentences 7 and 12?
38 The Lexical Syllabus

The importance of ‘way’

As the third most frequent noun in English, the word way is important in its own right. It is unlikely that a learner will get very far without the need to express the kind of meaning normally encoded in English as:

The best way to . . . is to . . .
   by _ing...
One way of _ing . . . is to . . .
   by _ing...

and so on. And these phrases provide a very typical environment for the phrases 'of -ing . . . ', 'by _ing . . . ', and of the use of the infinitive with to followed by part of the verb be as in:

The idea is to score as many points as you can.
One possibility would be to start from the beginning again.

The commonest patterns in English occur again and again with the commonest words in English. If we are to provide learners with language experience which offers exposure to the most useful patterns of the language, we might well begin by researching the most useful words in the language.