# 3 Discourse analysis and vocabulary

> 'When *I* use a word,' Humpty Dumpty said, in rather a scornful tone, 'it means just what I choose it to mean – neither more nor less.'
>
> 'The question is,' said Alice, 'whether you *can* make words mean so many different things.'
>
> **Lewis Carroll: *Through the Looking Glass***

## 3.1 Introduction

Bringing a discourse dimension into language teaching does not by any means imply an abandonment of teaching vocabulary. Vocabulary will still be the largest single element in tackling a new language for the learner and it would be irresponsible to suggest that it will take care of itself in some ideal world where language teaching and learning are discourse-driven. The vocabulary lesson (or part of a lesson) will still have a place in a discourse-oriented syllabus; the challenge is to bring the discourse dimension into vocabulary teaching alongside traditional and recent, more communicative approaches (e.g. Gairns and Redman 1986). Therefore, in this chapter we shall look at research into vocabulary in extended texts in speech and writing and consider if anything can be usefully exploited to give a discourse dimension to vocabulary teaching and vocabulary activities in the classroom. Most are already in agreement that vocabulary should, wherever possible, be taught in context, but context is a rather catch-all term and what we need to do at this point is to look at some of the specific relationships between vocabulary choice, *context* (in the sense of the situation in which the discourse is produced) and *co-text* (the actual text surrounding any given lexical item). The suggestions we shall make will be offered as a supplement to conventional vocabulary teaching rather than as a replacement for it.



## 3.2 Lexical cohesion

One recent attempt at studying vocabulary patterns above sentence level is Halliday and Hasan's (1976) description of lexical cohesion. Related vocabulary items occur across clause and sentence boundaries in written texts and across act, move and turn boundaries in speech and are a major characteristic of coherent discourse. The relations between vocabulary items in texts described by the Halliday–Hasan model are of two principal kinds: *reiteration* and *collocation*.

It is debatable whether collocation properly belongs to the notion of lexical cohesion, since collocation only refers to the probability that lexical items will co-occur, and is not a semantic relation between words. Here, therefore, we shall consider the term 'lexical cohesion' to mean only exact repetition of words and the role played by certain basic semantic relations between words in creating *textuality*, that property of text which distinguishes it from a random sequence of unconnected sentences. We shall consequently ignore collocational associations across sentence boundaries as lying outside of these semantic relations.

If lexical reiteration can be shown to be a significant feature of textuality, then there may be something for the language teacher to exploit. We shall not suggest that it be exploited simply because it is there, but only if, by doing so, we can give learners meaningful, controlled practice and the hope of improving their text-creating and decoding abilities, and providing them with more varied contexts for using and practising vocabulary.

Reiteration means either restating an item in a later part of the discourse by direct repetition or else reasserting its meaning by exploiting *lexical relations*. Lexical relations are the stable semantic relationships that exist between words and which are the basis of descriptions given in dictionaries and thesauri: for example, *rose* and *flower* are related by *hyponymy*; *rose* is a *hyponym* of *flower*. *Eggplant* and *aubergine* are related by *synonymy* (regardless of the geographical dimension of usage that distinguishes them). In the following two sentences, lexical cohesion by synonymy occurs:

(3.1) The meeting commenced at six thirty. But from the moment it began, it was clear that all was not well.

Here, *commence* and *begin* co-refer to the same entity in the real world. They need not always do so:

(3.2) The meeting commenced at six thirty; the storm began at eight.

In (3.2) *commence* and *begin* refer to separate events, but we would still wish to see a *stylistic* relationship between them (perhaps to create dry humour/irony). Decoding the co-referring relationship in (3.1) is an interpretive act of the reader, just as occurs with pronouns (see section 2.2). In (3.3), cohesion by hyponymy occurs:

(3.3) There was a fine old *rocking-chair* that his father used to sit in, a *desk* where he wrote letters, a nest of small *tables* and a dark, imposing *bookcase*. Now all this *furniture* was to be sold, and with it his own past.

The superordinate need not be an immediate superordinate in the family tree of a particular word; it can be a *general word* (see Halliday and Hasan 1976: Ch. 6). Instead of *furniture* we could have had all these *items/objects/things*, which are examples of general superordinates. Other general superordinates, covering human and abstract areas, include *people*, *creature*, *idea* and *fact*. Reiteration of this kind is extremely common in English discourse; we do not always find direct repetition of words, and very often find considerable variation from sentence to sentence in writing and from turn to turn in speech. Such variation can add new dimensions and nuances to meaning, and serves to build up an increasingly complex context, since every new word, even if it is essentially repeating or paraphrasing the semantics of an earlier word, brings with it its own connotations and history of occurrence. In the case of reiteration by a superordinate, we can often see a summarising or encapsulating function in the choice of words, bringing various elements of the text together under one, more general term. Reiteration is not a chance event; writers and speakers make conscious choices whether to repeat, or find a synonym, or a superordinate.

Discourse analysts have not yet given us any convincing rules or guidelines as to when or why a writer or speaker might choose a synonym for reiteration rather than repetition, though some research suggests a link between reiteration using synonyms and the idea of 're-entering' important topic words into the discourse at a later stage, that is to say bringing them back into focus, or foregrounding them again (see Jordan 1985). Other research claims correlations between boundaries of discourse *segments* (as opposed to sentences or paragraphs) and re-entering of full noun phrases instead of pronouns (see B. Fox 1987). We may also be dealing with a lexical parallel to the grammatical topicalisation discussed in section 2.3. In (3.4), we can observe the importance of the words *route* and *way* in the foregrounding of the topic in this short extract, which is *how to* or *ways of* getting a contract, as indicated by the headline:

(3.4)

**HOW to get a contract**

**THE NORMAL route is to build up a following through live shows, send in tapes to record companies and then wait until someone "discovers" you. But there are other ways ...**

(from *News on Sunday*, 14 June 1987, p. 22)



Such usage as this is very common in English discourse. However, in practice, since our knowledge is inadequate, language teachers must content themselves with observing each case as it arises and, for the moment, work on raising an awareness of such phenomena where awareness is lacking, and, most important of all, providing the lexical equipment in L2 and practice of the skills to enable learners to create texts that resemble naturally occurring ones themselves. It means that it *is* important to make learners aware that synonyms are not just ways of understanding new words when they crop up in class, nor are they some abstract notion for the organisation of lexicons and thesauri, but they are there to be used, just as any other linguistic device, in the creation of natural discourse.

Another implication for language pedagogy is that material writers who create their own texts or who simplify naturally occurring ones should remember that disturbing the lexical patterns of texts may lead to unnaturalness and inauthenticity at the discourse level; simplification may mean an unnatural amount of repetition, for example, compared with the variation between exact repetition and reiteration by other means found in natural texts.

An analysis of the following newspaper extract according to Halliday and Hasan's principles, shows lexical cohesion at work:

(3.5) **BRITAIN'S green and pleasant meadows yesterday became "killing fields" with the start of the fox cub hunting season.**

More than 6,000 young foxes enjoying their first flush of life will be hunted down in the next three months to give inexperienced young hounds a blood lust.

But the dogs will also suffer.

Anti-hunt campaigners estimate that 7,500 young hounds will be destroyed because they fail to make the grade.

And many experienced hounds will be killed because they are too old to hunt.

The cub hunting season is just a curtain-raiser to the traditional pastime of killing adult foxes.

(from *News on Sunday*, 2 August 1987, p. 10)

*Fox cub* is reiterated as the near-synonymous *young foxes*; *young hounds* is repeated, but also covered by the superordinate *dogs* in the third paragraph. *Destroyed* and *killed* are also synonymous in this context (paragraphs 3 and 4).

Learning to observe lexical links in a text according to Halliday and Hasan's model could be useful for language learners in various ways. For one thing, it encourages learners to group lexical items together according to particular contexts by looking at the lexical relations in any given text. One of the recurring problems for learners is that words presented by the teacher or coursebook as synonyms will probably only be synonymous in certain contexts and the learner has to learn to observe just when and where individual pairs of words may be used interchangeably.

Little is known about the transferability of these lexical features of text from one language to another. Some languages may have a preference for repetition rather than linking by synonymy (such as is often said of literary and academic styles in Spanish, for example); sometimes learners may find the transfer of these skills to be easy and automatic. In either case the learner may need to use a range of vocabulary that is perhaps wider than the coursebook or materials have allowed for. Additionally, an awareness of the usefulness of learning synonyms and hyponyms for text-creating purposes may not always be psychologically present among learners; there is often a tendency for such areas of vocabulary learning to be seen as word study divorced from actual use, or at best only concerned with receptive skills. Conventional treatments of vocabulary in published materials often underline this word-out-of-context approach. Redman and Ellis's (1989 and 1990) vocabulary materials are exceptional in this respect.

---

### Reader activity 1

Trace all subsequent lexical reiterations of the underlined words in the text below. Are the reiterations in the form of near-synonyms, antonyms or hyponyms/superordinates?

**Cruise guards 'were asleep'**

DOZING guards allowed a group of peace campaigners to breach a missile security cordon yesterday.

The women protesters claimed to have walked right up to cruise launchers.

As sentries slept, they tip-toed past sentries at 3am and inspected a cruise convoy in a woody copse on Salisbury Plain.

Greenham Common campaigner Sarah Graham said : "For the sake of making things more realistic, the copse was protected by soldiers dug into fox-holes.

"And there were dogs rather than the usual reels of barbed wire."

But, she claimed, the American airmen were dozing by the launchers.

"One was kipping beneath one of the vehicles," she added.

Eventually, one of the airmen "woke up" and spotted the women, who had been trailing the convoy from the Greenham Common base in Berkshire since Tuesday.

The Ministry of Defence confirmed there had been an incident.

Ten women had been arrested, charged with trespassing and released on bail.

(from *News on Sunday*, 2 August 1987, p. 15)

---

## 3.3 Lexis in talk

There is no reason why the model of lexical relations in text outlined above should not also be applied to spoken data (see Stubbs 1983: 22–3). When we do this, we find that interesting observations can be made concerning how speakers reiterate their own and take up *one another's* vocabulary selections in one form or another from turn to turn and develop and expand topics in doing so. We shall refer to this phenomenon as *relexicalisation.* Let us look at a piece of data from Crystal and Davy (1975) and analyse it according to the general principles of the Halliday–Hasan model:

(3.6) (Two women are talking about 'Bonfire Night', the night when many people in Britain have large bonfires and fireworks in their gardens.)

A: No, I don't think we can manage a large bonfire but the fireworks themselves er we have a little store of . . .
B: Oh yes, they're quite fun, yes.
A: Mm yes, the children like them very much so I think as long as one is careful, very careful (B: Oh yes) it's all right.
B: Mm.
A: But erm I ban bangers, we don't have any bangers (B: Yes) I can't stand those (B: Yes) just the pretty ones.
B: Sparklers are my favourites.
A: Mm Catherine Wheels are my favourites actually but er you know we have anything that's pretty and sparkly and we have a couple of rockets you know, to satisfy Jonathan who's all rockets and spacecrafts and things like this.

(Crystal and Davy 1975: 28)

In A's first turn, she concludes a few previous exchanges about bonfires and then shifts the topic to the closely associated *fireworks.* B accepts the topic and just says that fireworks are fun. A takes up B's use of *fun,* and relexicalises it as *like them* and then adds that one should be careful. B simply replies 'mm'. A (who seems to work hardest at this point in developing topics) returns to the fireworks themselves and talks of *particular* fireworks: *bangers* and *pretty ones.* B continues this with *sparklers.* A comes back with *Catherine Wheels,* then repeats *pretty* and *sparkly* and expands to *rockets.* At the same time she exploits the double association of *rocket* to bring in its near-synonym *spacecraft,* thus expanding the topic to talk about her child, Jonathan.

Meanwhile, other relexicalisations are discernible: *fun* in B's turn, which becomes *like* in A's, is taken up as *can't stand* in A's next turn, then as *favourites* by B, and finally as *favourites* again by A, representing, by moving from near-synonym to antonym and vice versa, the sub-topic of 'likes and preferences' with regard to fireworks. Another relexicalisation chain can be seen in the sub-topic of 'precautions and restrictions': *careful, ban, don't have* carry this strand over the turn boundaries. This small number of lexical chains accounts for *almost all the content items* in the extract. The intimate bond between topic development and the modification and reworking of lexical items already used makes the conversation

develop coherently, seeming to move from sub-topic to sub-topic as a seamless whole. In this way the scope of the topics is worked out between the participants, with neither side necessarily dominating. This accords with the ethnomethodological approach to discourse analysis, which sees conversation as a *joint* activity that has to be worked at. Topics unfold interactively, rather than 'existing' as static entities; Wardaugh (1985: 139–40) refers to topic as a 'consensual outcome'. This is quite clearly so here. Speakers can throw topics into the ring, but whether they are taken up or die depends on the other speaker(s); if one speaker insists on pursuing his/her topics, ignoring the wishes of others, this is precisely when we recognise deviance into monologue or complain later to our friends that 'X was hogging the conversation'. Utterances by one speaker are an invitation to a response by another (see Goffman 1976); the initiating utterance puts an obligation on the responding speaker to make his/her turn both relevant to the previous turn and a positive contribution to the forward moving of the discourse (see Vuchinich 1977). Relexicalisation of some elements of the previous turn provides just such a contribution to relevance and provides other important 'I am with you' signals to the initiator.

Topics unfold, and the vocabulary used by the speakers offers openings for possible development, which may or may not be exploited. The 'conversation' class where topics are pre-set may be a straitjacket to this natural kind of development; a safer course of action might be to see pre-set topics merely as 'starters' and not to worry if the discourse develops its own momentum and goes off in unpredictable directions.

---

## Reader activity 2

Look at this extract from Svartvik and Quirk's data and trace the repetitions and relexicalisations of the italicised items, in the way that was done for the fireworks text (the transcription is simplified here):

A: You're *knitting* . . . what are you knitting, that's not a tiny *garment*.
B: No (A: laughs) no it's for me, but it's very plain.
A: It's a *lovely* colour.
B: It's nice.
A: Yeah, I never could take to knitting except on these double-O needles with string you know, that's my sort of knitting.
B: Yeah.
A: It grows quickly.
B: Yeah I get very fed up.
A: It's just the process though . . . do you sew? I used to sew a lot when . . .
B: No I don't.



A: In the days when I was a human being.
B: I have aspirations to make marvellous garments you know.
A: Well it's so cheap you know, this is the thing.
B: Yes.
A: Particularly, I think you probably like the sort of clothes I like anyway, which is fairly simple, things like summer dresses which are just straight up and down you know, with a scoop neck.
B: Particularly with those shifts, I mean you're well away aren't you.
A: Yes, oh yes, terribly cheap.

(Svartvik and Quirk 1980: 83–4)

---

Other linguists' data, in analyses where they have been interested in discourse features such as agreement/disagreement patterns and everyday discussion, also show regularly recurring vocabulary patterns where speakers use synonyms, hyponyms and antonyms to perform conversational functions (see Pomerantz 1984 and Pearson 1986, for example). In Pearson's data, people did not typically agree or disagree with phrases such as 'I agree' or 'I disagree' (beloved of English coursebook writers); rather, there seemed to be a preference for simply using some sort of lexical relation between turns.

The way in which we can observe speakers moving from superordinates to hyponyms and from synonyms to antonyms and back again is a common feature of conversation and learners can be equipped to use this skill by regular practice. As with written texts, in English at least, speakers do not just repeat the same items endlessly. This may be so in all languages and the behaviour itself may be easily transferable (but see Hinds 1979, for interesting observations on the preference for direct repetition in Japanese conversation). However, to behave in this natural way in a foreign language, the learner needs to have a fairly rich vocabulary, and to have at his/her fingertips the synonyms, antonyms, etc. of the words that are 'in play'. Once again, the issue is how to relate abstract notions such as synonymy and hyponymy to discourse skills, rather than just teaching them as disembodied properties of word lists.

Encouraging recognition of the communicative value of these lexical relations can start at quite an early stage in language learning, as soon as the necessary vocabulary is encountered. Simple cue and response drills for pairwork can train the learner in immediately associating synonyms and antonyms, or a superordinate with its hyponyms, and vice versa (see Redman and Ellis 1989 for examples).

## 3.4 Textual aspects of lexical competence

A somewhat different type of lexical relation in discourse is when a writer or speaker rearranges the conventional and well-established lexical relations and asks us, as it were, to adjust our usual conceptualisations of how words relate to one another for the particular purposes of the text in question. In one way or another, our expectations as to how words are conventionally used are disturbed. A simple example is the following extract from a review of a book on American military planning:

(3.7) The depressing feature of Allen's documents is the picture which emerges of smart but stupid military planners, the equivalent of America's madder fundamentalists, happily playing the fool with the future of the planet.

(*The Guardian*, 13 November 1987: 15)

Here, two words, *smart* and *stupid*, frequently occurring in the language as antonyms, and therefore incompatible, are to be interpreted as compatible descriptions of the military experts. To do this we have to adjust our typical expectations of how the two words operate as a related pair. One reasonable interpretation would be that the experts are clever ('smart') but morally reckless ('stupid'); to interpret them as meaning 'intelligent but unintelligent' would clearly be a nonsense.

Similarly, groups of informants faced with the following advertisement text react with mild surprise if the last two words are first covered up and then revealed:

(3.8) Just brush one generous coat of Hammerite directly on to metal. Within 15 minutes it's dried to a smooth, hammered-enamel finish that shrugs off dirt and water just like a non-stick pan. You get all of this in a choice of ten attractive colours. Plus black.

(from *Weekend*, 23–29 May 1984, p. 19)

In many situations *black* is an unexceptionable member of the 'colour' set of adjectives (such that the remark 'he/she wears really attractive colours, blacks and reds, you know . . .' would be quite normal). Here we are expected to place *black* outside of the range of 'attractive colours' and to consider it as a separate entity. Such an adjustment probably has no great permanent implications for the place of *black* in our mental lexicon (though we might be unconsciously on our guard and less surprised if we



encountered the relationship of exclusion again, especially in the context of paints, perhaps) and, as in the case of *smart* and *stupid*, no necessary implications that such relations have language-wide validity.

Alongside these eye-catching disturbances of our lexical expectations are other, less obvious kinds of lexical readjustments. These are lexical relations that are valid in particular texts only, and whose interpretations may not correspond to dictionary definitions. The good reader/listener has to decide when words are being used as more or less synonymous (or in what Bailey (1985) calls 'functional equivalence') and, conversely, when those same words may be being used in a way that focuses on the *difference* in meaning-potential.

Discourse-specific lexical relations can be called *instantial* relations, borrowing the term from J. Ellis (1966) (see also Hasan 1984). They are found frequently in spoken and written texts, and are probably a universal feature in all languages. The problems learners tend to encounter with such uses are usually more psychologically-generated; it is not that they have never encountered ad hoc rearrangements of predictable lexical usage, but more that they come to texts (especially reading comprehension texts), with the expectation that words have rather fixed relationships with one another because they have correspondingly fixed meanings, and vice versa. The task of the teacher is mainly to raise an awareness that typical vocabulary relations are often readjusted in individual texts, and, of course, to assist learners where necessary in interpreting such reorderings. Instantial relations often represent important stylistic features in texts, either in the sense of creative lexical usage, or perhaps as devices of evaluation or irony or for particular focus (e.g. the *smart/stupid* relation); by definition, each case has to be interpreted individually.

---

### Reader activity 3

#### *Instantial relations*

*Financial* and *economic* are very often used synonymously (e.g. 'The government has closed down the unit for financial/economic reasons'). How are they used in the following text? What possible interpretations could be put on the writer's choice of the words? (The writer is criticising a proposal to close a railway line in the north of England.)

> The accountants can produce as many figures as they like to prove that there are financial reasons for closing it. But there are no economic reasons. Already the campaign to keep the line has inspired many new initiatives along its length. It is an asset only now being fully appreciated in every sense.
>
> (*Country Living*, May 1988: 19)

## 3.5 Vocabulary and the organising of text

A distinction is often made between *grammar* words and *lexical* words in language. This distinction also appears sometimes as *function* words versus *content* words, or *empty* words versus *full* words. The distinction is a useful one: it enables us to separate off those words which belong to *closed systems* in the language and which carry grammatical meaning, from those that belong to *open systems* and which belong to the major word classes of noun, verb, adjective and adverb. *This*, *that*, *these* and *those* in English belong to a closed system (as do the pronouns and prepositions) and carry the grammatical meaning of 'demonstratives'. *Monkey*, *sculpture*, *noise* and *toenail* belong to open-ended sets, which are often thought of as the 'creative' end of language. In between these two extremes is another type of vocabulary that has recently been studied by discourse analysts, a type that seems to share qualities of both the open and the closed-set words. Let us consider a paragraph taken from an article in a learned journal:

(3.9) Here I want to spend some time examining this issue. First I propose to look briefly at the history of interest in the problem, then spend some time on its origins and magnitude before turning to an assessment of the present situation and approaches to its solution. Finally, I want to have a short peek at possible future prospects.

(W. J. Kyle, *Annals of the GGAS*, University of Hong Kong, 1984, no. 12: 54–66)

What is this article *about*? Controlling pests on fruit trees? Designing examinations for secondary schools? The possibilities are countless. What we are lacking here is the vocabulary that would identify the *field* of discourse. These sentences tell us a lot about the structure of the article, but nothing about the author's subject matter. They tell us that the tenor is relatively formal (it is hardly likely that this is someone explaining informally to a friend why he/she has never liked boiled eggs), but with an element of informality ('a short peek'). They tell us that a problem and its possible solutions will be examined, and that one part of the text will deal with the past, another with the future. So the words in our example do quite a bit of lexical work (they are not as 'empty' as grammar words are often said to be), but, in another sense, we need to seek elsewhere in the text for their content, what we shall call their *lexicalisation*. In our mystery text, the *this* of 'this issue' tells us that we can look to the preceding text to find out what the issue is; the lexical meaning of *issue* tells us to look for something problematic, something that is a matter of public debate, etc. 'The problem' works in a similar way. *Assessment* will identify with a portion of the text where something is being judged or evaluated; *solution* will be matter which can be counterposed to the 'problem', and so on. So these words stand in place of segments of text (just as pronouns can); a segment may be a sentence, several sentences or a whole paragraph, or more. We, the reader, (or listener if our example had been, say, a lecture) match the words with the segments, and, if we have decoded the text correctly, can render an account of what 'the problem' is, or what 'the prospects' are, according to the author. We shall call words such as *issue*, *problem* and *assessment discourse-organising* words, since it is their job to organise and structure the argument, rather than answer for its content or field. They are examples of the general phenomenon of *signalling* discussed in Chapter 1. Further examples may be seen in this extract:



(3.10) Week by week the amount of car traffic on our roads grows, 13 per cent in the last year alone.

Each day as I walk to work, I see the ludicrous spectacle of hundreds of commuters sitting alone in four or five-seater cars and barely moving as fast as I can walk.

Our traffic crisis now presents us with the classic conservation dilemma – too many people making too much demand on inadequate resources.

There are four possible solutions: One, provide more resources, in this case build more roads and car parks; two, restrict the availability of motorised transport by artificially raising the price of vehicles and fuel; three, license only those with a good reason for needing motorised transport and prohibit unnecessary use; four, reduce the average size of motor vehicles, especially those used for commuting purposes.

(from *Cambridge Weekly News*, 22 September 1988, p. 11)

The reader may be curious to know what extract (3.9) was about: in fact it is a study of the pollution of Hong Kong's streams, coastal waters and seashore. Pollution as a subject could be presented to the reader in a variety of ways; the author might have presented a series of claims and counterclaims about pollution, or perhaps a general statement about types of pollution and then details of these types. Our author chooses to present it as a *problem*, with *responses* ('approaches') to the problem and an *evaluation* ('assessment') of responses, in other words as a problem–solution text (see section 1.10). This is clearly signalled to the reader in our quoted extract. So, as well as representing text-segments, some of the discourse-organising

words we are examining additionally give us indications of the larger text-patterns the author has chosen, and build up expectations concerning the shape of the whole discourse.

From this account of the work of certain words in organising discourses it will be apparent that the language learner who has trouble with such words may be disadvantaged in the struggle to decode the whole text as efficiently as possible and as closely as possible to the author's designs. If the discourse-organising words are seen as signals of the author's intent, then inability to understand them or misinterpretation of them could cause problems.

But just how many such words are there in a language like English? What is the size of the task facing the teacher and learner in this particular lexical area? Some linguists have attempted to provide answers, but probably no one has compiled a complete list. Winter (1977 and 1978) has provided checklists, which teachers and material writers may find useful, of what he calls *vocabulary 3*, a precisely delimited sub-set of this more general set of discourse-organising words. Here is a selection of the list from Winter (1978):

(3.11) achieve, addition, alike, attribute, basis, case, cause, change, compare, conclude, confirm, consequence, contrast, deny, depend, differ, distinction, effect, equal, exemplify, explanation, fact, feature, follow, form, general, grounds, happen, hypothetical, instance, instrumental, justification, kind, lead to, manner, matter, means, method, opposite, point, problem, real, reason, replace, respect, result, same, similar, situation, state, thing, time, truth, way, etc.

Francis (1986) focuses on what she calls *anaphoric nouns* and gives extensive examples of nouns that frequently occur to refer back to chunks of text in the way that 'this issue' did in our first example. Here is one of her lists:

(3.12)

| | | | |
|---|---|---|---|
| abstraction | analysis | approach | assessment |
| assumption | attitude | belief | classification |
| comparison | concept | concoction | confusion |
| consideration | deduction | diagnosis | distinction |
| distortion | doctrine | dogma | doubt |
| evaluation | evidence | examination | fabrication |
| falsification | fantasy | finding | formulation |
| hypothesis | idea | ideology | identification |
| illusion | inference | insight | interpretation |
| investigation | misinterpretation | misjudgement | misreading |
| notion | opinion | perspective | picture |
| pipedream | position | rationalisation | reading |
| realisation | reasoning | recognition | reflection |
| scenario | speculation | supposition | surmisal |
| tenet | theory | thinking | view |
| viewpoint | vision | | |

(Francis, 1986: 15)



Another useful source is Jordan (1984), which brings together a large number of texts and has a vocabulary index. These works are good sources for teachers and material writers interested in this area, but many teachers will find it as easy simply to collect examples of such words from the press, where non-narrative texts, of the type where the author is presenting views and arguments and where such vocabulary is most readily found, are plentiful. In vocational/specialist courses, the best sources are learners' own subject material.

It might, at this point, be worth reminding ourselves that discourse-organising words operate *predictively* in text as well as retrospectively: if a discourse organiser does not already have its lexicalisation in the earlier text we expect it to come later in the text and are on the lookout for it, at least the efficient reader is. In (3.10) above, *dilemma* and *solutions* point forward in the text and are lexicalised in the subsequent discourse.

Predictive skills are often emphasised in current reading skills materials (see, for example, Greenall and Swan 1986); the study of vocabulary and discourse organisation offers the possibility of a more structured approach to this kind of teaching and practice activity.

## Reader activity 4

The italicised words in the following texts represent either preceding or subsequent segments of the discourse. Identify those segments by underlining the appropriate words:

1. I am always being asked to agree with the proposition that the British are the most anti-intellectual people in Europe. What other European language contains that withering little phrase 'too clever by half'? Where else do thinkers squirm when they are called intellectuals? Where else is public support for the institutions of intellectual culture – the universities and the subsidised arts – so precarious?
   Behind *these questions* lies a deep-seated inferiority complex in the post-imperial British middle-classes about the parochial philistinism of their culture, . . .
   (Michael Ignatieff, *The Observer*, 25 February 1990: 17)
2. *The issues* which emerge have beset the personal social services for generations – accountability, relationships with voluntary bodies, what their role is, for example, but the context is different.
   (*New Society*, 28 August 1987: ii)

Winter's work, and its extension in that of Hoey (1983), Jordan (1984) and Francis (1986), raises some interesting questions. First there is the question of whether it is possible to delimit a *procedural* vocabulary of such words that would be useful for readers/writers over a wide range of academic disciplines involving varied textual subject matters and genres. The notion of a procedural vocabulary is currently under debate in applied linguistics (see Widdowson 1983: 92–4; Robinson 1988). The procedural vocabulary is basically words that enable us to *do* things with the content-bearing words or *schematic* vocabulary. Another unanswered question is what happens if the most common signalling words are not known by the learner? Is coherent text-decoding seriously impaired or are such words the icing on the cake, especially in the case of production? Thirdly, if all languages have such text-organising vocabulary, can the teaching/learning process capitalise on transfer in some way? Are there direct and reliable translations for words like *point, argument, issue* and *fact* to and from other languages? Do languages with cognate words (e.g. French *problème*, Swedish *faktum*, Spanish *cuestión*) have an advantage here, or do they harbour false friends?

These questions cannot all be addressed in a book of this limited scope, but the vocabulary teacher and the learner can embark on their own research within their own situation as part of the 'learning-to-learn' process.

---

## Reader activity 5

Look back over the last few pages of text and note how many times I have used discourse-organising words to structure my text. Were you conscious of my use of them at the time of first reading? If so/not, what implications might this have for how language learners approach reading texts?

---

## 3.6 Signalling larger textual patterns

So far, the discourse-organising words we have looked at in greatest detail have been illustrated in their role of representing segments of text, parcelling up phrases and whole sentences. But we also noted in section 3.5 that they often have a broader textual function too, and that is to signal to the reader what *larger* textual patterns are being realised. We shall now look further at this phenomenon. In section 1.10, we saw an illustration of a *problem–solution* pattern. Discourse organisers often contribute to our awareness that a problem–solution pattern is being realised. In the following texts, items have been picked out in bold to exemplify this point. In



the first example, only the headline, the first paragraph and the last paragraph of a rather long newspaper article are given to show how organising words have been used to 'wrap round' a long problem–solution text:

(3.13)

*Headline* TV Violence: No Simple **Solution**

*Opening sentence* There is no doubt that one of the major **concerns** of both viewers and broadcasters is the amount and nature of violence on our television screens.

(*main text*)

*Closing sentence* The chief 'lesson' of all our viewing, reading and discussion is that there is no simple **solution** to the **problem** of violence on television.

(*The Observer*, 16 November 1986: 42)

The words in bold predict (*solution* in the headline, *concern*) and reinforce (*solution, problem*) the problem–solution pattern of the longer text (omitted here for space reasons), in which various responses to the problem of television violence are discussed and evaluated.

(3.14) In the past, the search for other worlds has been hampered by two factors. First, planets are tiny objects compared with stars: for instance, the sun, a typical star, is 300,000 times more massive than the Earth. Second, planets do not shine but only reflect light dimly from stars.

But Dr Campbell and his colleagues got round this problem by using high-resolution spectroscopy to measure accurately variations in a star's light. Slight differences in a star's light showed that many were being pushed and pulled out of their paths by unseen planets.

(from *The Observer*, 5 July 1987, p. 4)

Here both *problem* and *hampered* contribute to activating the problem–solution pattern, while *got round* indicates a positively evaluated response.

We can now begin to see that a number of vocabulary items characteristically cluster round the elements of larger patterns in texts. Words that often occur in the environments of the elements of problem–solution patterns include the following:

*Problem* concern, difficulty, dilemma, drawback, hamper, hind(er/ance), obstacle, problem, snag

*Response* change, combat (vb), come up with, develop, find, measure(s), respon(d/se)

*Solution/result* answer, consequence, effect, outcome, result, solution, (re)solve

*Evaluation* (in)effective, manage, overcome, succeed, (un)successful, viable, work (vb)

Likewise, other items characteristically cluster round the elements of claim–

counterclaim (or 'hypothetical–real') patterns, items such as *claim, assert, state, truth, false, in fact, in reality*, etc. Such words have been picked out in the following text:

(3.15) Historians are generally agreed that British society is founded on a possessive individualism, but they have disputed the origins of that philosophy. Some trace it back to the middle ages, others link it to the rise of capitalism. But the consensus is that the cornerstone of this society has been the nuclear family – where man the breadwinner holds dominance over his dependent wife and children. The values of individual freedom, self-reliance, individual advancement and crucially, the obligation of family duty to look after one's own in time of need are central to its operation. Within strict limits and under careful regulation, helping those less fortunate than oneself has been seen as part of the individual's obligation to society.

But, although most would accept that these values have been dominant, they would also acknowledge that the development of capitalist society saw the parallel growth of another ideology. Against individualism with its emphasis on individual freedom has been counterposed collectivism with its egalitarian values, and stress on the view that one individual's freedom cannot be paid for by the denial of freedom to others. The 19th century growth of trades unions, the cooperative movement and organised socialist political movements are all evidence of this opposition to dominant ideology. Because of this recognition of collective rights and responsibilities, feminists have always seen the granting and safeguarding of women's rights as lying within this socialist tradition.

(from *New Society*, 28 August 1987, p. 10)

Jordan (1984) is a useful work for teachers/material writers wishing to look at how particular vocabulary items have a tendency to cluster in each different segment of text-patterns such as the problem–solution pattern. He gives reference lists for the many textual examples he presents in his book and has a coding system for whether particular words typically occur in the 'problem' section or wherever. Part of his word list for the claim and counterclaim (or hypothetical–real) pattern is listed below:

(3.16) Whenever a writer needs to indicate doubt or uncertainty he uses a signal of hypotheticality to indicate this. Here are examples of such signalling words in the examples.

| | | | |
|---|---|---|---|
| according to | estimated | might | seems |
| apparently | evidently | old wives' tale | should |
| appears | expected | perhaps | signs |
| arguably | forecast | potential | so-called |
| believes | imagine | probably | speculation |
| claimed | likely | promises to be | suggests |
| considered | look | reported | thought |
| could | may | says | |

(Jordan 1984: 148)



These recurrent features of textual patterning may be exploited in vocabulary teaching/learning as a top–down phenomenon: once conscious of a larger text-pattern, the learner can be brought to an awareness of the rich vein of vocabulary that regularly realises it. As a bottom–up phenomenon, learners can bring together in their vocabulary records items that regularly occur in similar textual environments, e.g. the typical 'response' vocabulary of problem–solution patterns. Such lists can be added to over time to build up a rich, textually-based lexicon. It is yet another alternative to the random vocabulary list and the decontextualised, semantically-motivated list.

## Reader activity 6

Pick out words in the following texts which are strongly associated with either the problem–solution pattern or the claim–counterclaim pattern:

1. All western countries face a crisis in coping with the demands made on welfare provision by their growing elderly populations. The problem of resource scarcity is a real one. But perhaps not all countries have adopted so rigorously [as Britain] the view that care should be based on the family model.

   Scandinavia, for example, provides residential facilities for elderly people not wishing to remain at home or to live with their families, and those facilities are often available for use by local pensioners on a daily basis. Elderly people in the United States have developed communities of their own, supporting each other and running them by themselves, as their answer to increasing dependency. Some have argued against these 'age-dense' solutions, likening them to ghettos, but research suggests a high degree of consumer satisfaction.

   Examples from other countries demonstrate that there are alternative ways of tackling the issues of caring and dependency. The family model of care with the high demands made on women and lack of choice and frequent loneliness for the dependents is not the only solution.

(from *New Society*, 28 August 1987, p. 12)

2. Local authorities believe strongly in the involvement of the public sector and the need for public planning. They think that it is more important to protect jobs which are already in their area than to attract more from outside. And since they hold that production is the key to economic revival, they think it is more important to sustain manufacturing industry than to switch to alternatives, such as the service industry.

   Central government, on the other hand, places more faith in the private sector for its schemes, and it considers that public planning hinders rather than helps redevelopment. It usually dismisses planning as 'red tape'. Government is also more interested in attracting new jobs than protecting old ones. Above all, it believes that the market decides what sort of jobs should and should not be done.

   (from *New Society*, 28 August 1987, p. 20)

---

## 3.7 Register and signalling vocabulary

In claiming that particular vocabulary items tend to cluster round certain elements of text patterns we are ignoring the important fact that *register* (see page 32) is closely tied to lexical selection. Among the signals of the *problem* element we proposed *problem, drawback* and *snag* (see page 79). Clearly we might not expect to find *snag* occurring in this way in a formal scientific report, nor perhaps *come up with* as a signpost for *response* (*develop* would be a more predictable choice). Therefore, as in all matters, the relationship between vocabulary and register needs to be brought out when studying textual signalling. Lexical choice within the identified clusters will depend on the context (textbook, mazagine, news report, etc.), the author's assumptions about the audience (cultured/educated/readers of the popular tabloid press, etc.) whether the style is to be read as 'written' or 'spoken', and so on. Most of the texts we have looked at so far have been toward the 'written/formal/cultured' end of the spectrum. Here are two more, this time with a more informal, colloquial tone. They are presented to illustrate the fact that discourse-signalling words need not necessarily be only rather 'dry' academic words taken from the Graeco-Latin vocabulary of English. The relevant words are underlined:



(3.17) Put ordinary exterior varnish on your doors and window frames and in no time at all you'll wish you hadn't.

Wood shrinks and stretches when the temperature and humidity changes. Ordinary varnish, doesn't, so it cracks.

If you don't strip it off and start again you'll be in real trouble, your wood will be open to attack from fungus and rot, and quite frankly, it will look awful.

(Advertisement for Cuprinol from *The Observer*, 12 July 1987, p. 5)

Alongside more neutral items like *develop* and *reduce the risk* are informal, direct addresses to the reader: *you'll wish you hadn't* and *quite frankly, it will look awful* which create a pseudo-conversational register in which the element of *problem* is realised.

(3.18) Decide to tackle that troublesome moss on your lawn and you could find yourself going round and around in circles. Or at least backwards and forwards to your local garden centre.

Conventional moss treatments simply won't keep moss away for any length of time. You apply it and shortly afterwards your moss blackens and dies. You think all of your worries are over. Far from it. The little so and so's will turn up again as sure as the proverbial bad penny.

You're back where you started. And left with the choice of getting down on your hands and knees to weed it out or traipsing off to the shops for some more moss treatment.

So if you want to save yourself heartache, backache and a considerable amount of shoe-leather, insist on Lawnsman Mosskiller from ICI. You'll be rewarded with a moss-free lawn for the rest of the season.

Mix the sachet with water, stir, and sprinkle over your lawn. It's that simple.

(from *The Observer Magazine*, 6 April 1986, p. 12)

Here idiomatic phrases are used as signals of the response and its occurrence after a previous negatively evaluated response ('conventional treatments'). Idioms are often a problem for the teacher insomuch as it is not always easy to find natural contexts in which to present them. Research by Moon (1987) suggests that writers and speakers use idiomatic phrases to organise their discourse and to signal evaluation, far more frequently than previous linguistic studies of idiomaticity have suggested. Idioms are good metaphors for the kinds of textual segments we have been looking at (problem/response, etc.). Consider how some of the following could be used in informal discourses to suggest the problem–solution pattern:

*(to be) in a fix* — *to be up against a brick wall*
*to come up trumps* — *(sth) does the trick*
*to have a crack at (doing something)* — *to have a brainwave*
*(to be) up a gum tree*

Speakers and writers use these in informal situations to perform the same kind of organising and signalling functions that the more formal vocabulary does in written argumentation.

We can now begin to see just how important certain vocabulary items are in organising discourses. Admittedly, we have concentrated on reporting, expository and argumentative texts, but learners frequently have to tackle quite daunting and lengthy examples of these in their textbooks, and syllabus specifications often demand that they be studied. They are precisely the types of text that come festooned with comprehension questions in coursebooks and exams, and are regularly cited as being 'difficult', 'boring' and 'demotivating' for students by teachers. They are the texts that are hardest to unpack. Significantly, the kind of discourse-organising vocabulary that has occurred in most of our example texts, the Graeco-Latin words found in argumentation and exposition, is typical of the kind of vocabulary that research has claimed produces a 'lexical bar', a serious obstacle to progress in education, for children learning their first language (see Corson 1985). We should not underestimate the difficulties second language learners may experience with these words, particularly those who do not come from a Romance- or Germanic-language background.

Discourse-organising words are best presented and practised in their natural contexts. Simply looking them up in a monolingual dictionary can lead to a circularity of abstract definitions. Note how even a good, modern learner's dictionary like the Collins COBUILD (1987) dictionary defines *problem* in terms of *difficulty*, and *difficulty* in terms of *problem*:

(3.19) **problem** /prɒbləm/, **problems.** **1** A **problem** is **1.1** a situation or a state of affairs that causes difficulties for people, so that they try to think of a way to deal with it. EG. *...how families can try to solve these problems... ...the social problems in modern society... I think we may have a problem here... She has a weight problem... The problem is that she can't cook.*

**difficulty** /dɪfɪkəˈltiˈ/, **difficulties.** **1** A **difficulty** is something that is a problem for you. EG. *There are lots of difficulties that have to be overcome... The main difficulty is a shortage of time.*
**2** If something causes **difficulty**, it causes problems because it is not easy to do or understand. EG. *This can cause difficulty... ...questions of varying difficulty.*

(from *Collins COBUILD English Language Dictionary*, pp. 1143, 391)

## 3.8 Modality

One contribution that the study of vocabulary in naturally occurring discourses has made is to point up the all-pervasiveness of modality in spoken and written language. Modality is often thought of as the province of the closed class of modal verbs (*must*, *can*, *will*, *may*, etc.) and treated as part of the grammar of English, but a large number of 'lexical' words (nouns, adjectives, verbs and adverbs) carry the same or similar meanings to the modal verbs. For this reason, modality is dealt with here in our chapter on vocabulary rather than in Chapter 2.



Two notable studies of modality in large amounts of discourse, Holmes (1983) and Hermerén (1978), show a wide range of uses of the traditional class of modal verbs and of a vocabulary of lexical items carrying modal meanings, from the classic *epistemic* modality (concerned with degrees of certainty and possibility) to the *root* modalities (volition, permission, obligation). Both Holmes's and Hermerén's data show that, put together, other word classes express modality more frequently than modal verbs. The vocabulary of modality includes verbs such as *appear*, *assume*, *doubt*, *guess*, *look as if*, *suggest*, *think*, adverbs such as *actually*, *certainly*, *inevitably*, *obviously*, *possibly*, and nouns and adjectives related to them (for a full list, see Holmes 1988). In terms of frequency, the verbs and adverbs are considerably more frequent than the nouns and adjectives.

All these words carry important information about the stance and attitude of the sender to the message; they are concerned with assertion, tentativeness, commitment, detachment and other crucial aspects of *interpersonal* meaning (as opposed to ideational, or content, meanings). In the Hallidayan model of register they form a part of the *tenor* of the discourse. If we take a later part of one of our earlier texts, extract (3.10), we can see how modal vocabulary represents another aspect of discoursal meaning over and above the organisational and more general signalling vocabulary already analysed. Modal items are picked out in bold:

(3.20) **Inevitably**, objections will be raised to the promotion of the motor cycle as the saviour of our environment.

It is dangerous: it **can** be but three-fifths of all serious motor cycling accidents are caused by cars. So, by transferring some drivers from cars to motor cycles, the risk **can** immediately be reduced.

Department of Transport statistics have **shown** that a car driver is nine times more **likely** to take someone else with him in an accident than a motor cyclist, so riding a motor cycle is **actually** making a contribution to road safety.

(*Cambridge Weekly News*, 22 September 1988: 11)

Discourse analysts have demonstrated that modality is fundamental in the creation of discourse; *all* messages choose some degree of modality, even if it is only to make a *neutral* choice of bald assertion (e.g. 'The cat sat on the mat', as compared with the heavily modalised 'I suppose it's possible the cat just may have sat on the mat'). Language teachers have always paid attention to the modal verbs but, Holmes (1988) shows, in her survey of four ESL textbooks, that the larger vocabulary of modal lexical items is often under-represented in teaching materials, and there does seem to be a need to redress the balance in light of what natural data shows.

## Reader activity 7

Underline words conveying modality in this text:

**FOOD AND HEALTH**

### Can citrus peel harm?

Did you know that lemon and orange peel is coated with wax and chemicals?

The skin of almost all citrus fruit sold in the UK is treated with fungicides to stop it going mouldy. And the glossy surface is the result of bathing the fruit in wax.

Could the fungicides used on citrus peel be harmful – particularly since there's some evidence from laboratory tests that, in sufficient quantities, they may produce cancers or mutations in animals?

The Government doesn't feel there is any need to worry because the levels of fungicide permitted are very low. The levels are based on the recommendations of UK and international advisory bodies for the amount that can be consumed daily without any significant effect.

(from *Which?* January 1989, p. 4)

## 3.9 Conclusion

The study of vocabulary in discourse is concerned with patterns in text generated by the vocabulary relations that are found over clause and



sentence boundaries, the role of certain words in organising discourses and signalling their structure, and the relationship between these features of textuality and the register of the end product. Such an approach also offers an alternative motivation for the construction of word lists to supplement the traditional semantic-field orientation. Students themselves can be encouraged to collect items along discourse-functional lines, something which becomes more and more important as they embark on composition writing and argumentation in general, and something which can offer an organised backdrop in learning areas normally left to organise themselves. Once more, though, the whole enterprise depends on adapting what is useful in discourse analysis to current practices, and on teachers and material writers paying greater attention to the insights offered by naturally occurring data.

### *Further reading*

The standard work on lexical cohesion is Halliday and Hasan (1976); Hasan has since revised their model (see especially 1984).

Overall, not much research has been done on vocabulary and discourse, but further discussion of instantial relations may be found in McCarthy (1987 and 1988), and in Carter and McCarthy (1988: Ch. 5).

Cruse's (1975 and 1977) papers on hyponymy are innovative in that they look at language in use, while P. H. Fries (1986) and Ellis (1987) look at instantial synonymy.

On the use of superordinates in discourse, Wisniewski and Murphy (1989) is interesting.

McCarthy (1990) looks at further vocabulary features that cluster around text-organising words, and Lindeberg (1986) links lexical relations with thematic development in text.

A further paper that considers the re-entering of full noun phrases as opposed to pronouns is Hinds (1979).

Benson and Greaves (1973: 54–68) offer practical suggestions for the analysis of lexical relations in texts, based on the idea of lexical sets, and their paper on 'field of discourse' (1981) ties up the Hallidayan idea of collocation with the topics and institutional focuses of texts.

For more on topics as negotiated by participants see Richards and Schmidt (1983), and Brown and Yule (1983: 89).

Hoey (forthcoming) contains a thorough analysis and a novel view of the functioning of lexical cohesion.

King (1989) takes further the discussion of discourse-organising vocabulary.

Stubbs (1986) is a good, general paper on modality in discourse.

For more on modality see Perkins (1983) and Westney (1986)