6 Specialised uses of vocabulary

When learners have mastered the 2,000–3,000 words of general usefulness in English, it is wise to direct vocabulary learning to more specialised areas. Depending on the aims of the learners, it is possible to specialise by learning the shared vocabulary of several fields of study, for example academic vocabulary. Subsequently, the specialised vocabulary of one particular field or part of that field can be studied. Because many courses focus on learners who will do academic study in English, we will look first at academic vocabulary.

Academic vocabulary

Academic vocabulary is variously known as 'generally useful scientific vocabulary' (Barber, 1962), 'sub-technical vocabulary' (Cowan, 1974; Yang, 1986; Anderson, 1980), 'semi-technical vocabulary' (Farrell, 1990), 'specialised non-technical lexis' (Cohen, Glasman, Rosenbaum-Cohen, Ferrara and Fine, 1988), 'frame words' (Higgins, 1966), and 'academic vocabulary' (Martin, 1976; Coxhead, 2000). The division of the vocabulary of academic texts into three levels – general service or basic vocabulary, sub-technical vocabulary and technical vocabulary – is commonly made (although this distinction ignores low-frequency vocabulary with no technical or sub-technical features). Dresher (1934) made such a three-part division when looking at mathematics vocabulary for native speakers. Other writers have independently made a similar distinction. Typically, academic vocabulary lists include words like *accumulate*, *achieve*, *compound*, *complex* and *proportion* which are common in academic texts and not so common elsewhere.

Flood and West (1950) posed the question 'How many words are needed to explain everything in science to someone who has little or no training in science?' They answered this question by compiling a dictionary for readers of popular science and determining how large a defining vocabulary was needed. The resulting defining vocabulary

https://doi.org/10.1017/CBO9781139524759.008 Published online by Cambridge University Press

numbered just under 2,000 words – 1,490 words which made up the defining vocabulary of the *New Method Dictionary* (West, 1935), and 479 additional words needed for scientific terms. A revised version of this vocabulary can be found in an appendix to West's (1953) *General Service List* (*GSL*). Sixty of the 479 words are scientific terms like *alkali*, *cell*, *nucleus* and *molecule*; 125 are semi-scientific terms like *absorb*, *bulb*, *image* and *revolve*; the remainder are non-scientific words. It seems that a well-selected vocabulary of 2,000–2,500 words could be used to write popular scientific English, defining needed terms as they occurred.

There have been several studies that have investigated the vocabulary needed for academic study. Two of them (Campion and Elley 1971; Praninskas 1972) assumed that learners already knew a general service vocabulary and looked at academic texts to see what words not in a general service vocabulary occur frequently across a range of academic disciplines. Two other studies (Lynn 1973; Ghadessy 1979) looked at the words that learners of English wrote translations above in their academic texts. There were considerable overlaps between these four lists and they were combined into one list, the *University Word List* (*UWL*), by Xue and Nation (1984; also in Nation, 1990). This combined list of academic vocabulary was designed to consist of words not in the *GSL* but which occur frequently over a range of academic texts. Some examples: *acquire*, *complex*, *devise*, *fallacy*, *goal*, *imply*, *intelligent*, *phase*, *status*. This academic vocabulary, which contains over 800 word families, gives an 8.5% coverage of academic texts. Its low coverage of non-academic texts shows its specialised nature: 3.9% coverage of newspapers and 1.7% coverage of fiction (Hwang 1989).

The *University Word List* has been replaced by the *Academic Word List* (Coxhead, 1998). This list of 570 word families is based on a 3,500,000 token corpus of academic English which is divided into four groupings – Arts, Science, Law and Commerce – with each grouping consisting of seven sub-groupings such as psychology, mathematics, history, etc. Both range and frequency were used in choosing words for the list, with all word families in the list occurring in all four groupings and occurring at least 100 times in the total corpus. The frequency of each of the words in the list was compared with their frequency in a 3,500,000 corpus of novels. This was done to see which words in the list were truly academic words and which were general service words not in West's (1953) *GSL*. The list appears to provide slightly better coverage of academic text than the *UWL* even though it contains fewer words. The list is divided into ten sub-lists of 60 words each based on range and frequency criteria (see appendix 1).

The importance of academic vocabulary

There are several reasons why academic vocabulary is considered to be important and a useful learning goal for learners of English for academic purposes.

First, academic vocabulary is common to a wide range of academic texts, and not so common in non-academic texts. One of the earliest studies to look at this (Barber, 1962) is typical of the many small-scale studies that followed it. Barber compared three academic texts ranging in length from 6,300 to 9,600 tokens. This was done before computers were available for such research and, although the corpus was small, the analysis was very time-consuming and very carefully done. The finding of academic words common to the texts influenced a lot of thinking about English for Specific Purposes. Several subsequent studies have confirmed that it is possible to create an academic vocabulary common to a range of academic writing (Campion and Elley, 1971; Praninskas, 1972; Hwang, 1989). There has been little research comparing the frequency of specific academic words in academic and non-academic texts (Cowan, 1974), but the studies that have done this (Coxhead, 1998) show a big contrast in frequency.

Second, academic vocabulary accounts for a substantial number of words in academic texts. There are two ways of measuring this: by looking at the number of tokens (coverage) academic vocabulary accounts for; and by looking at the number of types, lemmas, or word families. Sutarsyah, Nation and Kennedy (1994) found that academic vocabulary (the *University Word List*) accounted for 8.4% of the tokens in the Learned and Scientific sections (Section J) of the *LOB* and *Wellington corpora*, and 8.7% of the tokens in an economics text. Coxhead (1998) found that her academic word list (the *Academic Word List*) covered 10% of the tokens in her 3,500,000 running word academic corpus and around 8.5% in an independent corpus. These are substantial percentages given that a general service third 1,000 word list would only cover around 4.3% of the same corpus.

Farrell (1990: 31), using a different classification of general service, semi-technical and technical vocabulary, found that out of 508 lemmas occurring more than five times in his corpus of electronics texts, 44% of the lemmas were semi-technical, and 27.7% technical. Sutarsyah, Kennedy and Nation (1994) found that in a 295,294 token economics textbook, there were 1,577 general service word families, 636 *University Word List* families, and 3,225 other families which included proper nouns, technical words and low-frequency words.

The coverage of each of the sublists in Coxhead's *Academic Word List* shows how even this specially selected list contains words with a

Table 6.1. *Coverage of the* Academic Word List

Sublists and number of words	% coverage	Number of tokens per 350 word page
1 60 words	3.6%	12.3 words per page
2 60	1.8%	6.0
3 60	1.2%	4.2
4 60	0.9%	3.2
5 60	0.8%	2.7
6 60	0.6%	2.4
7 60	0.5%	1.7
8 60	0.3%	1.3
9 60	0.2%	1.0
10 30	0.1%	0.5
Total 570 words	10.0%	35.1

wide range of frequencies. Table 6.1 shows that the first sublist contains 60 word families and covers 3.6% of the running words of each page of an academic text which is equivalent to 12.3 tokens per page. That is, on average, just over 12 words per page of an academic text will be made up of words from sublist 1. Around 35 tokens per page of most academic texts will be from the *Academic Word List*.

Third, academic vocabulary is generally not as well known as technical vocabulary. In a small-scale investigation of difficulties found by second language learners reading academic texts, Cohen, Glasman, Rosenbaum-Cohen, Ferrara and Fine (1988) found that non-technical vocabulary like *essential*, *maintain*, *invariable* was more often unknown than technical vocabulary. Cohen *et al.* identified some problems with such vocabulary in addition to simply not knowing the words.

- It was sometimes used with a technical meaning and sometimes not, and learners were not always aware of this.
- Learners were often not aware of related terms being used to refer to the same thing. That is, they did not pick up instances of lexical cohesion through paraphrase.

Anderson (1980) also found that sub-technical terms were the words most often identified as unknown by her learners in academic texts. Many learners get low scores on the *University Word List* section of the *Vocabulary Levels Test* (see appendix 3). In a study with native speakers of English, Cunningham and Moore (1993) found that

the presence of academic vocabulary in questions made those questions more difficult to answer.

Fourth, academic vocabulary is the kind of specialised vocabulary that an English teacher can usefully help learners with. This is in contrast to technical vocabulary where the teacher can often do little because of: lack of background knowledge in the subject; the need to learn technical vocabulary while learning the content matter of the technical field and; the mixture of specialist disciplines within the same group of English students. From this perspective, an academic vocabulary list represents an extension of the general service vocabulary for learners with academic purposes. It is a list of words that deserves a lot of attention in a variety of ways from both learners and teacher, no matter what their specialist area of academic study.

Trimble (1985: 129–130) suggests that a difficulty with some academic vocabulary is that it takes on extended meanings in technical contexts, and in different technical contexts there may be quite different meanings. For example, *fast* means 'resistant to' in medicine, 'a hard stratum under poorly consolidated ground' in mining, and 'said of colours not affected by light, heat or damp' in paint technology.

For the words making up an *Academic Word List*, it is important to discover if the list consists of words whose form is similar across disciplines but whose meanings are quite different, or if it contains words that draw on the same underlying meaning for their different uses. One way of comparing whether the same word form is used with a similar meaning in a different subject area would be to use a rating scale like that used by Nagy and Anderson (1984). Another way would be to look at elements of meaning. For example, Memory (1990) used a scoring key for scoring recall of definitions where one or two marks were allocated for each 'meaning element' (usually a content word) of a definition. So, for the term *wealth*, the presence of the word *product* or a close synonym in the recall of the definition was given two marks, *any* (1 mark), *economic* (2), *tangible* (2), *useful* (1), *scarce* (1), *transferable* (1). The original definition was 'any economic product which is tangible, in addition to being useful, scarce and transferable.' (*ibid*.: 44).

Making an academic vocabulary list

Academic vocabulary lists are usually made by analysing a corpus of academic English. This can be done in several ways. One way is to take an area of specialisation such as electronics (Farrell, 1990) or medicine

(Salager, 1983 and 1984) and classify the kinds of vocabulary found. Farrell (1990) drawing on Cowan (1974) defines semi-technical vocabulary as formal, context-independent words with a high-frequency and/or wide range of occurrence across scientific disciplines, not usually found in basic general English courses. This definition seems to hedge on range ('and/or wide range') but Farrell's later discussion makes it clear that range is a critical part of the definition, although he did not seem to use it when constructing his own list. Farrell created a semi-technical list consisting of 467 types from Section J of the *LOB Corpus*, using his intuition to remove the general words. Section J of *LOB* contains around 160,000 running words (eighty 2,000-word texts).

Salager (1983) used a comparison of the frequencies in the Kučera and Francis (1967) count with frequencies of words in a 100,000 running word corpus of medical English to divide the vocabulary of a medical corpus into three categories: Basic English, Fundamental Medical English, and Specialised Medical English. Fundamental Medical English largely corresponds to the *University Word List* and includes items such as *evaluate*, *differ*, *presence*, *factor*, *serve*. Salager classified the Fundamental Medical English terms into functional and notional categories such as description of process, cause and effect, measurement, description of illness or injury, in order to see what role these words played in medical discourse. Salager's (1983) study is noteworthy because of its comparison of a specialised corpus with a diverse corpus to highlight specialised vocabulary, and its attempt to explore the role that sub-technical vocabulary plays in academic discourse.

Ward (1999) suggests that it is not necessary and perhaps not desirable to set up the three levels of general purpose vocabulary, academic vocabulary and specialised vocabulary for learners who have clear specialised goals right from the early stages of their study. Ward created a list of frequent words from an engineering corpus (without distinguishing general purpose, academic and specialised vocabulary) and then applied this list to an independent set of engineering texts. He found that a 2,000 word family vocabulary was sufficient to provide over 95% coverage of the texts. This was much better coverage than that provided by the 2,000 words of the *GSL* and the 836 words of the *UWL*. Early specialisation helps strip away items that are useful in less specialised uses of the language but which may not occur in the specialised texts.

Another way of making an academic vocabulary list is to take a diverse academic corpus and see what words occurring with wide range and reasonable frequency are not part of general service high-

frequency vocabulary (Campion and Elley, 1971; Praninskas, 1972; Coxhead, 1998).

Coxhead's count was based on a 3,500,000 running word collection of recent academic articles and books. It was divided into four main faculty divisions: humanities, science, commerce and law. Each faculty was further divided into seven disciplines – history, education, etc. Range and frequency criteria were used in describing what words would be in the *Academic Word List*. The list, which we have already met several times, is called the *Academic Word List*; it assumes knowledge of the *GSL*.

A third way to make an academic vocabulary list is to collect words that learners write first language translations above in their academic texts (Lynn, 1973; Ghadessy, 1979).

There tends to be substantial overlap between these types of lists (Xue and Nation, 1984), indicating that there exists a general academic vocabulary which causes problems for second language learners.

Sequencing the introduction of academic vocabulary

Worthington and Nation (1996) examined the occurrence of academic vocabulary (the *University Word List*) in several series each consisting of 12 texts to see if the natural occurrence of such vocabulary in texts was sufficient for: providing coverage of the whole list and; providing a suitably gradual introduction to the words by not having too many or too few new items from the list in each text. They found that there were two major difficulties involved in using the natural occurrence of vocabulary in texts to determine the quantity and sequencing of vocabulary.

- An impossibly large number of texts would be needed to cover all of the vocabulary of the *UWL*. If texts were used as a means of sequencing vocabulary, it would be possible to do this for only a part, say 50%, of the *UWL*. Other ways of meeting the remaining vocabulary would have to be devised. These might include adaptation of texts, learning from lists, using specially prepared exercises, or simply leaving it somewhat to chance by encouraging extensive reading. This may be less of a problem for the *Academic Word List*.
- A very large amount of unfamiliar *UWL* vocabulary is met in the first three or four texts. This is far too much to be usefully dealt with in a few lessons and so there would need to be vocabulary learning preparation before meeting these texts.

https://doi.org/10.1017/CBO9781139524759.008 Published online by Cambridge University Press

It is clear from the difficulties involved in using texts to sequence the introduction of vocabulary that there would need to be a three step approach to sequencing.

1. First, the learners would need a gradual introduction over about five texts to the high-frequency, wide range 100–200 items in the *UWL*. This could be done by judicious selection or partial simplification of academic texts. The glossing could be done outside the text by the addition of glossaries at the side of the page or at the end of the text (Jacobs, Dufon and Fong, 1994), or a form of elaboration could be used where the words are explained in the text itself (Long and Ross, 1993). The partial simplification would involve the replacement or glossing of *UWL* words not in the first 100–200 items in addition to the replacement of some of the words not in the first 2,000 and *UWL*. At this step the sequencing is based on frequency and range.
2. Then, about 12 or more unadapted texts could be used to cover a further 200 or 300 items resulting in coverage of about half of the *UWL*. At this step the occurrence of vocabulary in texts determines the sequencing of the vocabulary.
3. Then, because the unknown academic vocabulary load of the texts would not be so heavy, learners could be encouraged to do large amounts of extensive reading of academic texts, both within and outside their subject areas. This could be accompanied by decontextualised learning of *UWL* words and study through formal exercises such as those involving word parts. At this step both frequency and range information, and occurrence in texts are used independently of each other to determine the sequencing of items to be learned.

The assumption behind this sequencing is that the occurrence of vocabulary in texts is the initial opportunity to meet the words but these then need to be learned to some degree so that they are not unknown items when met in subsequent texts. It is not sufficient to assume that simply meeting the items in a text is enough to ensure learning. This meeting would have to be accompanied or followed up by intensive study and opportunity for use, so that the knowledge of each item of vocabulary would be cumulatively enriched.

The nature and role of academic vocabulary

There have been attempts to study the role that academic vocabulary plays in an academic text. At one level the Latinate nature of the vocabulary adds a tone of formality and learnedness. It is this aspect

that Corson (1985 and 1997) describes in his work on the lexical bar. Some writers have also tried to examine the kinds of language functions and notions that academic vocabulary represents. Strevens (1973) suggests a classification of concepts which are general to science and technology and which reflect and convey the philosophy and methodology of science.

> *Discrimination and description* imply concepts of *identity* and *difference, processes, states, changes of state, quantification*;
> *Classification* implies concepts of *taxonomies* and the *co-occurrence of features*;
> *Inter-relation* implies concepts of *causality, influence,* and *interaction;*
> *Explanation* implies concepts of *evidence, intuition, hypothesis, experiment, models, theory;* etc
>
> (Strevens, 1973: 226–227)

Martin (1976) classifies academic vocabulary into: (1) the research process, (2) the vocabulary of analysis, and (3) the vocabulary of evaluation. These categories correspond to parts of a typical report of experimental research.

In a fascinating and insightful paper, Meyer (1990) suggests that there is a process of delexicalisation or grammaticisation going on in English where words which used to carry a full lexical meaning are now becoming more like function words. These include words like *affecting*, *barring*, *concerning*, *fact*, *process*, *matter* whose jobs in some other languages are done by function words or inflections. These words are becoming more grammatical and less lexical. Meyer classifies these words into three major categories.

1. Vocabulary relating to the domain of the text and the linguistic acts performed in it. This includes words like *argue*, *examine*, *survey*, *recommendation* which tell us what the authors are doing in their texts and what they ascribe to other authors.
2. Vocabulary describing scientific activities. This includes words like *analyse*, *examine*, *survey*, *implementation*. They relate closely to the categories described by Strevens (1973) above.
3. Vocabulary referring to the subject matter of scientific activities. This includes technical vocabulary but is by no means restricted to that. Meyer describes three main groups as examples:

    - Lexical expression of tense, aspect, modality, etc.: *current*, *present*, *recent*, *ability*, *impossibility*, *likely*.
    - Classification of states of affairs: *change*, *development*, *process*, *structure*, *quality*. Meyer notes that many of these words seem to be taking on the role of classifiers, that is,

general words to characterise a group of related items or a state of affairs. Classifiers can act as shorthand anaphoric items, act as general terms to be elaborated on later, and act as a kind of proper name for something already defined.

- Relations between states of affairs: this is a very diverse group. It can include quantitative changes *expansion*, *increase*, *decline*, *reduction*, causal relations *arising*, *affecting*, *contribute*, set inclusion *include*, *comprise*, and many others.

The academic vocabulary of texts allows the writer to generalise talk about scientific activities. Viewed from this perspective, academic vocabulary performs important roles in helping academics do what they need to do. 'Context-independent' vocabulary is an important tool of the writer in doing learned and scientific things.

Testing academic vocabulary

The *Vocabulary Levels Test* (appendix 3) contains a section based on the *Academic Word List*. If a learner intends doing academic study in English in upper secondary school or at university, then a score of at least 25 out of 30 is desirable. If a learner has a lower score than this then study of the items in the *Academic Word List* will be very useful. Academic vocabulary needs to be used productively as well as receptively so it is important to monitor learners' productive knowledge of these words. The productive levels tests devised by Laufer and Nation (1999) and the Lexical Frequency Profile (Laufer and Nation, 1995), which measures the proportion of various types of words in learners' free writing, are useful measures for this.

Learning academic vocabulary

For learners studying English for academic purposes, academic vocabulary is a kind of high-frequency vocabulary and thus any time spent learning it is time well spent. It is therefore important to have lists of academic vocabulary to help in planning and assessing learning. The four major strands of a language course – meaning-focused input, language-focused learning, meaning-focused output, and fluency development – should all be seen as opportunities for the development of academic vocabulary knowledge. Thus there should be listening and reading activities that encourage the learning of academic vocabulary and also language-focused activities such as direct teaching, learning from word cards, and word part analysis. Academic vocabulary is

largely of Latin or Greek origin and so learners can use word part analysis to help learn the vocabulary. Chapter 8 on word parts contains a variety of suggestions and exercise types. Farid (1985) uses a word part approach to words in the Praninskas (1972) list.

Because academic vocabulary is useful in speaking and writing, learners need the opportunity to use it in meaning-focused output activities, that is in speaking and writing in academic contexts. Corson (1995: 149) argues that using academic (Graeco-Latin) vocabulary helps users by letting them put their knowledge on display. Productive use of academic vocabulary is an important component of academic success. This can be encouraged through the presentation of prepared formal talks, discussions based on texts, writing summaries and critical evaluations of articles, and reviewing the literature of a topic.

Being able to use words fluently is part of vocabulary knowledge. Being able to access words quickly means that more processing time is available for concentrating on what to say rather than how to say it. Fluency is encouraged by repeated opportunities to work with texts within the learner's proficiency. One way that fluency can be encouraged is through the use of issue logs, an idea developed by Nikhat Shameem and Alison Hamilton-Jenkins at the English Language Institute at Victoria University of Wellington. Each learner chooses a topic to follow and become an expert on over several weeks during a pre-university English course. These topics might be terrorism, the Bougainville–New Guinea conflict, global warming or Thai politics. Each learner regularly finds and reads newspaper reports on their topic, listens to TV and radio news and writes a weekly summary of recent events related to their topic. They present a weekly oral report to members of their small group who discuss their report. These activities involve the learners using the four skills of listening, speaking, reading and writing with repeated attention to the same topic area. They thus soon bring a lot of background knowledge to their reading and discussion – ideal conditions for fluency development.

Knowing academic vocabulary is a high priority goal for learners who wish to do academic study in English. After gaining control of the 2,000 high-frequency words, learners need to then focus on academic vocabulary. Knowing the 2,000 high-frequency words and the *Academic Word List* will give close to 90% coverage of the running words in most academic texts. When this is supplemented by proper nouns and technical vocabulary, learners will approach the critical 95% coverage threshold needed for reading.

For native speakers, knowledge of academic vocabulary is a sign that they have been involved in academic study of various kinds. The vocabulary is the result of the experience. For second language learners who do not know the academic vocabulary of English it is important to

determine if they have gained academic skills and experience in their first language. If they have, then direct learning of the *Academic Word List* is one of a variety of useful ways to get control of this vocabulary. If however, second language learners of English have not done academic study in their first language, simply learning academic vocabulary will not make up for this lack of experience. They need to learn the academic vocabulary as they develop skill and experience in dealing with the appropriate range of academic discourse.

The research on academic vocabulary is encouraging, although much still remains to be done. It has been shown that it is possible to devise lists of academic words which are small enough to be feasible learning goals and which provide enough coverage of academic text to make them a very valuable part of a learner's vocabulary.

Technical vocabulary

The motivation for distinguishing technical vocabulary from other vocabulary is similar to that for distinguishing academic vocabulary from general service words, that is, to identify words that will be particularly useful for learners with specific goals in language use – reading academic texts in a particular discipline, writing technical reports, participating in subject specific conferences, and so on.

Having distinguished such a group of words it is possible to see how they affect language learning goals, particularly the number of words that need to be known to be able to cope effectively with language in use. The approach taken here is to use percentage of text coverage as an indicator of this. Having distinguished such a group of words it is also possible to examine how they would be learned and the role of teaching in the learning process.

Distinguishing technical vocabulary from other vocabulary

In essence, a technical word is one that is recognisably specific to a particular topic, field or discipline. There are degrees of 'technicalness' depending on how restricted a word is to a particular area. These degrees can be shown by classifying and exemplifying technical vocabulary in four categories, with Category 1 being the most technical and Category 4 the least. The four categories depend on the criteria of relative frequency of form and meaning.

Category 1. The word form appears rarely if at all outside this particular field.

Law: *jactitation*, *per curiam*, *cloture*
Applied Linguistics: *morpheme*, *hapax legomena*, *lemma*

https://doi.org/10.1017/CBO9781139524759.008 Published online by Cambridge University Press

Electronics: *anode*, *impedance*, *galvanometer*, *dielectric*
Computing: *wysiwyg*, *rom*, *pixel*, *modem*

Category 2. The word form is used both inside and outside this particular field but not with the same meaning.

Law: *cite* (to appear), *caution* (vb)
Applied Linguistics: *sense*, *reference*, *type*, *token*
Electronics: *induced*, *flux*, *terminal*, *earth*
Computing: *execute*, *scroll*, *paste*

Category 3. The word form is used both inside and outside this particular field, but the majority of its uses with a particular meaning though not all, are in this field. The specialised meaning it has in this field is readily accessible through its meaning outside the field.

Law: *accused* (n.), *offer*, *reconstruction* (of a crime)
Applied Linguistics: *range*, *frequency*
Electronics: *coil*, *energy*, *positive*, *gate*, *resistance*
Computing: *memory*, *drag*, *window*

Category 4. The word form is more common in this field than elsewhere. There is little or no specialisation of meaning, though someone knowledgeable in the field would have a more precise idea of its meaning.

Law: *judge*, *mortgage*, *trespass*
Applied Linguistics: *word*, *meaning*
Electronics: *drain*, *filament*, *load*, *plate*
Computing: *print*, *program*, *icon*

Words in Category 1 are clearly technical words. They are unique to a particular field in both form and meaning. Yang (1986) suggests that these words could be found by computer analysis using figures based on frequency and range. Someone who knows these words is likely to have knowledge of that field well beyond knowing the words. Indeed, it is likely that these words can only be learned and really understood by studying the field. They could not sensibly be pre-taught. Words in Category 2 are clearly technical words as the more general meaning of the word when used outside the field does not provide ready access to its technical use.

Words in Categories 3 and 4 are less obviously technical because they are unique neither in form or meaning to a particular field. The words, particularly in Category 4, are readily accessible through their use outside the field. A glance at a list of Category 4 words is sufficient to quickly identify what field is being examined. Murphey (1992), in a study of pop songs, found that the word *love* was extremely frequent, in fact among the top ten words along with *the*, *be*, *you*. *Love* is probably not a technical word, although it meets the criterion for Category

4, because pop songs are not considered an area of technical knowledge.

Categories 2 and 3 indicate that range based on form alone is not sufficient to make sensible decisions between what are technical words and what are not. In many cases, meaning must also be considered. The cline which exists from Categories 2 to 4 raises the question of whether a technical word needs to have a technical meaning, that is a meaning which is different from its uses outside a particular field, and if so how different the meaning needs to be. The cut-off point distinguishing technical words from non-technical words could come after category 2, 3 or 4.

Several researchers have used relative frequency or range as a way of distinguishing technical vocabulary. Bečka (1972) distinguished three types of words in specialised discourse: grammatical words and two kinds of lexical words – non-terminological words and terminological words. Grammatical words (sometimes called function words) are a small group of around 270 types (176 word families, see appendix 6), if numbers are included. They are generally of very high-frequency and very wide range. They consist of words that are not nouns, verbs, adjectives and adverbs. They include words like *the*, *of*, *she*, *but*, and *might*. Only a few function words (*hence*, *hither*, *thither*, *thrice*, *whereas*, *whither*) fall outside the most frequent 2,000 words. Terminological words (which could also be called terminology or technical terms) are words whose meaning requires scientific knowledge. To use them well we must know the science involved (Bečka, 1972: 48). Terminological words, because of their specialised meaning, are generally of moderate to low-frequency in texts and more importantly have very narrow range. In groups of texts classified under a range of different disciplines, terminological words will be among the lower-frequency words and will occur in a limited range of texts and disciplines. The critical statistical signal that a word is a terminological word is that it has narrow range (low textual and register frequency) – it occurs in only a few texts and disciplines – in relation to its frequency of occurrence. For example, *bargain* and *styrene* have the same frequency in Francis and Kučera (1982), but *bargain* occurs in eight different texts in seven different disciplines. *Styrene* occurs in only two different texts in two different disciplines. *Styrene*'s narrow range marks it as a technical word. Bečka's data is a compelling argument for the need to take account of range in the construction of word lists.

A study of the 6,000 lemmas in the ranked list of higher frequency lemmas in Francis and Kučera (1982) however indicates that most words with low range compared to other items of the same frequency are not technical terms. However a corpus of 2,000-word texts may not be the best kind of corpus to reveal striking differences in range

between terminological words and non-terminological words. Longer texts would give technical terms a chance to attain a frequency comparable to moderate frequency non-terminological words.

Sutarsyah, Nation and Kennedy (1994), in a study of an economics text, found that a group of around 30 words that were obviously related to the field of economics accounted for over 10% (1 in every 10) of the running words in the text. These were words like *cost*, *demand*, *price* (see Table 6.2). Over half of the 34 words from Table 6.2 are in the first 1,000 of the *General Service List* (West, 1953) and four-fifths are in the complete *GSL* or the *University Word List*.

Making lists of technical vocabulary

There are two systematic ways of developing lists of technical vocabulary: through the use of technical dictionaries, and through the use of corpus based frequency counts. Technical dictionaries can be regarded as technical vocabulary lists, but they do not contain frequency information useful in assessing the relative usefulness of the words, and they give little indication of how the words listed were chosen.

Using a dictionary

Using a dictionary to create a technical vocabulary list involves the methodological problems of sampling, defining what is actually being counted (what is counted as a word) and classification described in Nation (1993c). There is also the problem of choice of dictionary, because working from a dictionary means that the base data has been compiled by someone other than the researcher and the principles underlying its compilation are not likely to be described or apparent. Does the dictionary adequately represent the field which is being examined? Is the dictionary up to date? Is it as complete as possible? Does it reflect a US or UK bias?

Using a corpus-based frequency count

It is possible to use an existing frequency count that distinguishes range figures from frequency figures – the best example is the *American Heritage Word Book* (Carroll, Davies and Richman, 1971) – or to carry out a range and frequency count comparing a general corpus with a specialised corpus (see Sutarsyah, Nation and Kennedy, 1994). Some studies dispense with the general count (Farrell, 1990), but then there is greater dependency on intuitive judgement. Farrell created a list of technical vocabulary for the field of electronics based on a 20,017 token corpus containing 1,258 lemmas. The division of

Table 6.2. *Words in the most frequent 1,000 of the economics text that occur much more frequently than in the general academic corpus (from Sutarsyah, Nation and Kennedy (1994))*

Word family	Frequency in the economics text	Frequency in the general academic corpus
Price	3,080	90
Cost	2,251	91
Demand	1,944	102
Curve	1,804	83
Firm	1,743	41
Supply	1,590	86
Quantity	1,467	53
Margin	1,427	24
Economy	1,353	172
Income	1,183	96
Produce	1,237	167
Market	1,104	110
Consume	955	70
Labour	1,004	131
Capital	907	50
Total	946	114
Output	861	50
Revenue	763	10
You	866	118
Profit	733	27
Production	772	84
Average	777	90
Goods	705	21
Product	749	106
Trade	621	85
Buy	521	35
Wage	522	75
Monopoly	454	13
Percent	450	41
Million	445	42
Household	360	41
Equilibrium	328	21
Choice	339	39
Elasticity	333	34
Total frequency	3,080.00	2,322

Notes:
Words not marked are in the first thousand words of the GSL (19 out of 34 words)
Words marked word are in the second thousand words of the GSL (3 words)
Words marked **word** are in the *University Word List* (6 words)
Words marked *word* are not in any list (6 words)

https://doi.org/10.1017/CBO9781139524759.008 Published online by Cambridge University Press

the lemmas into basic, semi-technical and technical terms seems to have been largely done on an intuitive basis although frequency and range information were also consulted. The data in the Farrell study agrees with other studies of specialised corpora (Sutarsyah, Nation and Kennedy, 1994).

- A small group of highly topic related words (*current*, *fig.*, *voltage*, *circuit*) occurs very frequently, accounting for at least 10% of the running words in the corpus.
- Several of the words in this high-frequency group are in West's *General Service List of English Words* or the *University Word List* with related meanings (*circuit*, *field*, *energy*, *plate*, *connected*, *supply*, *positive*, *flow*). They are not completely technical but are also known in some form in non-specialised language.
- A few of the highly frequent words stand out as reflecting the nature of the discourse (*fig.*, *if*, *we* – the most frequent pronoun).
- Several words are clearly technical in form and meaning but have a very low-frequency of occurrence in the corpus. This would make it difficult to distinguish technical words purely on the basis of frequency.
- Several technical words have very low range *within* a specialised corpus (Farrell, 1990: 30). This means that the occurrences of a particular technical word are not spread evenly through a text, but cluster in a particular chapter or section. This reflects the occurrence of some proper nouns in novels (Hirsh, 1992), with some characters appearing for a short time and then never being mentioned again.

Learning technical vocabulary

Several writers (Cowan, 1974; Higgins, 1966; Barber, 1962) consider that it is not the English teacher's job to teach technical words. The words are learned through study of the field. As we shall see later, the use of general service words and academic words as technical words (*resistance* in electronics; *wall* (cell wall) in biology; *demand* in economics) means that the English teacher may be able to make a useful contribution to helping the learners with technical vocabulary. Strevens (1973: 228) points out that learners who know the scientific field may have little difficulty with technical words; a teacher who does not may have a great deal.

Godman and Payne (1981: 37) argue that a technical term only makes sense when other related terms are also known. This is perhaps another way of saying that knowing a technical word involves

knowing the body of knowledge that it is attached to. Flowerdew (1992: 208) notes that definitions in science lectures to non-native speakers occur systematically. The lecture may be organised around definitions of key terms in that topic area. This discourse role of definitions underlines the point that knowing the technical vocabulary is very closely related to knowing the subject area.

Considering the large numbers of technical words that occur in specialised texts, language teachers need to prepare learners to deal with them. Chapter 1, on the goals of vocabulary learning, presented some information on the statistical nature of technical words. From the learner's point of view, unknown technical words usually cannot be ignored when reading because they are closely connected to the topic being discussed. They are also difficult to guess from context if the reader does not already have a good background in that technical area. For the same reason, looking the word up in a dictionary does not bring much satisfaction. Clearly, learning technical words is closely connected with learning the subject.

Although English teachers are not usually well equipped to work with technical texts and the technical vocabulary they contain, they can help learners get accustomed to the idea that different uses of words may have a shared underlying meaning. The *wall* of a living cell shares important features with the *wall* of a house. Visser (1989) devised the following kind of exercise to deal with this. Learners can work individually or in pairs on the exercises.

interpret/ɪntɜːrprɪt/verb If you **interpret** something in a particular way, you decide that this is its meaning or significance. *Even so, the move was interpreted as a defeat for Mr Gorbachev . . . The judge says that he has to interpret the law as it's been passed . . . Both of them agree on what is in the poem, but not on how it should be interpreted.* **How would you interpret the meaning of this sign?** ▶	**interpret/ɪntɜːrprɪt/verb** If you **interpret** what someone is saying, you translate it immediately into another language. *The woman spoke little English, so her husband came with her to interpret . . . Three interpreters looked over the text for about three or four hours and found that they could not interpret half of it.* **Interpret this sentence into your language: 'I really like chocolate cake.'**	**What is the core meaning of this word?**

https://doi.org/10.1017/CBO9781139524759.008 Published online by Cambridge University Press

These exercises are easy to make and as well as improving knowledge of particular words, they get learners used to the idea that words 'stretch' their meanings. The sample sentences come from the *COBUILD* dictionary.

Memory (1990), in a study of 15- and 18-year-old native speakers, looked at whether technical vocabulary was best taught before, during or after reading and found no significant difference. Learning was tested by recall of definitions. The argument for learning technical vocabulary during reading emphasises the importance of seeing how a technical term fits into a framework of knowledge. It may thus be more revealing to also test using semantic mapping or some other measure that looks at how knowledge of a technical term is integrated into a field of knowledge. Learners should approach specialised vocabulary strategically, considering whether particular words are worth learning, and considering how they can be most efficiently learned.

The main purpose in isolating an academic or technical vocabulary is to provide a sound basis for planning teaching and learning. By focusing attention on items that have been shown to be frequent, and in the case of academic vocabulary of wide range, learners and teachers can get the best return for their effort.

Technical vocabulary however is only one kind of specialised vocabulary, and its occurrence is affected by factors that influence the use of all vocabulary. We will now look at these factors.

Vocabulary in discourse

So far in this book we have looked mainly at vocabulary as isolated words or in phrase and sentence contexts. But the main role of vocabulary is to convey messages in extended spoken and written texts. We will now look at the part played by vocabulary in discourse.

Vocabulary use in a text arises from the communicative purposes of the text. There are two related aspects to consider. First, vocabulary use signals and contributes to the uniqueness of the text, that is, what makes this text different from all other texts. Second, vocabulary use carries general discourse messages which are shared with other texts of similar types. Thus, when we examine what vocabulary use does in a text, we can look at the special features of the text, and we can also look at how these special features are examples of general language constraints and discourse requirements. We will look at the general discourse messages that vocabulary can carry and see how these can

affect particular texts. Table 6.3 lists the communicative messages of vocabulary in a text, classified according to Halliday's (1994) three major divisions of field, tenor, and mode.

Let us now look in more detail at each of the three sets of discourse functions.

Vocabulary and the information content of the text

Function words

It might be expected that because the function words of English are a small, largely closed group their frequency would be constant across a range of texts. However, this turns out not to be so. Although it is possible to predict that a small number of function words will account for a significant proportion of the running words of a text, the particular nature of the text will determine how frequent each function word is and its frequency relative to other function words.

The approximately 270 function word types (176 word families) account for 43–44% of the running words in most texts (Johansson and Hofland, 1989; Francis and Kučera, 1982). The unusually high-frequency of some function words in a text may indicate important features of the discourse. In the economics textbook used in the Sutarsyah, Nation and Kennedy (1994) study, *you* had a frequency much higher than its frequency in the general academic corpus, because the writer typically addresses his message directly to the reader to involve the reader in the topic: 'You have just been named chief economic strategist for OPEC.' Newton and Kennedy (1996) found different occurrences of prepositions and conjunctions in split information tasks compared with shared information tasks.

Topic related vocabulary

Goodman and Bird (1984) argue that there are two major kinds of word frequency studies, each giving quite different information. The most common kind of word frequency study examines a large range of texts to establish general service high-frequency words. A more neglected kind of frequency study looks at the frequency of words within a particular text. If we want to understand why and how particular kinds of words are used, we must study individual texts intensively. The frequency with which words occur in a text is a result of the characteristics of the particular text itself.

The most immediately striking finding when looking at a frequency ranked list of the vocabulary in a text is the way that topic related

Table 6.3. *The discourse functions of vocabulary*

The information content of the text

1. The vocabulary reflects the topic of the text through the frequent use of particular words.
2. The vocabulary shows the formality of the text through use of *Academic Word List* vocabulary and other Latinate vocabulary.
3. The vocabulary shows how technical the subject matter is through the use of technical vocabulary and deep taxonomies.
4. The vocabulary shows the writer or speaker's ideological position (Fairclough, 1989: 113–116), often through metaphor.

The relationship between the writer or speaker and the reader or listener

5. The vocabulary shows the power relationships and frequency of contact relationships between the writer and reader or speaker and listener through the use of vocabulary over the range of colloquial, spoken vocabulary to very low-frequency, learned vocabulary.
6. The vocabulary shows the writer or speaker's attitude to the subject matter or to others through the use of vocabulary over the range of emotionally involved to uninvolved.
7. The vocabulary shows the writer's wish to make the text accessible or inaccessible to certain readers or the writer's wish to communicate with the already initiated (Corson, 1985) through the selection of vocabulary and through defining in the text.
8. The pronoun use shows the writer's stance with regard to the audience.

The organisation of the text

9. The vocabulary signals the rhetorical stages or semantic structure of the text (e.g. problem-solution) (McCarthy, 1991).
10. The vocabulary shows the most important sentences in terms of drawing the main topics together (Hoey, 1991).
11. The vocabulary shows the connections between parts of the text through lexical cohesion (Halliday and Hasan, 1976), and use of 'grammaticised' words (Winter, 1978; Meyer, 1990).

words occur among the very high-frequency words. A brief glance at the most frequent content words in the list is usually sufficient to determine what the text is about. The following list contains the most frequent content words from part of a well known story.

little	25
pig	22
house	17
said	14

https://doi.org/10.1017/CBO9781139524759.008 Published online by Cambridge University Press

wolf	9
build	8
straw	7
man	6
catch	5
bricks	4
built	4
now	4
sticks	4

Typically, the most frequent content words in a text occur with a frequency per 1,000 words that is very much higher than their frequency per 1,000 words in other texts or in a collection of different texts. Sutarsyah, Nation and Kennedy (1994) found that a group of 34 words in an economics textbook were so frequent that they accounted for 10% (one word in every line) of the running words in the text. These words occurred with a frequency of up to sixty times the frequency with which they occurred in a more general corpus of similar size. Here are some of those words: *price*, *cost*, *demand*, *curve*, *firm*, *supply*, *quantity*, *margin*, *economy*, *income*, *produce*, *market*, *consume*, *labour*, *capital*, *total*.

Each text has its own topic vocabulary which occurs because of the message the text is trying to convey. The vocabulary gives the text part of its unique flavour. This has several important messages for language teaching. First, in the production of simplified material, it is important that any vocabulary grading scheme used is flexible enough to allow the use of topic related vocabulary that may not be in the lists used to guide the grading of the material. There can be rules regarding the repetition of these additional words to make sure that they have a chance of being learned and that they do not act as a burden to the reader. If a particular word occurs only once then it may be a burden but if it is repeated several times in the book then the initial learning effort is repaid by the opportunity to use that learning again when the word reoccurs. Most well designed graded reader schemes have rules of this kind.

Second, when learners are being asked to speak or write on a topic, their language production is likely to be more apt if they are given the chance to meet relevant topic related vocabulary before they produce. This can be done in a variety of ways – through topic related reading, discussion, direct teaching, or accompanying support materials.

Third, from a course design perspective, learners may need exposure to a range of topics if they are to develop a rich vocabulary. In the

Sutarsyah, Nation and Kennedy (1994) study of an economics text, only 548 of the second 1,000 words of the *GSL* occurred, compared to 796 in the more general corpus.

Fourth, as we have seen in Table 1.7 (see p. 17), different texts contain quite different amounts of academic vocabulary. This shows that the general topic of the text influences the type of vocabulary that occurs. In an unpublished study, Jenkins (1993) developed a vocabulary of children's books, and in a similar study Hwang (1989) found evidence of a newspaper vocabulary. Courses focusing on a limited range of text types could benefit from the development of a specialised vocabulary.

Fifth, teachers need to be careful when focusing on vocabulary in intensive reading. The content words that occur most frequently in a particular text may not be useful words when learners face a different text. Teachers may need to give most attention to less frequent words in a particular text that are high-frequency words across a range of texts. Today's teaching needs to help tomorrow's tasks.

Vocabulary and the relationship between the writer or speaker and reader or listener

Corson (1997) presents arguments to support the view that use of academic vocabulary is taken as evidence of being in control of the academic meaning systems, and is thus essential to academic success.

Academic vocabulary is overwhelmingly Graeco-Latin and is not easy to learn because words refer to abstract ideas, they are infrequent, and their forms do not reveal their meaning. Thus learners need to have 'a rich acquaintance with the specialist areas of discourse in which they appear, as well as frequent and motivated contact with the words themselves' (*ibid.*: 701). Corson argues that meeting words receptively is insufficient for using them well. They need to be used in motivated talk about text. Not all learners have access to this experience.

Studies involving the *University Word List* and the *Academic Word List*, have shown the very uneven spread of this vocabulary across different types of writing. It is uncommon in fiction (1.7% text coverage), moderately frequent in newspapers (3.9% text coverage), and very frequent in academic texts (8.5% text coverage). The frequent occurrence of this vocabulary is a sign of the formal academic nature of a text, or in Corson's (1997) terms, that the text is drawing on different meaning systems from those texts with little academic vocabulary.

Some work has been done on examining the way vocabulary reflects academic meaning systems. One area of attention has been in the use

of reporting verbs as in 'Barrington (1967) *states* that . . .' Thompson and Ye (1991) look at the way the wide range of reporting verbs in English can be used to reflect evaluation of the citations reported. There have also been attempts to relate the kind of academic vocabulary used to what academic discourse does: citing, evaluating, hypothesising, contrasting, relating and explaining (Strevens, 1973; Martin, 1976; Meyer, 1990).

The amount of technical vocabulary in a text and presence or absence of explanation of this vocabulary is a sign of the intended audience for the text. It is not always easy to decide what is a technical term, and there are degrees of 'technicalness'.

Vocabulary and the organisation of the text

The frequency of topic-related words is only one aspect of their occurrence in a text. In an insightful book, Hoey (1991) shows that by examining the number of lexical links between sentences in a non-fiction text, it is possible to: identify the sentences which are central to the topic of the text (these also tend to provide a reasonable summary of the text); identify the sentences which are marginal to the topic of the text; and identify where a topic is introduced and ends.

These lexical links include repetitions of words (in either the same form or in inflected or derived forms), paraphrase of various kinds (which includes synonyms and hyponyms), substitution (including pronouns) and ellipsis (Hoey, *ibid.*: 83). Essentially, sentences which are central to the topic of the text have more lexical links to other sentences, that is, they share more vocabulary that refers to the same ideas. In terms of discourse analysis, a major strength of Hoey's lexically based analysis of relationships is that it shows relationships between sentences that may be separated by several intervening sentences. These links often occur with the topic-related words and they can form 'lexical chains'. They can occur within a text and also between speakers in a conversation (McCarthy, 1991: 69).

Learners need to be able to see the links between the various forms of topic-related vocabulary; for example, they need to see that *biologist* and *scientist* are in fact referring to the same person in a particular text. For this reason, Hoey (1991: 241) suggests that topic-related words which form links should be given priority when glosses are provided to accompany a text. This would have the effect of helping the learner quickly make sense of the text, because the sentences central to the topic would be understood.The learning of unknown words would also be helped if their relationship with synonymous known items was

clarified. In addition, links through synonymy and paraphrase not only show shared aspects of meaning, but also highlight differences. Each new link can be part of a developing enrichment (McCarthy, 1991: 66). The results of Hoey's study can be used to justify:

- non-linear note-taking from the text
- looking for lexical links in order to understand the structure of the text
- not needing to understand every sentence to get the important ideas in a text
- the importance of stressing lexis as a prerequisite for reading
- teachers focusing on word families and lexical sets
- writers making clear connections between related parts of their text using lexical repetition

Discourse-organising vocabulary

McCarthy (1991: 78–84) and McCarthy and Carter (1994: 105) drawing on the work of Winter (1977, 1978) and Hoey (1983) show how certain words are strongly associated with certain patterns of information. These patterns involve stages in a piece of discourse such as: (1) stating the problem, (2) suggesting solutions, (3) evaluation of the solutions. Vocabulary like *problem*, *crisis*, *dilemma*, *issue* is associated with the problem stage, while words like *address* (vb), *justifiable*, *effective*, *manage*, and idioms are associated with the evaluation stage. McCarthy (1991) gives most attention to the problem/solution/evaluation and hypothetical/real (claim/counter-claim) patterns, but as McCarthy suggests, there are numerous other patterns whose parts may be signalled by the occurrence of certain vocabulary. These include:

1. the various topic types (Johns and Davies, 1983; Nation, 1993b) such as description of physical structure and characteristics (what something is like), instruction (how to do something), state/situation (what happened) and process (what happens)
2. the various genres (Derewianka, 1990) such as narratives, arguments, instructions, information reports, recounts and explanations
3. the classical rhetoric classifications such as argument, narrative, exposition and description
4. the various clause or conjunction relations (Winter, 1977 and 1978; Hoey, 1983; Halliday and Hasan, 1976; Nation 1984) such as cause and effect, contrast, exemplification, and inclusion,

especially when they relate several sentences rather than just clauses within a sentence

Some of this discourse-organising vocabulary consists of words that act a little like pronouns in that they refer back or forward in the text to another part of the text. These have been called 'anaphoric nouns' (Francis, 1994) and more generally 'discourse-organising words' (McCarthy, 1991: 75). They include words like *question*, *issue*, *assumption*, *hypothesis*, *position*, *case*, *situation* when they refer to another piece of text. Here are some examples from one text (Parkin, 1990: 101):

- 'If the supply of a good falls, its price rises. But by how much? To answer this *question*'
- 'You are trying to decide whether to advise a cut in output to shift the supply curve and raise the price of oil. To make this *decision*'
- 'Let us compare two possible (hypothetical) *scenarios* in the oil industry'

Meyer (1990) sees this discourse-organising vocabulary as becoming to some degree 'delexicalised', that is, depending more for its meaning on what it does or refers to in the text than what it carries with it. When learners meet these words in texts they need to be sensitive to their many functions which include referring to other parts of the text and signalling a stage in the discourse.

Ivanič (1991) calls these words 'carrier nouns' and typifies them as countable abstract nouns which are like pronouns in that they have a constant meaning and a variable context-dependent meaning. Ivanič (*ibid.*: 108) notes that these nouns often play an important role in exam questions ('Describe three factors that . . .') because they can be accompanied by a number but they do not give anything away about the content of the answer. Because these nouns are not topic specific, they are important candidates for a general academic vocabulary like the *Academic Word List*. Their strengths as discourse organising vocabulary are that they have a referential function and variable meaning like pronouns but, unlike pronouns, they can be modified by demonstrative pronouns, numbers, and adjectives, they can occur in various parts of a sentence and they have a significant constant meaning. Francis (1994) describes the function of these nouns as **'labelling'**: they tell the reader what to expect when they occur before their realisation and they encapsulate and classify what has been said when they occur after their realisation. They thus play an important role in the organisation of discourse.

Table 6.4. *Examples of different markers of the same clause relationship*

Vocabulary 1	Vocabulary 2	Vocabulary 3
though	nevertheless	concede
if, unless	otherwise	condition
so that	for this purpose	purpose
whereas	however	contrast

Winter (1977) notes that the relationships between clauses and sentences are largely unsignalled, but when they are signalled there are three kinds of vocabulary that do the signalling, which he refers to as Vocabulary 1, Vocabulary 2 and Vocabulary 3.

Vocabulary 1. Subordinators like *after*, *although*, *as*, *at the same time as*.
Vocabulary 2. Sentence connectors like *accordingly*, *in addition*, *all the same*, *also*.
Vocabulary 3. Lexical items like *achieve*, *affirm*, *alike*, *cause*, *compare*, *conclude*, *consequence*, *problem*.

Winter argues that this third group, although it consists of nouns, verbs and adjectives, has many of the characteristics of closed class or function words. The words in this group make up a small and fairly closed set, are to varying degrees delexicalised, and their meaning is realised by words occurring before or after them. Winter (*ibid.*: 20) lists 108 headwords for Vocabulary 3. Of these, 92 can be seen as paraphrases of words in Vocabulary 1 and Vocabulary 2 (see Table 6.4).

These 108 headwords are not the complete list as many of them can be expressed by synonyms, but the relationships they signal make up a closed set. Because these delexicalised words can play an important signalling role in clause relationships, they play an important signalling role in discourse structure.

Marco (1998) sees 'procedural' vocabulary (lexical words which structure discourse and establish meaning relationships) as consisting of two main groups: procedural organising vocabulary and procedural defining vocabulary. Procedural organising vocabulary is involved in clause relations and the structure of schemata. Procedural defining vocabulary includes formal signals of the act of defining (*is defined as*, *means*) and signals within the act of defining, namely category words and descriptions of attributes which relate to the parts of

the classical definition pattern. Teachers need to be sensitive to the discourse functions of these discourse organising words and draw attention to them in intensive reading.

Words in discourse

Let us now look at a short piece of academic text to see how vocabulary occurs in the text and to pull together the points made about vocabulary so far in this chapter. The text is taken from *Macroeconomics* (Parkin, 1990). Here is the text with the words marked in bold which occur with a very high-frequency in this book but with a much lower frequency in other texts. This is the book used in the Sutarsyah, Nation and Kennedy (1994) study whose results are shown in Table 6.2 (p. 202).

Chapter 5 Elasticity
OPEC's Dilemma
If the **supply** of a **good** falls, its **price** rises. But by how much? To answer this question, **you** will have to don a flowing caftan: **You** have just been named chief economic strategist for OPEC – the Organization of Petroleum Exporting Countries. **You** want to bring more money into OPEC. Would **you** restrict the **supply** of oil to raise **prices**? Or would **you produce** more oil? **You** know that a higher **price** will bring in more dollars per barrel, but lower **production** means that fewer barrels will be sold. Will the **price** rise high enough to offset the smaller **quantity** that OPEC will sell? As OPEC's economic strategist, **you** need to know about the **demand** for oil in great detail. For example, as the world **economy** grows, how will that growth translate into an increasing **demand** for oil? What about substitutes for oil? Will we discover inexpensive methods to convert coal and tar sands into usable fuel? Will nuclear energy become safe and cheap enough to compete with oil?

In this chapter, **you** will learn how to tackle questions such as the ones just posed. **You** will learn how we can measure in a precise way the responsiveness of the **quantities bought** and sold to changes in **prices** and other influences on **buyers** or sellers.

Price Elasticity of **Demand**
Let us begin by looking a bit more closely at **your** task as OPEC's economic strategist. **You** are trying to decide whether to advise a cut in **output** to shift the **supply curve** and raise the **price** of oil. To make this decision, **you** need to know how the **quantity** of oil **demanded** responds to a change in **price**. **You** also need some way to measure that response.
Two Possible Scenarios
To understand the importance of the responsiveness of the **quantity** of oil **demanded** to a change in its **price**, let us compare two possible (hypothetical)

https://doi.org/10.1017/CBO9781139524759.008 Published online by Cambridge University Press

scenarios in the oil industry, shown in Fig.5.1. In the two parts of the figure, the **supply curves** are identical, but the **demand curves** differ.

Focus first on the **supply curve** labelled *So* in each part of the figure. This **curve** represents the initial **supply**. Notice that *So* cuts the **demand curve** in both cases, at a **price** of $10 a barrel and a **quantity traded** of 40 **million** barrels a day.

Now suppose that **you** contemplate a cut in **supply** that shifts the **supply curve** from *So* to *S1*. In part (a), the new **supply curve** *S1* cuts the **demand curve** *Da* at a **price** of $30 a barrel and a **quantity traded** of 23 **million** barrels a day. In part (b), the same shift in the **supply curve** results in the new **supply curve** cutting the **demand curve** *Db* at a **price** of $15 a barrel and a **quantity traded** of 15 **million** barrels a day. (Parkin, *ibid.*: 102–103)

Firstly, note the very frequent occurrence of the marked words. There are 34 words in the book in this group (see Table 6.2) and these 34 word families account for 10% of the running words in the text. There is on average one in every line of Parkin's book.

Secondly, notice the part of the text where these words do not occur. This is where a range of examples is being presented. Examples help bring a message alive but they also impose a vocabulary load because they move outside the normal vocabulary of the text.

Thirdly, notice that one of these highly frequent words is a function word – *you*. This is the only function word in the list of 34 words. Notice that it is frequent because of the way the writer treats the relationship between himself and the reader. He directly addresses the reader, involving the reader in the text and directing the reader in a polite version of the imperative ('you will have to learn . . .').

Let us now look at the lexical chains in the text. Notice the relationship between the lexical chains and the very high-frequency vocabulary. Here are two related chains.

1. price rises
   raise prices
   a higher price
   will the price rise high enough
   raise the price
   change in price

2. supply of a good falls
   restrict the supply
   lower production
   smaller quantity that OPEC will sell
   a cut in output
   shift in the supply curve

Notice that by the end of the text the discussion has moved from actions (restrict the supply) to abstract representation (shift in the supply curve); this is the point of the whole book: learning the principles behind economic activity.

Notice the variety of forms and uses to convey the same idea.

Because this text is introductory, the writer is very aware of and friendly to the reader (*you*). As a result the vocabulary used is largely accessible and not highly technical. There are few words that have forms unique to the field of economics (*elasticity* perhaps). Most are slightly narrowed uses of common words: *price*, *supply*, *demand*, *margin*. The whole book has a vocabulary of only 5,438 word families indicating once more its role as an accessible, introductory text.

The example text contains some discourse-organising words – *question*, *decision*, *scenarios* – that could be related to the problem/solution/evaluation pattern. There are also several anaphoric nouns that are very clearly formally related to what they refer to: *production*, *growth*, *response*. This clear relationship reflects once more the writer's intention to keep the text accessible.

This chapter has tried to show that vocabulary does more than convey particular meanings; it plays an important part in making a text cohesive and coherent so that the text conveys a range of different messages to the reader or listener.