A Corpus-based Investigation of the Lexis of the Postgraduate Engineering Textbooks with Reference to the Needs of Southeast Asian Students

Jiawei Liu

Submitted for the degree of Ph.D
Centre for English Language Teacher Education
University of Warwick

December, 1998

To my husband, Tom

CONTENTS

ACKNOWLEDGEMENTS

I would like, first and foremost, to thank my supervisor, Dr. Hilary Nesi for her continued support. During my slow and often interrupted writing of this thesis, she has never failed to give me guidance, especially in times of uncertainty. For this, I am deeply indebted to her.

I also wish to express my gratitude to Julia Khan, the Joint Director of Centre for English Language Teacher Education, who has given me time to discuss my earlier studies at the beginning of the research. I am also grateful to Meriel Bloor for her interest and sound advice that she has provided. Dr. Kia Karavas-Doukas and Dr. Pauline Rea-Dickins kindly helped with one of my studies.

My appreciation also goes to Dr. Amanda Dowd, the co-ordinator of the Msc.Engineering programme at the Engineering Department, who helped with collecting the data of the Engineering textbooks, and to the staff at the Engineering Department who kindly provided me with the feedback that I needed for my study results. I also acknowledge the assistance of the Data Entry Unit from the Computing Services who prepared the data for me.

Last but not the least, I wish to thank Anne, Pat, Louise and Lorraine for their constant help during my study at CELTE, and after I left there. They have given me various kinds of assistance, for which I am very grateful.

All the many errors which undoubtedly remain are entirely my own responsibility.

ABSTRACT

This research is mainly concerned with establishing the vocabulary learning needs and goals of the Engineering students from Southeast Asia studying at British universities. The research was motivated by the needs to enhance the reading skills of these students. Subtechnical and technical vocabulary are the focus of this invesgitation.

The research is based on data derived from a 536,051 word corpus of text from recommended Engineering textbooks. The relative frequency and range of lexis within the corpus was found to be a good criterion for identifying subtechnical and technical vocabulary. The students proved to have a better receptive knowledge of subtechnical than technical vocabulary. The research suggests that there is a need for collaborative work between ESP teachers and subject teachers to help the students with technical vocabulary.

The thesis is divided into nine chapters. Chapter One is a review of literature to the research. It clarifies various definitions and concepts, describes the research approach, and provides a framework of the thesis. Chapter Two investigates my subjects' overall vocabulary knowledge. Chapter Three introduces some preliminary data that contrasts the received opinions in ESP regarding technical and subtechnical vocabulary. For further investigation of these two types of vocabulary, Chapter Four describes the data on which empirical studies are based. Chapter Five analyses the data. Chapter Six presents the empirical studies and concludes that students' receptive knowledge of subtechnical vocabulary is better than their technical vocabulary. Chapter Seven examines the reasons why techncial vocabulary was problematic. Chapter Eight summarises the research findings and proposes pedagogical implications in the teaching of subtechnical and technical vocabulary to the specified group of learners. And Chapter Nine draws conclusions, discusses limitations of the research and makes recommendations for future research.

Chapter One

Definitions and Framework

1.1 Introduction

This research is a corpora-based study of the lexis postgraduate Engineering students encounter while studying at British universities. The aim of the research is to provide a possible basis for organizing the teaching of Engineering English vocabulary to these students.

Vocabulary teaching presents considerable problems to those concerned with English courses for students of Engineering at this level, particularly when a common Engineering English programme for students of various disciplines of Engineering is required. None of the available word lists seem to be appropriate for this purpose. For example, West's **General Service List of English Words** (1953) (see 1.3 for a full account of this list) is mainly intended for 'general purpose' English. Moreover, most frequency lists contain only high frequency words. In this study, an attempt is made to draw up more realistic and useful frequency lists based on an investigation of the vocabulary occurring in relevant Engineering texts, using a frequency and range criterion. These frequency lists serve as examples of how frequency and range criteria can be used to draw similar lists with larger corpora.

The purpose of this chapter is to clarify some of the basic definitions used in the research, and to present the research purposes and methodology, so that a theoretical framework of the thesis can be established.

The first step in any study of lexis is to define what a word is. Section 1.2 of this chapter therefore describes various ways of defining a word as well as different aspects of word knowledge. Section 1.3 reviews several standard word lists used in the study of foreign language learner's general vocabulary knowledge. Section 1.4 discusses vocabulary in ESP, which is the key area of investigation in this research. Section 1.5 describes my research subjects. Section 1.6 presents the framework of the thesis. Finally section 1.7 discusses the research questions and research methodologies.

1.2 Definitions and concepts

This section discusses some definitions and issues that will be frequently referred to in this research, such as 'word', 'word knowledge', and 'vocabulary learning'. Iit also draws a distinction between 'receptive' and 'productive' vocabulary knowledge.

1.2.1 Word

In this thesis a word is defined in the manner suggested by Carroll, Davies and Richman (1971: xiii) 'as a string of graphic characters bounded left and right by space'. This is often referred to as the orthographic definition of a word. The authors acknowledge the limitations of the definition: 'This definition is insensitive to differences in meaning and function; it treats all words spelled the same way as the same word' (ibid, vi). The LOB corpus uses a similar definition, 'A graphic word is defined as a sequence of alphanumeric characters surrounded by spaces' (Hofland and Johansson, 1982: 7). The authors also acknowledge its deficiencies, 'Our rank list shares the weaknesses of other similar listings. As it is based on graphic words, it does not specify ... the frequency of the various senses of lexical items' (ibid, 20). According to this definition the separate occurrences of verbs such as *argument* (as in **Information Technology**) and *argument* (of an opinion) are treated as one occurrence of the same item. This is far from satisfactory, as semantic information is essential for a study to be pedagogically useful. A second issue that arises is that when the orthographic definition is used related derivational and inflectional forms such *as varying, varies, vary, variation, varied, variable, various* and *variety* are listed separately in a frequency list. Also, multi-word items which form a single semantic unit are broken up and listed separately.

In spite of all these disadvantages, however, the orthographic definition remains a good and practical working definition, because it is fairly straightforward and easy to apply consistently. Sinclair (1991) suggests that 'The best way is ... to work for as long as possible with data in a form close to its physical occurrence, and only to introduce abstractions where these are unavoidable', otherwise, '... special routines will be required to retrieve the original words, and ... we are distorting the text.' The orthographic definition is used in many studies such as Campion and Elley (1971), Hofland and Johansson (1982) and Yang (1986).

The terms 'word form', 'graphic word', 'token', 'type' and 'lemma' are also used in this research. For example, my small corpus consisted of 208,473 word forms. When I measure the length of a text I count the number of word forms, many of which occur more than once. Some researchers prefer to use the term 'graphic word' or 'token'. 'Word form', 'graphic word' and 'token' are interchangeable. I count every instance of the same word form in a text as belonging to one 'type'. For example, of the 208,473 word forms in my small corpus, there were only 10,348 different types. Sinclair uses the term 'lemma' to represent all the inflected forms of the same word type (1985: 84). For example, the lemma carry represents *carry, carries, carrying,* and *carried.* The lemma is an abstract or theoretical construct which is realized by word forms. However, the question remains whether the lemma should be used to represent word forms in frequency studies, especially if no uninflected forms appear in the data. Sinclair (1985: 84) points out, 'Lemmatization looks fairly straightforward, but is actually a matter of subjective judgment by the researcher. There are thousands of decisions to be made'. Farrell (1990) used the lemmatization policy in his study of the English of Electronics, and noted that in some cases inconsistent decisions need to be made.

To summarise, 'word' in this thesis refers to an orthographic word rather than a 'lemma'. It is interchangeable therefore, with 'word form' and 'token'. A word represents every instance of the same word form within a text.

1.2.2 Word knowledge

Word knowledge may be taken to imply the ability to recognize or use a word. The various components of word knowledge are the knowledge of: **form, position, function** and **meaning** (Nation, 1990). Both *spoken* and *written* **forms** that need to be considered. Knowing a word involves being able to recognize it when it is heard (what does it sound like?) or when it is seen (what does it look like?). It also involves being able to pronounce it and write it correctly. **Position** includes the knowledge of *grammar* and *collocation*: the grammatical patterns of a word, the words or types of words expected before or after this word, and the words or types of words that must be used with this word. Collocations are words that often occur together. Knowing a word involves having some expectation of the words that it will collocate with. **Function** refers to the knowledge of *frequency* and *appropriateness* of a word. Knowing a word means being able to decide whether this word is frequently occurring or rare. It also involves knowing whether it is old-fashioned, whether it is limited to regional usage, whether it is only used in spoken or written English and whether it is colloquial, formal, neutral or impolite. Lastly the meaning of a word is the concept it expresses and its associations. Knowing the meaning of a word involves knowing the meaning of other words which express the same or similar concepts. It also includes being able to decide which meaning is most suitable for the context that it occurs in.

There are two types of word knowledge: 'receptive' knowledge or 'passive' knowledge and 'productive' or 'active' knowledge (Melka Teichroew, 1982). Receptive knowledge of vocabulary is the ability to recognize the word when it is seen. This includes being able to distinguish it from words with a similar form and being to able to judge if the word form sounds right or looks right. Receptive knowledge of a word also involves having an expectation of what grammatical patterns the word will occur in. For example, knowing the word *suggest* involves the expectation that this word will be followed by an object and that this object can be in the form of a clause . Productive knowledge of a word, according to Nation

(1990: 32), 'includes receptive knowledge and extends it'. It refers to the ability to actively produce the word in speech or writing. It includes, for example, the knowledge of how to pronounce a word, how to write and spell it, and how to use it in grammatical patterns.

Faerch, Haastrup and Phillipson (1984) place receptive and productive vocabulary knowledge on a continuum and define it as a 'continuum between ability to make sense of a word and ability to activate the word automatically for productive purposes'. This definition means that for each particular learner there may be a varying number of words which are found at each specific point on the continuum. Thus, at one end of the continuum there are words that the learner has not come across before either in speech or in writing, but which he can understand when encountered. This kind of vocabulary is referred to as 'potential vocabulary' (Berman, Buchbinder, and Beznedeznych 1968, cited in Palmberg, 1990), which means learners can make sense of this vocabulary by making lexical inferences using interlingual, intralingual or extralingual cues to meanings (Carton, 1971). Towards the other end of the continuum there are words that the learner has already been confronted with during the learning process, and that he can either only understand (passive real vocabulary) or both understand and use (active real vocabulary) (Palmberg, 1990).

In an analysis of the components of word knowledge, Cronbach (1942) identifies four stages involved in understanding a word. These are generalization (being able to define the word); application (selecting an appropriate use of the word); breadth of meaning (recalling the different meanings of the word); and precision of meaning (applying the word correctly to all possible situations). These steps are related to increasing knowledge of a word, which takes into account such aspects as the difference between being able to define it, selecting an appropriate use, and being able to use it.

In the analysis of receptive knowledge of vocabulary, Dale (1965) proposes a four-stage continuum:

Stage 1	I never saw it before.
Stage 2	I've heard of it, but I don't know what it means.
Stage 3	I recognize it in context - it has something to do with ...
Stage 4	I know it.

The analysis of word knowledge, especially the theory of a 'continuum', sheds light on the research into vocabulary learning. In a way it helps us to understand learner behaviour and the structure of learners' mental lexicons.

The mental lexicon of the learner is the potentially complex network of associational links between individual words stored in learner's memory. Some of the links, for example those that are based on semantic relationships, topical and situational links, are native-like. Other links, for example those ones based on formal similarities between foreign language words and words in the mother tongue, are typical of foreign language learners. There are also links that are based on experiences, there are even links that are created (imaginatively) for learning purposes. In order to activate the mental lexicon, learners should first and foremost be aware of the potential power of these associational links, and must practise the ability to form the associational links into networks of high valency - networks where most of the incorporated words are linked up to as many other words as possible (Meara, 1984). According to Meara (1984: 231), the learners' mental lexicon comes into two major parts, with a phonological/orthographical code that identifies the basic form of a word and a semantic entry that specifies its meaning. Forster (in Gairns and Redman 1986: 88) considers the mental lexicon as a 'large master file' that consists of a number of 'peripheral access files' where information about the spelling, phonology, syntax and semantics of each incorporated word is stored.

In a theoretical model Bialystok and Sharwood Smith (1985) introduced two concepts that are central to all vocabulary learning and use: lexical knowledge and lexical control. Lexical knowledge is defined as 'the way or form in which words are represented or stored in learners' mental lexicons'. Lexical control, on the other hand, is considered as 'the processing system used by learners for controlling that knowledge during actual performance'. As far as lexical knowledge and lexical control are concerned, foreign language learners and native speakers are different in that the mental lexicons of foreign language learners do not enable them to use English with the fluency that characterises the words they use in the first language. The reason is that, as Meara (1984: 231) suggested, foreign words are typically stored in learners' mental lexicons in incomplete ways with, for example, half of the entry blank or with parts of the word missing. Gairns and Redman (1986: 88-89) report findings that word frequency and recency of use are two of the most important factors that affect the storage and retrieval of foreign language words in a learner's mental lexicon. The most frequently used words are easier to retrieve. Similarly, recently used words are often more readily accessible than less recently used ones. Apart from frequency and recency, other factors that affect the difficulty of a word could be its pronounceability, its part of speech, its similarity to known words, and its being learned receptively and productively. It is believed (Nation, 1982) that pronounceability rather than syllable length of a word has a strong effect on difficulty, and it is also claimed when a word is easier to pronounce, it is easier to learn.

1.2.3 Vocabulary learning

There are two types of vocabulary learning: direct and indirect vocabulary learning. When a conscious effort is made to learn vocabulary, direct vocabulary learning is taking place. Examples of direct vocabulary learning are: learning lists of word forms and their meanings, doing vocabulary learning exercises, or studying affixes and roots. In indirect vocabulary learning, new words are learned while reading or listening, usually as a result of information provided by the context.

Most vocabulary learning activities are indirect. This is probably because the amount of vocabulary involved and the complexity of learning make it impossible for a language course to teach all the vocabulary with ease within a limited period of time. Saragi, Nation and Meister (1978) found that after reading a novel, learners could recognize the meanings of 76% of the ninety new words tested. They concluded large quantities of vocabulary can be learned indirectly. Many language teachers believe that words should be presented in contexts. For example, Judd (1978: 73) argued that '... vocabulary should be taught in context'. He justified his argument by continuing, 'Words taught in isolation are generally not retained. In addition, in order to grasp the full meaning of a word or phrase, students must be aware of the linguistic environment in which the word or phrase appears'. The first part of his justification is not true, because it is known that words taught in isolation can be retained. The second part of his justification is somewhat vague. Although context is important in helping to understand the meaning of a word, it is not clear how precisely the 'linguistic environment' helps the actual learning of a word. In most cases a linguistic environment only presents one particular learning situation. Compared to the translation or dictionary definition of the word, the meaning that the linguistic environment provides is very restricted and limited. Meeting the word in different environments rather than one environment helps learners to grasp the meanings of the word. In this sense, it is the number of contexts or frequency of encounters with the words in context, rather than context itself, that helps the learning of a word. Learning from a word list can be equally or more beneficial and efficient. Meeting a word in context can help learners to learn the part of speech of a word more easily, however, this information is typically available in word lists, too. In the analysis of vocabulary learning methods, Nation (1983) reports that learning words in context would not help the learning of a word, as the content of text is remembered better than its formal element. He concludes that except by providing different environments, it is difficult to see how contexts help the learning of new words. Also, the results of these studies depend on what aspects of word knowledge were tested.

Earlier studies by Morgan and Bailey (1943), Morgan and Foltz (1944), Lado, Baldwin and Lobo (1967) and Gershman (1970) found no significant difference between learning words in isolation and learning words in context. Seibert (1930) found that the method of learning words in isolation consistently gave better results.

In spite of these findings, the issue of whether vocabulary should be learned in context or isolation cannot be settled before the following points are taken into consideration. First of all, the term 'context' needs to be defined. Context for Crothers and Suppes (1967) for example, consist of neutral (i.e. non-defining) sentences. For Seibert (1930) context is a defining sentence, with translations following the words to be learned. For Morgan and Bailey (1943), Morgan and Foltz (1944), Holley and King (1971) and Holley (1973) context means the presence of a story, with the words presented in a passage. Lado, Baldwin and Lobo (1967) define context as anything added to the word in question, such as putting the word in a sentence, providing cognates to accompany the meaning, and placing several words in the context of a story. Therefore, the term context is open to interpretation until the ways in which various types of context help learning can be decided. Secondly, we must investigate what learners actually do when they learn vocabulary in isolation, and when they learn it in context. Gershman (1970: 3690b) drew attention to similarities between the two processes:

> ... despite differences in the overt responses required for completion of the tasks, there were no basic differences in what the learners did covertly, which was to establish the bond between the target word and its English equivalent, i.e. to establish a word-meaning association.
>
> For example, presenting words totally in context did not prevent the learner from translating the target word through his own language.

Thirdly, the length of time allowed for learning a word needs to be considered. The decision of how much time spent on the new words and how much time on the context is important when comparing learning words in isolation (such as a word list) to learning in context. Morgan and Bailey (1943: 56) comment on the effect of vocabulary learning in context:

> Context may bring less frequent use of dictionary, reduce the time spent and, as a total result of these two factors, reduce the amount of learning as measured by vocabulary recall tests. Thus, if vocabulary learning is the aim, the use of contextualized material might be less efficient than the use of word lists.

This section discusses two approaches to vocabulary learning: direct and indirect vocabulary learning. Indirect vocabulary learning can be encouraged by exposure to a large amount of reading and listening materials. It would be ideal to have a well organized extensive programme of reading in a language course. However, it is difficult to give the right advice and instruction on how to make use of context; no method is available to measure the particular knowledge that context might give, therefore there is no way to tell the extent to which context will aid learning. As a result, direct vocabulary learning plays an important part in vocabulary development. Such learning is not a substitute of indirect learning, it is a complementary approach which speeds vocabulary development.

1.3 Word lists

A word list is a list of the basic and most important words in a language, it is generally intended for use as a basis for language teaching or for the preparation of teaching materials. Word lists are usually based on frequency counts. A frequency count is compiled by making a list of the words in a particular text or group of texts and counting how often and where they occur. Two of the most widely used frequency counts are: **The Teacher's Word Book of 30,000 Words** by Thorndike and Lorge (1944), and **A General Service List of English Words** by West (1953).

The Teacher's Word Book of 30,000 Words lists the most frequently occurring items in a corpus of four and a half million words. For each of these items it includes a general frequency count per million words, and frequency counts of the words in the following sources:

1. readers, textbooks, the Bible, and the English classics,
2. recent and popular magazines,
3. books recommended for boys and girls in grades three through eight,
4. miscellaneous reading matter for juveniles and adults (excluding school readers and textbooks).

For example, the entry for the word *abandon* is:

	G	T	L	J	S
abandon	38	119	150	130	285

In the first column after the word *abandon* is a number stating the occurrences per million words. 1 = at least one occurrence per million and not so many as two per million; 2 = at least two per million and not so many as three per million; and similarly up to forty nine. In the case of the word *abandon*, it has thirty eight occurrences per million. In the other four columns are numbers giving the number of occurrences in approximately 4.5 million words of (T) the Thorndike general count of 1931, (L) the Lorge magazine count, (J) the Thorndike count of 120 juvenile books, and (S) the Lorge-Thorndike semantic count. The T counts emphasize frequency in readers, textbooks, the Bible, and the English classics. The

L counts include only recent and popular magazines. The J counts include only books recommended for grades three to eight. The S counts use a miscellany of juvenile and adult reading, but omit school readers and textbooks. The words marked 'AA' belong in approximately the first thousand for frequency; words marked 'A' belong in approximately the second thousand; words marked 'forty nine' to 'thirty' are approximately the third thousand; words marked 'twenty nine' to 'nineteen' belong in approximately the fourth thousand; and words marked 'eighteen' to 'fourteen' belong in approximately the fifth thousand. The word *abandon* therefore belongs in the third thousand word band.

Regular plurals, comparatives and superlatives and verb forms in *-s, -d, -ed,* and *-ing,* past participles formed by adding *-n,* adverbs in *-ly* that occur less than once in a million words and equally rare adjectives are counted under the same word form. So are words which have special meanings when capitalized and words usually capitalized which have special meanings when not capitalized. Some adjectives in *-ing* or *-ed* and some nouns in *-ing* are entered separately, the word form being followed by *adj.* or *n.* For example, there is an entry:

abandoned (adj)	3	11	14	12	27

In general, participial adjectives and verbal substantives are included under the main word form. In the above examples, the first number is the most important, indicating the frequency of occurrence in the general count, while the following numbers indicate the frequencies of occurrence in the four different areas. This word count enables us to know the frequency of occurrence of each word. It gives specific information about words that occur in different types of standard English reading matter in the 1940s. It also provides a useful guideline in the treatment of words for foreign language learning and research purposes. Those words from 'AA' down to '14' are an important source for drawing word lists for testing and other purposes.

Roberts (1965: 10) comments, 'With the publication of this compilation, word counts of reading vocabulary had gone about as far as they were ever apt to go ... the 1944 count is regarded by many as the epitome of word counts.' Nation (1990) highly recommends this word count too, and regards it as 'the most widely known' (1990: 20). One of his vocabulary tests uses word samples from **The Teacher's Word Book of 30,000 Words**. The study by Campion and Elley (1971) (see section 2.2) also used this frequency count, as did studies by Zettersten (1979), Gui (1987), Meara and Jones (1987) and Meara and Buxton (1987). All these studies will be reviewed in Chapter Two.

Thorndike and Lorge indicated range as well as frequency of the word forms listed. A similar strategy was adopted by West in his **General Service List of English Words** (1953). Inspired by his perception of the needs of second language learners of English, West's work was based on a report to the Carnegie Conference in the 1930s on vocabulary selection for teaching purposes. The conference was held:

> to discuss the part played by corpus-based word lists in the teaching of English as a foreign language. The conference brought together some of the most influential English language teachers of the day ... and leading US linguists and educators. The product of their deliberations ... outlined the principle that items with a likely frequency of occurrence in texts should be first to avoid memory overload and confusion and to lighten the learning burden. ... and the Report broke new ground by highlighting the importance of very high frequency function words in text. ... The need to concentrate initial teaching on high frequency items became an article of faith for the most professional of language teachers from the 1940s ... (Kennedy 1992: 337-339)

West's main criteria for the selection of vocabulary for learning are that the frequency of each word in written English should be indicated, and information should be provided about the relative prominences of the various meanings and uses of each word form. These criteria provide useful guidance for teachers to decide which words and which meanings should be taught first. The list consists of 2,000 words with semantic and frequency information drawn from a five million word corpus. It is claimed that knowing these words gives access to 80% of the words in any written text (Carter, 1987: 163). This stimulates motivation, as learners realize that these words are extremely useful in a wide range of contexts.

The following is a typical example of an entry for the word *head* in the **General Service List**.

HEAD	2216e		
head, n.	(1)	*(part of body)*	
		A hat on my head	
		Head and heart (= *intelligence and feeling*) (5%)	
		Will cost him his head (= *life*) (2%)	
		Wiser heads; count heads (= *persons*) (1%)	
	Phrase:	Over head (= *above*)	**63%**
	(2)	*(top)*	
		Head of a bone, plant, page, procession	**13%**
	(3)	*(idea of leadership)*	
		Head of the school	
		Head clerk	
		At the head of the whole business	**14%**
head-/		Head-rest, head-hunter, etc.	**1%**
	Phrases:		
		Head to foot	**0.2%**
		Head or tail	**0.3%**
		Keep, lose one's head (= *control*)	**0.4%**
	[On your head be it, 0.4%;		
	Off his head = mad, 0%]		
head, v.		Headed the list	
		All came in headed by Mr. X	
		Gold headed stick	**5%**
	[Heading for = going towards, 0.4%]		
heading, n.		Newspaper heading, etc.	**2%**

In this example, 2261e indicates the estimated number of occurrences in five million words. The detailed break-down of each individual derivative provides a considerable amount of help for teachers in selecting words and meanings to teach.

Although these two counts provide useful information for a wide range of purposes such as teaching and research, their weaknesses should not be ignored. First of all, neither of the counts was based on modern day texts. **The Teacher's Word Book of 30,000 Words** was prepared in the 1940s, and **A General Service List of English Words** was prepared in the 1930s and revised in 1953. The sense of 'essentialness' of certain words could be different now compared to the time when the two counts were compiled. For example, a modern corpus may include *computer* as a high frequency word, as opposed to words such as *crown* and *canal* that have high frequency in the above two lists. The two counts are based on the written language only, therefore their frequency and contexts in the spoken language are not included. Also, the word counts offer no proper divisions of genre. For example, the Thorndike and Lorge corpus was sampled from a variety of sources such as the Bible, the 'English classics' and popular magazines. Despite these disadvantages, standard word counts contain highly useful information about words, valuable for teaching, learning and research purposes.

1.4 Vocabulary in ESP: technical and subtechnical vocabulary

1.4.1 Introduction

This section reviews studies of subtechnical and technical vocabulary in the study of English for Engineering , in particular it presents different definitions of these two types of vocabulary, which my own definitions will be based on.

The analysis of the Engineering textbooks to be described in this thesis follows the tradition of register analysis in ESP, which took place in the 1960s and early 1970s. Based on the principle that the English of Engineering, for example, constitutes a specific register different from that of Economics, register analysis aims to identify grammatical and lexical features of these registers and use this information in developing materials. Some practitioners express little interest in the analysis of grammatical register, '... register analysis revealed that there was very little that was distinctive in the sentence grammar of Scientific English (Hutchins and Waters, 1987: 10). The analysis of lexical aspects, however, is considered to be far more relevant, as suggested by Sager, Dungworth and McDonald (1980), 'The non-specialist reader recognises the grammatical forms and structure of an SE (Special English) text as those normally used in written English. Many of the lexical items when he encounters there, however, are likely to be either completely new to him or at least used with a meaning different from that with which he has previously been familiar' (1980: 230). This is also expressed by Johansson (1975), '... the crucial criteria of any given register are to be found in its grammar and its lexis. Probably lexical features are the most obvious ... Purely grammatical distinctions between the different registers are less striking ...'.

Another approach to language description is genre analysis, which took early register analysis much further by drawing attention to the relationship between grammar, lexis and their combined communicative events. For some writers, 'genre' is the same as 'text type' or register. As with register analysis, genre analysis looks at the operation of language within a complete text, seeing the text as a system of features and choices. Selection is made according to the text writer. Salager-Meyer et al (1989) made an analysis of medical English scholarly papers, and found that these could be divided into editorials, research papers and case reports, referred to as 'sub-genres' and 'text types'. The results suggest a systematic difference between each text type or sub-genre according to the attitude of the writer to the reader, offering pure description in the case reports, advice and suggestion in the research papers, and judgment, value and instruction in the editorials. These writers seem to indicate that editorials, research papers and case reports are sub-genres of the 'genre' of medical English, thus taking discipline or domain as the primary distinguishing factor between texts. Swales (1981) used the term 'genre' to imply much more than 'text type'. He has the following definition of 'genre': 'a more or less standardized communicative event with a goal or set of goals mutually understood by the participants in that event and occurring within a functional rather than a personal or social setting'. The standardization of the event implies some regulation by the professional community, and the 'mutual understanding' suggests some induction into that community. In his later work (1986, 1988) he introduced the concept of 'discourse community', the members of which 'share common public goals', have 'mechanisms for intercommunication between members' and have 'discoursal expectations' leading to the development and use of distinctive text types involving specialised terminology'. Thus for Swales, genre involves not only text types but also the role of the text in the community which produces it. In a discussion of what might be involved in genre analysis, Miller and Selzer (1985) suggest that the analysis of any particular text should consider the 'generic' element (for example, whether it is broadly of the report, memo or letter genre), the institutional element (for example, which company department it derives from) and the disciplinary element (for example, the topic and academic discipline). Dudley-Evans, in an introduction to genre analysis (1987: 2), suggests that ESP needs a system of linguistic analysis that demonstrates differences between texts and text types. 'Genre analysis may be used as a classificatory system, revealing the essential differences between the genre studied and other genres and also between the various sub-genres'. He is also concerned with similarities between texts, using a system of clause relations (see 1.4.4) to 'describe relations that are found in all texts' (Dudley-Evans, 1987: 5). He considers genre analysis to be prescriptive (for example, able to provide pedagogically useful information), whereas register analysis is descriptive.

Although genre analysis is an important development in the description of language, other approaches may still be of use, for example that of discourse analysis. For the study of vocabulary, corpora-based statistical or frequency analysis is commonly used.

Frequency studies, usually based on corpora, have become more popular because the data from the studies enable language teachers or practitioners to identify lexical items which learners are likely to encounter frequently, and also to identify common syntactic, semantic and collocational patterns. In other words, frequency studies provide factual information about the frequency of words, word senses, structures, and collocations within specified genres. A corpus provides information about lexical use in naturally-occurring discourse rather than information based on a native speaker's intuition, which sometimes does not match the actual patterns of use. This is well put by McEnery and Wilson (1996:):

> The importance of corpora in language study is closely allied to the importance more generally of empirical data. Empirical data enable the linguist to make statements which are objective and based on language as it really is rather than statements which are subjective and based upon the individual's own internalised cognitive perception of the language.

Carter (1987: 182) notes that lexical studies 'offer invaluable data for vocabulary materials development'. Fillmore (1992) puts it this way, 'The most convincing part of the case for using a corpus was that it makes it possible for linguists to get the facts right. Authenticity was the key word'.

Biber, Conrad and Reppen (1991) identify two major advantages to the use of corpora for linguistic analysis.

> 1. They provide a large empirical database of natural discourse, so that analyses are based on naturally-occurring structures and patterns of use rather than intuitions and perceptions, which often do not accurately represent actual use.
> 2. They enable analyses of a scope not feasible otherwise, allowing researchers to address issues that were previously intractable. This is particularly true of computer-based text corpora, which can be analysed using (semi-) automatic techniques. Such analyses can examine much more language data than otherwise possible, including more texts, longer texts, a wider range of variation (texts from different language varieties), a wider range of linguistic characteristics, and the systematic co-occurrence patterns among linguistic features.

Some people might consider it a weakness that a large number of relatively rare words usually appear in corpus-based frequency lists. 'Statistical studies of word distribution show that ... high percentage of words occur only once' (Richards 1974: 72). In the present research, for example, 8,871 types occurred less than five times, although this only made up 15,507 tokens, 3% of the entire corpus. Only 6,122 types were found to occur more than five times. This totalled 520,544 tokens, however, making up 97% of the entire corpus.

Although the corpora based approach has its defects, this approach is an important means of developing language materials, as Kennedy (1992) forcefully argued:

> language teachers, syllabus designers, and materials writers are willfully ignoring compelling frequency evidence already available. Whatever the imperfections of the simple equation 'most frequent' = 'most important to learn', it is difficult to deny that frequency information becoming available from corpora has an important empirical input to learning materials.

Many lexical studies of technical and subtechnical vocabulary are based on corpora. Sections 1.4.2 and 1.4.4 review some of these studies.

1.4.2 Technical vocabulary and technicality

Robinson (1991: 28) places vocabulary in three categories: specialist vocabulary (which I call *technical vocabulary*), subtechnical (sometimes called *semi-technical* or *general scientific/technological*) vocabulary, and general or non-academic vocabulary. Technical vocabulary is fundamental to specialized texts because it is through the use of such vocabulary 'that scientific knowledge can be differentiated from everyday language' (Pueyo and Val, 1996: 253).

In order to understand why scientific writing is difficult, Halliday and Martin (1993, 14-15) argue that grammar and words should not be separated from each other, 'it is the total effect of the wording - words and structures - that the reader is responding to, and technical terms are part of the overall effect'. Although grammar is important in the process of forming a new term, this research is more concerned with the lexis that helps the formation of technical terminology. I agree with Halliday and Martin that technical terms cannot be defined in isolation; each technical term has to be understood as part of a larger framework, and each one is defined by reference to all the other terms. Pueyo and Val (1996: 254) consider the following to be the function of technical terms:

> Technical terms are used in technical discourse to compact information and to change the nature of the more common sense meanings. Technical terminology is a linguistic resource that science has developed in order to reinterpret and re-classify the world in a different way from normal practice in natural spoken language, and to put into words very abstract concepts, processes and events.

The process of creating new terminology is known as the technicalizing process, or technicality. For example, some terms can be formed by adding the Greek word 'poly' in order to name new substances, such *as polyamide, polycarbonate, polyester* and *polymer*. Technicality is considered to have a field creating function. Field refers to the way the experimental world is dividing up into institutional areas of activity, such as Mechanics, Electricity, Robotics, Linguistics and so on. All fields have the capacity to name the entities that concern them. Thus, each field develops its own vocabulary, and it is easy to recognize the field by simply looking at the lexis. Technicality is the 'resource a discipline uses to name and order its emic (i.e. meaningful and relevant) phenomena in a way distinctive to that field. Through technicality, a discipline establishes the inventory of what it can talk about, and the terms in which it can talk about them' (Martin, 1993).

Technical terms are considered important by some ESP practitioners (Martin, 1976; Swales, 1976; Inman, 1978; Loots, 1986; Baker, 1988; King, 1989). Hoffman (1986:49), for example, believes that 'in the teaching of LSP attention is sharply focused on the acquisition of special vocabularies'. Similarly, Yang (1986: 93) claims that 'scientific/technical terms ... should be part of the teaching of EST'. However, some ESP practitioners disagree and suggest that technical vocabulary does not deserve such attention in the ESP classroom. Higgins (1985: 32), for example, states firmly that 'It is not the job of the English teacher to teach technical vocabulary'. It is commonly believed among ESP practitioners that 'what can cause a problem for the non-native user of English is not specialist ... vocabulary' (Robinson, 1991: 28).

Despite the disagreement as to whether or not technical vocabulary should be taught in the ESP classroom, most writers and researchers agree that the same technical vocabulary should not be taught to students of different subject specialisms. Nation (1990: 14) describes technical vocabulary as 'words related closely to the topic of the text'. As for the frequency and range of technical words, he commented that within a certain topic or subject area these technical words 'may occur several times, but they are unlikely to occur in texts outside that subject area.'

Yang (1986) used more vigorous tests to identify technical vocabulary, which he defines as words that have 'a peak frequency of occurrence in one or several fields, but never occur, or have very low frequency of occurrence, in other fields' (Yang, 1986: 98). According to Yang, 'Where there is a PEAK ..., there the chances for that word being a term in that specific field is fairly high.' For example, the word *acid* has a peak frequency in Chemistry, therefore it is considered a term in that field. Yang's definition can be used as a tool to distinguish texts with different subject matter. For a full account of his study, see section 1.4.4.3.

Another corpus based study on subtechnical and technical vocabulary was by Farrell (1990). Farrell identified technical words by their high frequency and low range and he noted that technical words 'have a relatively significant frequency' and 'a surprisingly low range' (Farrell, 1990: 30). Farrell identified within his technical word list two distinct types of technical words: highly technical words which do not belong in the general language, for example, *anode* and *cathode*, and words such as *energy, force, wave* and *signal* that occur in general language, but also have a technical sense. See section 1.4.4.3 for a full account of Farrell's study.

1.4.3 Lexical familiarization of technical vocabulary

Anderson and Freebody (1981) claim word knowledge is strongly related to reading comprehension. Faerch, Haastrup and Phillipson (1984) consider word knowledge as a 'continuum between ability to make sense of a word and ability to activate the word automatically for productive purposes'. This means learners 'know' certain words to different degrees: there are 'potential', 'receptive' and 'productive' groups of vocabulary (see section 1.2.2).

Potential vocabulary refers to new words that learners have not come across before, but which they can understand when encountered. The understanding of the new words is usually facilitated by lexical inferences such as intralingual cues. According to the theory of learners' mental lexicons (section 1.2.2.), the intralingual cues could automatically activate the potential complex networks of associational links between individual words stored in learners' memory or learning mechanism, these are networks in the mental lexicon where most of the incorporated words are linked up to as many other words as possible.

Technical vocabulary that is new to learners could belong to the category of 'potential vocabulary'. If provided with intralingual cues, learners could use these cues to get to the meaning of the vocabulary. Technical vocabulary is 'highly subject matter specific' (Yang, 1986: 93). A technical term is related to other terms concerned with the same phenomenon, or related phenomenon, or to terms concerned with hypotheses or technical concepts. In a group of related terms, any one term cannot be completely defined until all the other terms are defined. Hence a term is only part of the meaning of a group of terms. For example, the concept of <u>flow of electric current</u> is associated with a group of terms such as *resistor, resistance, conductor, conductivity* and *superconductivity*. The meaning of any one of these terms involves the meaning of all the other terms in the group. This shows the complexity or the technicality of technical vocabulary, and the importance of explaining, defining or using any means to facilitate the understanding of technical vocabulary from the textbook writers' perspective. Some researchers (for example, Bramki and Williams, 1984) have examined how technical vocabulary is presented to readers in texts. The following subsection describes the ways in which technical vocabulary is explained within subject specialist reading material, this is called lexical familiarization.

Lexical familiarization is typically found in textbooks, where the writer's relation to the reader is that of specialist to non-specialist. Williams (1980) defines lexical familiarization as the 'author's intention to familiarize the newly-introduced lexical items to the target reader, by verbal, illustrative or numerical devices, or any combination of these means.' Lexical familiarization occurs when it is the author's intention to help the reader with difficult vocabulary. It in effect follows Huckin's suggestion (1981): 'Those who write to non-specialists must be sure to explicate the most important concepts in their texts, by using examples, definitions, analogies or other forms of illustrations; in other words, use the familiar concepts to explain unfamiliar ones.' Therefore, lexical familiarization is an intentional contextual aid, which is explicitly provided by the author when writing for a specific readership (Bramki and Williams, 1984: 170). It is important to emphasize that it is the 'author's intention' to provide readers with sufficient familiarity so that the readers can continue reading. To illustrate this point, Williams (1985: 124) hypothesises that the author in effect pauses and thinks: 'My reader probably does not know the meaning of the term I have just used. I must explain it in some way without disrupting the flow of my writing.' He gives an example to illustrate this point:

> A large percentage of the human race still lives in very small self-sufficient peasant communities. These people experience great poverty, but they can provide, on an individual basis, for their own survival. They have a degree of *economic independence*.
> (from Williams, 1985: 124)

In this example, the author is concerned that his readers might not be familiar with the specialist term *economic independence*. He therefore provides sufficient familiarization by incorporating the phrases *self-sufficient* and *provide ... for their own survival.*

Lexical familiarization is different from lexicographic definition in that the latter provides the reader with a full, precise and usually decontexualised definition of the lexical item concerned. Therefore the writer's aim is to provide an accurate, all-embracing explanation. When using lexical familiarization, the textbook writer does not intend to provide the reader with a 100% accurate explanation, instead, his or her concern is only to give enough familiarization to enable the readers to continue reading with understanding.

Bramki and Williams (1984) conclude that lexical familiarization is less complete than definition, and it is more reader-oriented as the author's aim is to help the reader acquire enough conceptual understanding of the new term to proceed. Brown (1980) put this issue more lucidly, '... even if he (the reader) does not get the full meaning, he gets enough to proceed for a while.'

Bramki and Williams (1984) found six categories of lexical familiarization using a corpus of Chapters One to Four of **Introductory Economics** (Stanlake, 1976). These six categories are 'exemplification', 'explanation', 'definition', 'stipulation', 'synonymy' and 'illustration'.

1. Exemplification

Exemplification provides the reader with an instance (or instances) of what the newly-introduced term refers to. The following is an example given by Bramki and Williams:

> *Durable consumer goods* include such things as books, furniture, television sets, motor cars and domestic electric appliances.
>
> ... A simple arithmetical example will make the point clear. Suppose that there are two leather workers, Jones and Smith, each producing shoes and handbags. In 1 week, Jones can make either 10 pairs of shoes, or 10 handbags ... This is a very simple account of the important principles of *comparative advantage* ...

As can be seen, the first example is an exemplification expressed through a list of familiar items while in the second example the exemplification is expressed through arithmetic. Signalling of lexical familiarization by exemplification is most commonly by *such as, for example, is typified by, e.g., include, provide an example of* and *such things as.*

2. Explanation

Explanation provides the reader with a sequence of words (a phrase, a sentence, several sentences) which is equivalent or opposed in meaning to the newly-introduced term. Explanation is achieved in three different ways.

The first is by direct explanation, for example:

> *Saving* is the act of foregoing consumption.

In this example, the explanation is semantically equivalent to the term being familiarized.

The second way is by contrasting the newly-introduced term with, for example, a situation that the author believes the reader already understands, for example:

> A large percentage of the human race still lives in a very small self-sufficient peasant communities. These people experience great poverty, but they can provide, on an individual basis, for their own survival. They have a degree of *economic independence*.

And the third way is by operational description, for example:

> *Monopolies.* It is a feature of the market economy, especially in more recent times, that firms tend to increase in size and power ... This tendency towards market domination by giant firms reduces or removes the limiting role of competition and gives the large firm power to exploit the consumer by change prices as well as above costs.

According to Bramki and Williams (1984) the most common signals of this category are *i.e., means that, is known as, is taken to, means, refers to* and *concerns.* However, they also claim that familiarization by explanation is frequently unmarked. They give the following example:

> Whatever the reason the fact is that man finds himself in a situation of scarcity. He cannot have all the things he wants.

3. Definition

Swales (1971) and Lambrou (1979) agree that definition comprises three elements: the term (T), the class (or genus) to which the term belongs (C), and the differentia (the distinguishing features) (D). Flowerdew (1992) noted that definitions are judged on their completeness as well as their universality, whereas Applied Linguistic research is usually concerned with definitions in contextualised circumstances. Bramki and Williams (1984: 177) give the following examples of definition found in their Economics corpus:

> Economics is essentially a study of the ways in which man provides for his material wellbeing.
> Any activity which helps to satisfy want is defined as production.
>
> *Consumer* goods are those commodities which satisfy our wants directly.

> One major objective of science is to develop theories. These are general statements or unifying principles which describe and explain the relationship between things we observe in the world around us.

Signalling is by means of the definition formula *T is/are C which D, is/are called, is/are known as, refers to, is/are understood to be, the term X is used to describe* and *may be defined as.*

4. Stipulation

Konecni (1978: 378) defines stipulation as 'a type of definition which indicates that the term being defined has its particular meaning only in a given situation, and that it does not necessarily have the same meaning in other situations.' Bramki and Williams (1984) give an example of *land* which is stipulated in the Penguin Dictionary of Economics:

> 'Land' in *Economics*, is taken to mean not simply that part of the earth's surface not covered by water, but also the 'free gifts of nature', such as minerals, soil fertility, etc.

Thus, stipulation restricts the use of a term to a particular subject field and specifies its meaning within that field. It is signalled by means of *X uses (term) to describe Y, (term) as X sees it, X restricts the meaning of Y to, from the point of view of X* and *to X.*

5. Synonymy

According to Bramki and Williams (1984), synonymy familiarizes by providing the reader with a more familiar lexical item which has almost the same meaning. Synonymy is deliberately used by native speakers when they wish to simplify for others (Blum-Kulka and Levenston, 1983). However, Carter and McCarthy (1988) point out that true synonyms (100% interchangeability) are very rare.

The following are examples given by Bramki and Williams from the Economics corpus:

> The market system of economic organization is commonly described as a *free enterprise*, or *laissez-faire* or *capitalist system*. We shall use all these terms to stand for a market economy.
>
> Working capital consists of ... This kind of capital is sometimes called *circulating capital.*

For Bramki and Williams, synonymy is signalled by *X stands for Y*, *X is referred to as Y*, and *X or Y*.

6. Illustration

Illustration is a predominantly non-verbal method of familiarizing through tables, pictures, charts, diagrams and so on. Konecni (1978) defines illustration as 'a rhetorical technique which refers to the combination of text and visual aid used in informational writing to clarify concepts.'

All these categories of lexical familiarization are used by authors to present technical terms to readers in a way that enables them to proceed with the reading without too much disruption. The criteria used by Bramki and Williams to categorize instances of lexical familiarization are not always consistent, and sometimes they do not seem to be able to offer a clear-cut judgment. For example, they point out that the most common signalling devices of explanation are: *is known as* and *refer to,* but these devices are also used to signal for definition.

There may be differences between explanation and definition expressed through the same devices, but Bramki and Williams do not make any distinction regarding this. It is therefore difficult to distinguish familiarization expressed through explanation and definition by following the guidelines they give.

1.4.4 Subtechnical vocabulary

Subtechnical vocabulary has been the subject of considerable attention in the teaching of English for Specific Purposes, although it has not always been well-defined. Many different types of words have been classed as subtechnical by different writers. This section discusses various definitions of subtechnical vocabulary, which I have drawn on to form my own definitions in the studies reported in this thesis.

It has been claimed by many ESP practitioners that the type of vocabulary known as subtechnical vocabulary is important in the teaching of ESP because subtechnical vocabulary items occur with high frequency across disciplines, but are not often taught on General English courses. They are therefore the cause of problems for students studying in the medium of English (Trimble, 1985: 128). Although most writers seem to agree that subtechnical vocabulary is 'neither technical nor general words' (Baker, 1988), the choice of words to fit this category varies. Widdowson (1983), for example, emphasized the importance to EAP of 'procedural vocabulary': discourse organizing words which have little meaning in themselves, but which are used to structure the information carried by more topic-specific (technical) lexis. This emphasis was echoed by Baker (1988), and

King (1989) in their descriptions of subtechnical vocabulary. Higgins (1985), on the other hand, chose words according to their perceived difficulty for EAP students; and Johns and Dudley-Evans (1985) used the same approach to identify subtechnical words. A different approach again was taken by Kennedy and Bolitho (1984); they focused on vocabulary which also occurs in General English, but acquires a different meaning within an academic context. These words, like those identified by Higgins, were chosen for the EAP syllabus because they were thought to create particular difficulties for learners.

Cowan (1974) defined subtechnical vocabulary as those words that are context independent and occur with high frequency across disciplines. Cowan's definition of subtechnical vocabulary has often been used by later writers, such as Inman (1978), Trimble (1985) and Farrell (1990). This definition however, is inadequate in describing which type of word exactly fits the subtechnical category. It is unclear for example, whether it includes very common General English words. Inman based her definition of subtechnical vocabulary on Cowan's and she included very frequent words such as *large* and *high* in her lists of words with high distribution across different fields. Farrell (1990) used a similar definition to Inman's (his subtechnical words were 'context-independent words which occur with high frequency across disciplines)', but he eliminated very frequent words from his subtechnical list.

Baker's (1988) account of subtechnical vocabulary was more precise than Cowan's. She gave six different interpretations of what subtechnical vocabulary is. She argued that in technical writings, vocabulary falls into three categories: specialized items, general items and items between the two, which she called subtechnical. According to Baker, the term 'subtechnical' covered a whole range of items which were neither highly technical and specific to a certain field of knowledge nor obviously general.

Baker's six interpretations of the term 'subtechnical vocabulary' were:

1. items which express notions general to all or several specialized disciplines.
2. items which have a specialized meaning in one or more disciplines, in addition to a different meaning in general language.
3. items which are not used in general language but which have different meanings in several specialized disciplines.
4. items which are traditionally viewed as general language vocabulary but which have restricted meanings in certain specialized disciplines.
5. general language items which are used in preference to other semantically equivalent items to describe or comment on technical processes and functions.
6. items which are used in specialized texts to perform specific rhetorical functions. These are items which signal the writer's intentions or his evaluation of the material presented.

Baker's interpretations of subtechnical vocabulary summarized the various definitions given by other ESP practitioners and linguists. Higgins' (1985) and Johns and Dudley-Evans' (1985) definitions accord with the first category of Baker's subtechnical items in that the words they identified as subtechnical were common to several subject disciplines. Baker's second and fourth categories were in line with the views of Kennedy and Bolitho (1984) and Trimble (1985) who have described subtechnical vocabulary as words with two different meanings in technical and General English.

Another group of linguists has emphasized the discourse organizing role of subtechnical vocabulary. This ties in with Baker's categories 5 and 6, which focus on discourse functions. Writers such as Martin (1976), Winter (1977, 1982), Hoey (1983), Francis (1986), Widdowson (1983), Robinson (1988), and McCarthy (1991) have looked at words which function as discourse organizing items.

A third group of writers have approached subtechnical words more objectively, defining subtechnical words in terms of the frequency and range of words within a corpus of texts. No intuition need be involved in their identification. Barber (1962), Inman (1978), Yang (1986) and Farrell (1990) have adopted this method of defining subtechnical vocabulary.

These three different approaches of defining subtechnical vocabulary will be discussed in more detail.

1.4.4.1 Subtechnical vocabulary as difficult items

Several ESP teachers have considered subtechnical vocabulary as words that present particular learning difficulties for EAP learners. Higgins (1985) attempted to define subtechnical vocabulary by considering a list of 'frame' words, which according to him were definitely not technical terms but were frequently used in technical writings. He also stated that these words caused difficulties for students. Higgins' frame words formed a rather miscellaneous list. Some words such as *similar* and *identical* are procedural and can be found in Winter's (1977) list of Vocabulary Three items. Others such as *theory* can be found in Francis' (1986) lists of anaphoric nouns. There were also topic-specific words such *as boil, heat, chemicals, symptom, diagnosis, relapse, heal*, and *cure*. These words do not usually have any discourse organizing function (although in some contexts the nouns may have a summarizing or anaphoric role). The lack of consistency is very obvious. Swales (1985: 32) commented, 'It is not entirely clear what Higgins precisely means by 'frame' words; clearly some are semi-technical words of wide range; and others known to be problematic in the local situation'. Higgins' criteria for identifying subtechnical items were based on the 'hardness' of words for Thai science students.

Like Higgins, Johns and Dudley-Evans (1985) also considered subtechnical vocabulary as words that cause difficulties for students. They analyzed a sample of recorded lectures given by subject specialists in the area of Transportation and Environmental Planning and Plant Biology, and classed as subtechnical those words or terms in their corpus which were drawn from the 'common core' of English but took on a special significance in a number of different subjects. Sample words such *as affect, effect, yield, deplete* and *withstand* were claimed to cause difficulties for students.

These writers were concerned with the difficulty that subtechnical vocabulary might cause for EAP students. Kennedy and Bolitho (1984) also regarded subtechnical vocabulary as difficult vocabulary, but they used a less intuitive approach by looking at the meaning potential of words. They accepted as a starting point that subtechnical vocabulary consisted of those words which were not specific to a subject specialism but which occur regularly in scientific and technical texts. Words such as *reflection, intense, accumulate, tendency, isolate, cycle, conductor* (in terms of sense) and *dense* had the above characteristics. These words were believed to cause difficulties for learners because they took on a specialized meaning within a scientific or technical context. The specialized meaning was different from the 'general' English meaning of the word. The learner may know the 'general' meaning already and may be confused when he or she meets it in a context with a different meaning. For example, the word *cycle* has different meanings from the everyday meanings of *bicycle* in *carbon cycle* and *a cycle of electricity*. Likewise the word *conductor* has different meanings in General English and in the field of Electrical Engineering.

Trimble's (1985) definition of subtechnical vocabulary bears a strong similarity to that of Kennedy and Bolitho. Like Kennedy and Bolitho, Trimble regarded as subtechnical those General English words with special meanings in specific scientific and technical disciplines. Trimble recognized Cowan's definition of subtechnical vocabulary, but used the term 'subtechnical vocabulary' to refer to 'those words that have one or more 'general' meanings and which in technical contexts take on extended meanings'. He gave as example words: *base* (with different meanings in the fields of Botany, Chemistry, Electronics and Navigation, apart from the meaning in General English), and *dog* (with different meanings in the areas of Construction, Machining, Mechanical Engineering, Petroleum Engineering and Railroading as well as General English).

This definition raises several questions. Although Trimble took Cowan's definition as a starting point, Trimble's definition goes so far beyond Cowan's that it overrules it. Cowan's subtechnical words such as *absence, acceptance, accident, accelerate,* and *accomplish*, were a very different set of words from Trimble's subtechnical words such as *dog, fascia* and *transport*. Cowan's words were context independent words which occurred in fairly equal measure across disciplines (this can be seen from his definition of subtechnical vocabulary: words that are context independent and occur with high frequency across disciplines), while Trimble's words occurred with high frequency in only one or two fields, although they might occur occasionally in other fields. In other words, the word *dog* is not likely to occur with an even distribution in a wide range of texts. It is more likely that *dog* occurs in the specific field(s) in which *dog* is the topical word. In this sense, Trimble's subtechnical vocabulary was context dependent rather than 'context independent'.

Definitions of subtechnical vocabulary which hinge on the perceived difficulty of words are not very useful, since any decision regarding the belief of difficulty for a heterogeneous group of learners must largely be intuitive. A word might be difficult for some types of learner, but not for others. Kennedy and Bolitho (1984) and Trimble (1985), like Higgins (1985), identified difficult words as subtechnical. However, their approach was more objective because word meanings were taken into consideration. Words were identified as subtechnical if they had a General English meaning and a specific meaning within a technical field. This method of identifying subtechnical words can be replicated but may not provide us with a large set of words, as homonyms have been found to form only a small percentage of words in technical writings (Sager et al, 1980: 235).

1.4.4.2 The discourse organizing role of subtechnical vocabulary

Widdowson (1983) describes procedural words as highly 'indexical', having very little lexical content, and being impossible to separate from the contexts in which they occur. Martin (1976) describes the use of procedural words to discuss the research process, which consists of formulating, investigating, analyzing, drawing conclusions, and reporting results. Similarly, McCarthy (1991) identifies words such as *issue, problem,* and *solution* as discourse organizing words, because 'it is their job to organize and structure the argument, rather than answer for its content or field' (McCarthy, 1991: 75). Winter (1977) classifies lexis into two broad categories: closed system items and open system items. His 'Vocabulary Three' is a set of open system lexical items which are able to facilitate a reader's interpretation of a discourse. These items have 'similar semantic properties to closed system items in sentence connection' (Winter, 1977:2), and 'constitute a special vocabulary of context for the clause relations of English; they are words which can function as special signposts of what a word means in sequence with its adjoining sentences' (Winter, 1977:4; Hoey, 1983:19). One of the most important signposting functions of Vocabulary Three items is that they enable the reader to anticipate or predict the information in the discourse. In this aspect they function like the true closed system items, the subordinators and sentence connectors (Winter's Vocabulary One and Two respectively). For example, the Vocabulary Three item *contrast* is similar to the Vocabulary One subordinator *whereas* and the Vocabulary Two sentence-connector *however.* All these words are signals of the clause-relation 'comparative denial'. The characteristics of Vocabulary Three items are discussed in more detail in Chapter Five.

Winter points out that it is possible to signal the same relationship using Vocabulary One, Two or Three. Apart from Hoey (1983, 1991, 1993), Winter's work is also developed by Crombie (1985) who argues that second language teaching pedagogies and syllabuses should be based on explicit attention to the lexicalisations of intersentential semantic relationships in the texts.

Francis (1986) concentrates on a group of nouns which play a similar organizing role in discourse. She calls these nouns anaphoric nouns, which are 'primarily interactive organizational signals with the signposting function' (Francis, 1986:3). This function is one which Winter's Vocabulary Three also recognizes. The characteristics of anaphoric nouns are that they can be used metadiscursively to talk about the ongoing discourse. They can function as a pro-form 'and as such be an anaphorically cohesive device, referring metadiscursively to a stretch of a discourse preceding it in terms of how the writer chooses to label or interpret the latter for the purposes of his/her argument.' (Francis, 1986: 3, 4). Within the discourse, anaphoric nouns are presented 'as the given information in terms of which the new propositional content of the clause or sentence in which it occurs is formulated' (Francis, 1986: 4).

Hyland (1998) uses the word 'metadiscourse' to describe words or devices used to organise the discourse, engage the audience and signal the writer's attitude. He defines metadiscourse as 'those aspects of the text which explicitly refer to the organization of the discourse or the writer's stance towards either its content or the reader' (Hyland, 1998: 438). The term is used to refer to non-propositional aspects of discourse which help to organise prose as a coherent text and convey a writer's personality, credibility, reader sensitivity and relationship to the message (Crismore et al, 1993). Metadiscourse is the author's linguistic and rhetorical manifestation in the text in order to 'bracket the discourse organisation and the expressive implications of what is being said' (Schiffrin, 1980: 231).

Baker (1988) uses the term subtechnical to refer among other things to items 'which signal the writer's intentions or his evaluation of the material presented' (Baker, 1988:92). Examples she gives of such items are:

> after, disease, case, treatment, during, cases, showed, results, reported, days, test, months, weeks, concentration, cell, cells, type, developed, diagnosis, blood, tube, caused, hospital, examination, symptoms and syndrome

Some words in this list such *as patient, disease, case, treatment, cell, diagnosis, blood, cells* and *tube* are topic specific. Other methods of categorization might consider them technical vocabulary, because they convey specific concepts which relate to particular frames of reference. It is possible, however, that such words may also function within the discourse as anaphoric nouns, and may therefore carry out a discourse organizing role.

King (1989) considers nouns from five student projects, aiming to find out what lexis was common to the five projects. Examples of these nouns are:

> behaviour, case, change, conclusion, figure, force, form, hand, introduction, part, problem, project, reason, site, state, steel, system, technique, theory and time.

King reports that there is an overlap of his list with Francis' anaphoric nouns. Words such as *conclusion, factor, point, theory, aspect, fact* and *question* occur in both lists.

Widdowson (1983) and Robinson (1989) identify discourse organizing words with 'core vocabulary'. As Carter and McCarthy (1988) and McCarthy (1990) point out, core vocabulary items tend to be frequent but not difficult items. McCarthy writes, 'In any given lexical field, core words tend to be the most frequently occurring ones' (McCarthy, 1990: 49). The core word in a semantic set tends to be the most frequent item, as *fat* in the semantic group of *obese, overweight, plump, podgy* and *stout.* Core words are also used to paraphrase or define other words and thus have a procedural function. Widdowson's procedural words with 'high indexical potential' are likely to occur in a very wide range of contexts, words of this kind are: *make, move* and *use.*

Although there is clearly some overlap between discourse organizing vocabulary items and those items identified as 'subtechnical' by other writers, most EAP practitioners consider subtechnical vocabulary as something apart from common defining vocabulary which is largely made up of very high frequency words taught in the early stages of a language learning programme.

1.4.4.3 Subtechnical vocabulary identified in corpora

Two approaches have been discussed regarding the definition of subtechnical vocabulary. The first approach is concerned with identifying subtechnical vocabulary on the basis of how difficult certain items are for learners, Higgins' is an example of this kind. This approach is difficult to replicate, as it is hard to decide on the difficulty level of a word. The second approach defines subtechnical vocabulary in terms of its discourse organizing function. This approach is important to discourse analysts, but is likely to include core vocabulary in the subtechnical category, because core vocabulary items are used to link the ongoing discourse. This subsection focuses on a different approach to defining subtechnical vocabulary, an approach based on the criteria of frequency and range. The works of Barber (1962), Inman (1978), Farrell (1990) and Yang (1986) will be discussed.

Barber (1962) analyzed three texts from Electronic Engineering, Biochemistry, and Instrumental Optics, with a total of 23,400 words. He created a 'general scientific' vocabulary list which included:

nouns:	*equation, method, phenomenon, principle, process, series;*
qualifiers:	*accurate, complete(ly), considerable(y), constant, definite, positive, similar;*
verbs:	*alter, approach, assume, consist, indicate, obtain, occur, require, vary.*

All these words were reported to be high frequency words in Science and Engineering texts, and were not contained in West's **General Service List of English Words** (1953). Barber's work is useful in that it is based on an objective approach of frequency and range, unlike some of the subjective and intuitive approaches reviewed earlier. However Barber's work did not take range into serious consideration. This can be seen from the way he defined the subtechnical vocabulary in his corpus: 'high frequency words not contained in West's General Service List of English Words (GSL) 1953 (Barber, 1962).

Inman (1978) had a much larger corpus of 114,460 running words which covered ten subject areas. Her definition of subtechnical vocabulary was based on Cowan (1974), who described subtechnical vocabulary as 'context independent words which occur with high frequency across disciplines'. Inman classed 20% of the words in her corpus as technical (with high frequency low distribution across fields), 9% as function words, and 70% as subtechnical words. Her function words were closed set grammatical words. Inman did not create a separate category for high frequently occurring General English words, which explains why the proportion of subtechnical words in her study was high. It was also possible that she included low frequency items in her subtechnical list (probably because of the size of her corpus), because she did not state any minimum cut-off point for her subtechnical vocabulary.

The precise make-up of Inman's subtechnical vocabulary list remains unknown, as the complete results of her research do not seem to be available in any published form. In her article in Trimble (1985), she listed only ten subtechnical vocabulary items: *high, system, result, process, function, form, temperature, large, solution* and *structure*, forms which according to Inman occurred with high frequency across fields. Nevertheless, Inman's method of identifying subtechnical vocabulary based on corpus analysis paved the way for later works, especially the study conducted by Farrell (1990) who used Inman's lemmatization rules for the presentation of his data.

Farrell's study aimed to define 'subtechnical vocabulary' and using this definition to draw up a list of subtechnical items in the area of Electronics. Farrell worked with a 20,000 word corpus, which consisted of ten 2,000 word samples of running text in the area of Electronics. He produced a list of the lemmas with an overall frequency and range figure, with the lemmas listed in rank order of frequency and range, as the following example shows:

1. the	*2034/10*
2. is	*879/10*
3. a	*850/10*
4. of	*692/10*
5. to	*439/10*
6. and	*394/10*
7. in	*367/10*
8. current	*284/10*
9. it	*207/10*
10. fig.	*207/7*

(1990: 74, Appendix 3)

By dividing the main lemmatized list he had three word lists: general words, semi-technical words and technical words. Farrell's general words included words which learners who come to an ESP course might know, these words were grammatical words such as *the, be, a, of, to, and, in, it, this, by* and *as*, some common verbs such *as use, have, make, move, see, give, call, pass, need, find* and *open*, common adjectives such as *small, long, large, strong, important* and *difficult*, and concrete nouns such as *paper, page, picture* and *test*. Farrell's criteria for identifying these general words was their 'commonness', although he did not say explicitly that he identified these general word items on the basis of their high frequency and wide range.

Farrell's subtechnical words totalled 3,487 tokens, covering 17.43% of the total occurrences of 20,000 tokens. Farrell defined subtechnical 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; words with high frequency across scientific disciplines' (Farrell, 1990:11)

This is an example of Farrell's subtechnical vocabulary:

fig.	supply	alternating	difference
shown	output	high	metal
produced	direction	effect	balance
connected	flow	increased	turn
change	varying	low	type
measure	emitter	deflection	horizontal

(1990:79, Appendix D)

Of the total of 508 lemmas (types) occurring more than five times in Farrell's corpus, 224 were classed as semi-technical, 143 as general, and 141 as technical. Semi-technical lemmas were 44% of all lemmas, technical lemmas were 27.7%. Farrell's figure for technical vocabulary was very close to Inman's (Farrell's was 27.7%, and Inman's 21%) while his figure for subtechnical vocabulary was not as high as hers. This was because Farrell used a different selection procedure for subtechnical vocabulary. Farrell classified words as technical rather than subtechnical if they were used in a variety of scientific disciplines but occurred with a specialized meaning in Electronics (for example, *base, resist, field, potential, frequency, positive* and *negative)*. Inman's third category of words was function words rather than general words, and many words which would seem likely to be identified under a General English section were in Inman's subtechnical section. Words of this kind were: *close, open, important* and *difficult*.

Inman and Farrell seem to agree on the same approach to identifying technical vocabulary as those items that occur with high frequency but low range across disciplines. Farrell's study adopted a more rigorous approach when selecting subtechnical vocabulary. Unlike Inman who only separated function words from subtechnical words, Farrell separated General English words from his subtechnical list, too. This is probably a more useful approach (when using a corpus based study to compile word lists) because as mentioned earlier, General English words tend to be the items taught in the early stages of a language programme.

However, Farrell's method of eliminating 'general words' is unclear. He accepted as a starting point that general words were those items that learners would know before they embark on an ESP course, but it is not clear how he decided which words were general. He simply removed 'common grammatical words', 'common verbs', 'common adjectives' and 'concrete nouns'. Thus Farrell's method appears to have been to some extent dependent on personal experience and intuition. Both *high* and *low* for example, were included in Farrell's semi-technical word list; this indicates these two words did not belong to his 'common adjectives' category, although his common adjectives included *small, long* and *large*, which belong in the same semantic set. All these words are included in the most frequent 5,000 words of Thorndike and Lorge (1944). Farrell also counted 'concrete nouns' as general words but included *metal, tube, wind* and *flies* in his subtechnical list. These words are concrete nouns, and all occur in the first 5,000 words in Thorndike and Lorge (1944). The word *metal* might be a topic specific word, too, and thus it might be argued that it should be in the technical word list. Farrell's subtechnical word list also included grammatical words such as *since, thus, such, though, hence, therefore* and *during*, which might have been placed in a general word list because they are frequently occurring closed set items.

A further problem is that Farrell's definition of subtechnical vocabulary is misleading. As mentioned earlier, his subtechnical words included those 'Formal, context-independent words with a high frequency and/or wide range of occurrence across scientific disciplines, not usually found in basic General English; words with high frequency across scientific disciplines' (1990: 11). This definition can be paraphrased as follows: semi-technical words have a high frequency **and** wide range, or have a high frequency **or** wide range of occurrence. This definition does not seem to regard frequency and range as two independent parameters, but rather as alternatives.

A more objective definition based on frequency and range has been offered by Yang (1986). Yang based his analytical work on a scientific English corpus which consisted of nine texts and totalled about 300,000 words. The nine disciplines of Yang's corpus were as follows: Electronics, Electricity and Modern Physics, Electricity and Magnetism, Electromechanics, Thermodynamics, Mechanics, Biochemistry, Physical Chemistry, and Medicine (Yang, 1986: 102). The aim of his study was to develop a computerized method of identifying technical terms in scientific texts, therefore one of the focuses of the work was concerned with what did or did not constitute a technical term. According to Yang, words with high distribution and high frequency were function words, words with high distribution and lower frequency were subtechnical words, and words with a peak frequency in one or several fields were considered technical terms.

Subtechnical vocabulary was identified as items having a high distribution across different disciplines. Although Yang's definition of technical vocabulary did not distinguish the frequency and range differences between subtechnical and technical vocabulary, an important contribution of Yang's work was that he did not force-fit words into the only one category, because this does not always reflect the language. A word may occur across several disciplines, but at the same time have a peak frequency in one particular discipline. Yang regarded this type of item as both technical (in the particular discipline in which it peaks) and subtechnical (across disciplines). For example, *solution* and *work* were subtechnical words in Yang's corpus because of their high distribution. They were, however, also regarded as technical terms in the fields of Chemistry and Mechanics, because they had a peak frequency of occurrences in these two fields. Examples of subtechnical words from Yang's list were:

> *absolute, accuracy, act, basic, basis, bearing, calculate, carry, conclusion, decrease, define, definition, effect, effective, electrical, face, fact, factor, feature, give, increase, introduce, know, possible, problem, result* and *total*

The fact that Yang's work is purely based on frequency and range can be regarded as both an advantage and a disadvantage. The advantage is that he employed objective criteria to identify subtechnical and technical vocabulary. All the previous work, including Farrell's, lack this consistent objective criteria. Farrell's work for example, seemed to have involved some subjective decisions regarding the elimination of general words. The disadvantage of this method was the disadvantage which all frequency studies have. Using the frequency and range criterion it is possible to miss some subtechnical or technical terms, simply because their patterns of frequency and range do not match the one recognized by the method. Similarly it is also possible to include some words which on intuitive criteria would seem to be neither subtechnical nor technical, but nevertheless conform to the recognized frequency and range pattern for technical or subtechnical words in the corpus under analysis. Decisions regarding how high 'high frequency' is, and how wide 'wide range' is still remain subjective. By changing the cut-off point slightly, the category into which a word falls can be altered.

Yang's criteria of frequency and range are the most satisfactory reviewed so far. The fact that no intuition or personal judgment was involved in categorizing words makes his study easy to replicate, and his idea that some words could fall into both the subtechnical and technical categories was a particular contribution to methods of categorizing words. Therefore, Yang's criteria will be used as the primary means to categorize words in my own studies.

1.5 Research Subjects

This research aims to investigate some issues related to English for Specific Purposes. More specifically, it aims to show how an investigation of vocabulary needs can contribute to the design of a course in English for postgraduate Engineering students. My subjects were taking a Masters programme at the University of Warwick and were from Southeast Asia: China, Hong Kong, Taiwan, Thailand and Malaysia. Each year a large number of such students come in the Engineering Department. Table 1.1 shows the course participants from the Southeast Asian region in the past four years:

Table 1.1 The enrollment of Southeast Asian students for the Msc. Engineering course			
Year	**Number of students from Southeast Asia**	**Total number of students**	**Percentage (%)**
1993	129	196	66
1994	125	201	62
1995	126	193	65
1996	103	196	53

Most of the subjects have completed their undergraduate studies in their home countries (or regions), where English is not the medium of instruction. In China, for example, Mandarin is the predominant medium of both written and spoken communication for all communities, at home, at work, in schools and tertiary institutions. Students have little if any exposure to English. The situation in Hong Kong and other Southeast Asian countries is similar: Cantonese is the 'mother-tongue of approximately 98% of Hong Kong's citizens' and Cantonese serves the 'essential social interactive function within the Hong Kong community, and between Hong Kong and other Cantonese communities in Southeast Asia' (Working Group, 1989: 4). In addition to this, written Chinese acts as a common means of communication and as a way of sharing common cultural and literary values within the wider Chinese community. In Hong Kong, Cantonese is used in the home, on the street, in the community, in the world of work, at school and in the media. Naturally the use of English in Hong Kong is more limited. In this respect, the situation is similar to that in China: English is more a 'foreign' language than a 'second' language. There is limited environmental support outside school for it, and students tend not to take advantage of English language newspapers and other reading materials, or of English medium TV and radio programmes. As a result, '... the majority of Hong Kong students experience great difficulty in using English as a medium of learning and in developing what are considered to be effective communication skills in English in school' (Working Group, 1989), and there is a widespread expression of belief that the standard of English is declining.

Taiwan and Thailand are in a similar situation to that of China. Mandarin in Taiwan and the Thai language in Thailand are the medium of both written and spoken communication. Students from Malaysia may have a little more exposure to the English language. However, only a few students from Malaysia participated in my studies, and the majority of the participants were from the other countries (or regions) mentioned above.

1.6 Framework of the thesis

Using corpora based lexical analysis, this research aims to establish the vocabulary learning needs and goals of the Engineering students from South East Asia studying at British Universities. The lexical analysis of the research is based on samples of Engineering textbooks, and mainly focuses on subtechnical and technical vocabulary, with the attempt to discover which type of these two categories of vocabulary the students have a better knowledge of. Studies of subtechnical and technical vocabulary seem to offer useful pedagogical implications in the teaching of these two types of vocabulary to ESP learners.

The thesis reports an analysis and three studies. The analysis aims to discover the characteristics of subtechnical and technical vocabulary identified on the basis of frequency and range, and to investigate the validity of the frequency and range criteria as a means of identifying categories of vocabulary in ESP textbooks. Study One and Study Two examine the Engineering students' receptive knowledge of subtechnical and technical vocabulary. And Study Three is a follow up study - an examination of whether lexical familiarization plays a part in the subjects' recognition of technical vocabulary.

1.6.1 Characteristics of technical and subtechnical vocabulary

As can be seen from 1.4, subtechnical vocabulary has been a controversial issue among researchers and ESP practitioners. First of all, the definition of subtechnical vocabulary has been rather inconsistent. One of the earliest definitions was given by Cowan (1974) who considered subtechnical vocabulary to be 'context-independent words'. This definition has been used by Inman (1978), Baker (1988) and Farrell (1990) in determining the nature of a subtechnical word in their studies. This definition, however, leaves many issues unresolved.

Although many researchers and ESP practitioners have discussed the notion of subtechnical vocabulary, few have attempted to elaborate a comprehensive list of subtechnical items on the basis of empirical studies, neither has subtechnical vocabulary been precisely defined. On the basis of the definitions given in 1.4, the present research aims to examine the characteristics of subtechnical and technical vocabulary, using samples of the Engineering textbooks that the students are expected to encounter. The research examines, for example, the behaviour of these types of words beyond the pattern of frequency and range.

1.6.2 Validity of frequency and range

Frequency studies in ESP have been conducted since the 1960s. Early and important frequency studies include Barber (1962, see 1.4). Frequency studies are first and foremost of a quantitative nature, based on the assumption that it is the significantly frequent occurrence of certain forms and structures that characterises scientific writing. As a consequence 'statistical methods play an important role in selecting an inventory for teaching purposes. ... It is the word and the phrase levels that yield the best results, i.e. lists of typical lexical items which may serve as a highly effective teaching/learning minimum' (Hoffman, 1981).

The second aim of the thesis is to investigate whether frequency and range are good criteria in the study of ESP textbooks, and particularly as a means of identifying technical and subtechnical vocabulary. As discussed in 1.4, one way of deciding whether a word is subtechnical is to rely on language teachers' experience or judgment of how difficult a word is. Another method of identifying subtechnical vocabulary is from the perspective of its discourse organizing functions. Both methods have been used by researchers to draw up lists of subtechnical and technical vocabulary. This research will draw up word lists on the basis of frequency and range, and compare the features of these words with those drawn up by using other criteria, to see whether the objective method supports some researchers' intuitive judgments. If it does, computer generated lists can be used instead of other researchers' subjective lists, because they will be complete, and relevant to the specific needs of any given group of learners, and do not require any insight or experience.

1.6.3 Technical and subtechnical vocabulary: the receptive knowledge

Word knowledge is strongly related to reading comprehension. In a survey of learners' reading problems, Yorio (1971) found that the subjects regarded vocabulary as the greatest obstacle. More recently a survey was carried out by Li (1989) among Chinese postgraduates and visiting students in eleven universities in Britain. 72.5% of the participants admitted that vocabulary is the main problem for their study and research.

Anderson and Freebody (1983) and Dunmore (1989) claim that lexical knowledge is an important component for reading ability. Nation and Coady (1988: 98) came to a similar conclusion based on findings from a number of empirical studies that '... vocabulary knowledge would seem to be the most clearly identifiable sub-component of the ability to read'. This is echoed by Elliman (1980) who claimed that when the proportion of unknown words rises above 10%, reading comprehension is severely affected. The importance of vocabulary is also realized by ESP practitioners, such as Salager (1983: 55) who proposed that 'it would be sensible to consider vocabulary as a key element in teaching ESP.' Vocabulary could be more problematic for ESP readers, because they need to read specialist texts, in which case they need to deal with vocabulary that they do not meet in General English language reading. In discussions of subtechnical vocabulary, it is often suggested that items of this type present particular difficulties for learners. Robinson (1980), for example, suggested that:

> Coursebooks ... perhaps do not need to concentrate on the very specialised vocabulary items as students will get these from other sources. Rather it is the sub-technical level which is often difficult (1980: 7).

Similarly, Kennedy and Bolitho (1984) suggested that the technical vocabulary of a subject specialism arises in context in the specialist classes and is not the language teacher's responsibility. Subtechnical vocabulary, or 'words which are not specific to a subject speciality but which occur regularly in scientific and technical texts' are words which learners frequently find difficulty with, and should be given high priority in the language programme (1984: 57). Higgins (1985) also pointed out that 'It is not the job of the English teacher to teach technical vocabulary; it consumes too much time, and he will probably not do it well' (1967: 32). Instead Higgins suggested that the English teacher should concentrate on 'frame' words, a list of words drawn up in association with subject teachers, and defined as 'words which were causing difficulty, words which, although not technical terms, are frequently used in technical writings'.

These claims, however, are either based on intuition, or personal experience. None of these claims are made on the basis of empirical analysis, therefore we cannot confirm whether they reflect a real ESP situation. This research, therefore, investigates whether subtechnical vocabulary is indeed the subset of vocabulary with which learners find greatest 'difficulty', as claimed by Higgins (1985), Robinson (1980), and Kennedy and Bolitho (1984).

In this thesis the word 'difficulty' is indicated by learners' problems recognizing the word. Not recognizing a word suggests that a learner is experiencing some problems installing this word in his/her mental lexicon. This research only examines learners' receptive knowledge of technical and subtechnical vocabulary, rather than their productive knowledge.

1.6.4 Lexical familiarization

I assume that lexical familiarization in textbooks helps readers to acquire receptive knowledge of technical vocabulary. For my final study, I examine the extent to which words selected for my previous studies are lexically familiarized in the textbooks. A modified version of Bramki and Williams's criteria was used to categorize the different types of lexical familiarization occurring in my data.

1.7 Organization of the thesis

The present research aims, first of all, to examine the vocabulary of the textbooks that Southeast Asian postgraduate Engineering students encounter in their studies. This examination is carried out using frequency and range criteria. Secondly, the research aims to empirically enquire whether students have a better receptive knowledge of subtechnical or technical vocabulary.

To fulfill these objectives, the following research questions will be addressed:

1. What are the characteristics of technical and subtechnical vocabulary identified in the corpus according to the criteria of frequency and range?

2. What is the validity of the frequency and range approach in the study of technical and subtechnical vocabulary?

3. Do ESP students have better receptive knowledge of technical vocabulary or subtechnical vocabulary?

4. To what extent and in what way is technical vocabulary lexically familiarized in textbooks? How does lexical familiarization affect students' receptive knowledge of technical vocabulary?

Chapter Two discusses issues of overall vocabulary size. The chapter discusses and evaluates ways of measuring learners' vocabulary size, and reports the measurements of the vocabulary knowledge of overseas Engineering students at Warwick University. Chapter Three presents some initial data which forms the hypothesis of the research. Chapter Four presents the data on which the main studies are based. Chapter Five analyses the data in order to answer the first two research questions presented above. Chapter Six describes studies of students' technical and subtechnical vocabulary in order to answer my third research question. Chapter Seven examines the lexical familiarization of technical words identified in the previous study. Chapter Eight summarises the research findings and proposes some pedagogical implications. And finally Chapter Nine draws conclusions, discusses limitations of the research and makes recommendations for future research.

Chapter Two

Research Subjects: The Vocabulary Size

2.1 Introduction

As stated in Chapter One, my third research question concerns the receptive knowledge of technical and subtechnical vocabulary. To investigate this, I needed to develop lists of these two types of vocabulary. For this purpose, those familiar words already known by the students needed to be identified first, so that those words could be excluded from the lists of subtechnical and technical vocabulary. Decisions therefore, needed to be made as to what words the subjects in this study already knew. This chapter first of all describes various methods teachers and researchers have used to examine what vocabulary learners know or have learned, then it presents a study to assess the vocabulary size of the subjects that participated in this research.

2.2 Studies of vocabulary size

There has been a great deal of research into the vocabulary size of native speakers, extending back to the end of the last century. However, as Anderson and Freebody (1981) pointed out in a review of the research, it is difficult to have much confidence in the results of these studies because of their wide variation, resulting from different calculation methods.

According to Seashore and Eckerson (1946), an educated adult might know more than 150,000 words. This figure does not give too much information, because of the problems of defining 'word' and the difficulty of finding a reliable procedure for assessing vocabulary knowledge. Since Seashore and Eckerson were pioneers of a method widely used in studies of vocabulary size, it might be useful therefore to review their study and findings, although they are thought by later researchers to have overestimated their subjects' total vocabulary knowledge.

Seashore and Eckerson defined a word as an item listed in Funk and Wagnall's **New Standard Dictionary of the English Language** (1937), which contains approximately 450,000 entries. They reduced this to 370,000 by omitting alternative meanings. About 166,247 of these were 'basic words', such as *loyal*, and the remaining 204,018 were derivatives and compounds, such as *loyalism, loyalize, loyally* and *Loyal Legion*. They obtained a representative sample of the total by extracting the third word in the first column of every left-hand page. This gave a list of 1,320 words, which was divided into four. College students were tested on their ability to define the words on each list and to use them in illustrative sentences.

Seashore and Eckerson found that on average their subjects knew 35% of the common 'basic words', 1% of the rare 'basic words', and 47% of the derivatives and compounds. Applying these proportions to the overall number of words in the dictionary, they found that the average college student knew approximately 58,000 common 'basic words', 1,700 rare 'basic words' and 96,000 derivatives and compounds. The overall total was 150,000 words.

Lorge and Chall (1963) have pointed out a number of weaknesses in Seashore and Eckerson's methodology. For example, the students might have guessed the meaning and use of derivatives from a knowledge of the 'basic words' to which they are related. Also, a high proportion of high frequency words was sampled. According to Lorge and Chall, the Seashore and Eckerson study contained more than twice as many words within the first 20,000 frequency range as would be expected to be representative. The effect of a big dictionary is another problem: the bigger the dictionary used, the more words people are found to know, probably because bigger dictionaries include more homonyms. Finally it is also difficult to know what level of knowledge is being tested: one person claiming to know a word might think of one meaning, while another person might think of another meaning.

An early study of foreign language learner's vocabulary size was by Zettersten (1979). This study was concerned with cross-national comparisons of vocabulary knowledge. It was designed to test the English vocabulary proficiency of first year students of English at universities in Denmark, Finland, Norway, and Sweden. Word samples were chosen from **The Teacher's Word Book of 30,000 Words** (1944). The result of this study showed that the first year students of English in all four countries had roughly the same recognition vocabulary. However, no precise figures of the vocabulary size were reported.

Zettersten's study influenced a later study by Gui (1987), which originated in a survey to estimate Chinese students' vocabulary size as a basis for national syllabus design and curriculum development. Gui selected 200 words through random sampling from the first 10,000 words from **The Teacher's Word Book of 30,000 Words** (1944). These 200 words were then divided into eight different groups corresponding to eight different frequency bands of English words. Specifically, there were twenty words in each of the first seven groups, making up 140 words. The eighth group had sixty words. Multiple choice questions were set in Chinese. Each word had four different translations of which only one was correct.

286 Chinese students were tested, of different subject specialisms and different levels. Some were studying English as the main subject, others not; some were university first year students, others were graduates preparing to go to an English speaking country for further studies. The results of this study showed that a secondary school student in China knew approximately 1,200 English words and a graduate student knew approximately 6,000 English words.

One of the problems with this study was again the question of what the researcher considered a 'word', and whether this was only a base or head word, or also an inflected form. I examined Gui's list of 200 words, and found that there were base words such as *problem* and *cover*, proper names such as *Philistine* and *Catholic*, compounds such as *well-known* and *businesslike*, derivatives such as *indignantly*, and inflections such as *lain*. Therefore it seems Gui's selection criteria were rather inconsistent.

Another problem with this study concerns the multiple choice format. The presentation of the four choices was also rather inconsistent. For example, different parts of speech were used as distracters for the same item, some distracters seemed likely to give clues to the testees, and others were very similar in meaning so that the testees would have difficulty in distinguishing them.

Other studies on native speaker vocabulary size were Kirkpatrick (1907) and Diller (1978). These two studies were very similar. In the first place, both studies used the same format: testees were asked to identify three categories of words - words they were absolutely sure of, words they were not sure of, and words they did not know. Secondly both tests were testing receptive knowledge of words. However, the results were different: Kirkpatrick reported that his subjects had a vocabulary size of 20,120; while for Diller, the estimated vocabulary for his subjects was 216,000 words. Neither Kirkpatrick nor Diller made a distinction between basic words and derivative words while sampling from the dictionary, it was possible therefore, that they both included derivative forms of words. The reason that they reached widely varying estimates was probably because of the dictionaries they used. Kirkpatrick used an abridged dictionary which contained only 28,000 words, whereas Diller used a dictionary that consisted of 450,000 entries.

The yes/no checklist method that they both used was very straightforward and easy to conduct and was adapted in later studies by Meara and Jones (1987) and Meara and Buxton (1987). However, the method used by Kirpatrick and Diller had its limitations. In the first instance, individual habits of thinking or judging were probably the largest factor in tending to make the test unreliable. Some subjects may have thought they knew a word simply because this word aroused a feeling of familiarity, while other subjects might have marked words as known only if they were confident they could give a correct definition. This difference of understanding cannot be avoided because Kirkpatrick and Diller did not give a clear indication of what was meant by 'knowing' a word. They gave no indication that any discussions of the meaning of 'knowing' took place before the test.

Meara and Jones (1987) developed a test based on frequency counts. Like Kirkpatrick (1907) and Diller (1978), Meara and Jones also used the checklist format, but they made an important improvement to the scoring method.

In sampling the words to be included in the test, Meara and Jones considered all grammatically inflected forms (regular plurals, regular *-s, -ing,* and *-d* forms of verbs, regular comparatives and superlatives) as variants of the head word. All derivatives were counted as separate words. Like Zettersten, Meara and Jones also used **The Teacher's Word Book of 30,000 Words** (1944). A sample was drawn by taking every tenth word at random and a total of sixty words from each band of 1,000 was drawn. As mentioned earlier, a yes/no format was used. What distinguished their studies from other studies (Kirkpatrick, 1907; Diller, 1978) was Meara and Jones' use of non-words as a measure to penalise 'dishonest' testees. The non-words were not real English words, but they were plausible and obeyed the phonological, graphological and morphological rules of English. Testees were asked to read each word on the list and to tick the ones they knew. Subjects were informed of the fact that there were non-words, therefore they knew that they were taking a risk in claiming to recognize a word which they were not absolutely sure of. In this way, the chance of dishonesty was reduced.

The main advantage of the checklist method is that a large number of items can be tested in a short period of time. Meara and Jones' test can be quickly administered and marked, so a large amount of data can be gathered relatively easily.

It should be noted that a yes/no test can only test the receptive aspect of vocabulary knowledge, and is not an adequate basis for determining whether someone really knows a word. Such a test is basically a test of familiarity. Meara and Jones' test assumes that familiarity with an item is an adequate indicator of language knowledge for placement purposes.

2.3 Nation's Vocabulary Levels Test

This section describes Nation's Vocabulary Levels Test (1990), which will be used to measure the general vocabulary size of the learners described in 1.5.

The Vocabulary Levels Test (1990) was a study based on academic word lists, therefore this section first of all reviews four standard word lists, all of which were compiled for academic purposes. Some of the word lists are based on a count of words that occur in university textbooks, other word lists are based on other academic writing materials such as published lectures. These word lists take into account the range of disciplines in which the words are found as well as the frequency of occurrence. Most academic word lists are compiled by excluding both high frequency general words (such as those in **A General Service List of English Words** or in the first 5,000 word band of **The Teacher's Word Book of 30,000 Words**) and low frequency narrow range words which consist largely of technical terminology.

2.3.1 Academic word lists

The word list developed by Campion and Elley (1971) was intended to represent the vocabulary commonly encountered by first year university students in New Zealand. It was selected from a sample of nineteen academic disciplines: Accountancy, Anthropology, Biology, Chemistry, Economics, Education, English, French, Geography, Geology, History, Law, Mathematics, Mechanics, Philosophy, Physics, Psychology, Political Science and Sociology. Three different sources were chosen for sampling: textbooks, published lectures, and examination papers. For textbooks, Campion and Elley chose only the first 200 pages in each textbook to ensure an approximately equal contribution from each discipline. In case the books in question contained fewer than 200 pages, words were drawn from an additional book in the same discipline. Only the first five lines on each page of a textbook were inspected, 'to avoid overloading the list with infrequently used words, occurring several times on one page'. As for published lecture notes, approximately 3,000 word forms were inspected from the same nineteen subject areas. Finally for examination papers, all the words in a sample from each of the nineteen disciplines were scanned. Table 2.1 shows the details of Campion and Elley's corpus:

Table 2.1 Composition of the Campion and Elley corpus	
Source	**Number of word forms**
Textbooks	234,000
Lectures	57,000
Examination papers	10,800
Total	**301,800**

In order not to 'overload the list with the simple and basic vocabulary found in any standard word count', they decided to omit the most frequently encountered 5,000 words, as listed in the Thorndike and Lorge count (1944). This decision was based on the assumption that these words were familiar to most students entering New Zealand universities. They also omitted technical terminology (for example, *microamperes, ribonucleic*) encountered in the study of these disciplines. Neither were foreign expressions (such as *sine die, raison d'etre*) or proper names (such as *Wellington, Einstein*) included. All plural forms were converted to the singular unless they had acquired a meaning distinct from the singular (for example, *data, cross-purposes*); all verbs were converted to the infinitive form, with occasional instances of the past tense of a verb used as an adjective (for example, *conditioned, stratified, ill-fated, interrelated, landed*); comparative and superlative forms of adjectives and adverbs were also reduced to the positive form. With these criteria, the final list yields 3,200 different words or types out of the 301,800 word forms.

Praninskas (1972) used a similar approach to Campion and Elley's to develop the American University Word List drawn from ten university level textbooks of various disciplines: Biology, Chemistry, History, Literature, Mathematics, Philosophy, Physics, Psychology, Rhetoric, and Sociology. This list, which totalled 272,466 word forms, was compiled for non-native speaker students studying

English. In the Praninskas list, words appearing in West's General Service List of English Words (1953) were excluded, with the same rationale as Campion and Elley that these words were already familiar to most non-native speaker students entering universities. Unlike Campion and Elley, Praninskas did not report how technical words were treated, and whether technical words were included in or eliminated from the list.

Lynn (1973) used a different approach to develop a list that he claimed to be 'both needed and unknown' to university students in Singapore. Instead of using textbooks as sample materials, Lynn recorded students' annotations in textbooks. The rationale behind this was based on the principle that a word list was only useful if it contained words both needed by the students and unknown to them at the same time. Lynn put this technique to use in a study of students in a college of commerce in Singapore. Fifty two books and four cyclo-styled hand-outs (from approximately fifty students) were examined, and a record was made of every instance where a student had annotated a book by adding a Chinese or English equivalent. A total of 10,000 such annotations was recorded. The lexical items most frequently annotated were identified as those items which were needed and yet were unknown. It was reported that the most striking feature of the resulting list is the absence of technical words. It was found that 'textbook English' words (non-technical terms from the academic register) presented the greatest problems to the students.

A similar word list was developed by Ghadessy (1979) by using the same approach as Lynn's. Ghadessy also believed that the words most often glossed by students were words students needed but found difficult. To record these items, Ghadessy took twenty textbooks in related science disciplines: eight books were for Chemistry, seven were for Biology, and five were for Physics. The total number of annotated items was 478,700 word forms, of which 111,7 were different words or types.

The difference between Ghadessy's word list and Lynn's list was that the former provided information on frequency of occurrences and the frequency with which a word had been glossed by the students. It thus indicated how many students did not know these words.

The University Word List (Xue and Nation, 1984) was mainly derived from Campion and Elley's list (1971) and the Praninskas list (1972). The two lists were combined and included only 'base' forms, which was also the criteria adopted by Campion and Elley: plural forms were converted to the singular; verbs were converted to the infinitive form; and comparative and superlative forms of adjectives and adverbs were also reduced to the positive form. After the two lists were combined, they were checked against Lynn's list and Ghadessy's list. It was found that 70% of the words in each of the Lynn and Ghadessy lists overlapped with the combined list. The high frequency non-overlapping words in these two lists were added to the combined list. The final University Word List contains 737 base words, even though derivative forms were sometimes more frequent than the base forms.

Table 2.2 Sources of the University Word List	
Entries from different lists	**Words included in the combined list**
500 most common words in the Campion and Elley list	323
The Praninskas list	291
The 3,200 word lists in Campion and Elley	64
The Lynn list	54
The Ghadessy list	5
Total	737

All these word lists were developed with regard to the specific needs of a particular type of students, although the methods of doing this were not exactly the same. The Campion and Elley list for example, was intended to represent the 'commonly encountered' vocabulary of the nineteen academic disciplines, and therefore was developed on the basis of the general usefulness of the words: these words occur across all the disciplines and will be encountered by the students in their academic studies. Technical terminology was eliminated from their list, because technical words only occur within specific disciplines. Campion and Elley's idea of eliminating general and technical words influenced subsequent studies such as Farrell (1990). My own study of subtechnical vocabulary also adopted this idea.

Lynn's finding that most annotated items were non-technical rather than technical was particularly interesting to me. It seems that he was one of the few practitioners who empirically studied students' vocabulary difficulties and reached this conclusion. Many other claims (for example, Higgins, 1985; Trimble, 1985) that students do not find technical words difficult are based on intuition. In the following chapter of this thesis, I describe a small scale interview aiming at obtaining information about students' vocabulary 'difficulties', so that further more in-depth investigation can be carried out.

These academic word lists are very useful resources for teaching and testing purposes. Nation's Vocabulary Levels Test (1990: 261) took word samples from The University Word List described above. The next section describes this test in more detail, together with a survey that I conducted to estimate the vocabulary size of my own subjects.

2.3.2 My subjects' vocabulary size

My own interest in foreign language learners' vocabulary knowledge originated from the comments made to me by some Engineering students who were studying at postgraduate level at the University of Warwick. They attributed all kinds of language problems to vocabulary deficiency. I often heard them say, 'I would understand this if I had a larger vocabulary', or 'My vocabulary is too small'. In order to gain a more thorough insight into their vocabulary knowledge, I conducted a survey among students studying in the UK. The students who took part in the survey were mainly from Mainland China, but some came from Hong Kong.

The survey consisted of two parts. The first part was a questionnaire, aiming to find out some factual information about the subjects, such as their age, the course they were studying, and the length of their stay in the UK. The second part of the survey was a Vocabulary Test, with the purpose of reaching an estimate of their vocabulary knowledge. I chose to use Nation's Vocabulary Levels Test (1990). This test was chosen because it was straightforward, easy to administer and mark. I decided to use the test to measure the subjects' knowledge of the 2,000, 3,000, and 5,000 most frequent words and also their knowledge of academic words. Therefore, the 10,000 (low-frequency vocabulary) word level was not part of the test.

Nation's Vocabulary Levels Test (1990) was designed to assess knowledge of three kinds of words: high frequency words (at the 2,000, 3,000, 5,000 word levels), academic vocabulary (represented by words from the Campion and Elley list), and low frequency vocabulary (at the 10,000 word level). The high frequency words (2,000, 3,000 and 5,000 word levels) and the low frequency words (at the 10,000 word level) were chosen from the **Teacher's Word Book of 30,000 Words** (Thorndike and Lorge, 1944). The academic vocabulary was chosen from the Campion and Elley list. As described in Chapter Two, this list does not include the most frequent 5,000 General English words or the technical words.

The test employs a word-definition matching format, but in a reversal of the standard practice, the testees are required to match the words to the definitions. In this way the definitions are the test items rather than the words. At each of the five levels, there are thirty six words and eighteen definitions in groups of six and three respectively, as in the example below:

apply	
elect	
jump	- choose by voting
manufacture	- become like water
melt	- make
threaten	

The subjects were asked to choose the right words from the left column to match the phrases or definitions in the right column. All the words in each group are the same part of speech, this is to avoid giving any clue as to meaning based on form. This is an improvement on Gui's (1987) choices of words. Words and definitions related in meaning are not grouped together so that subjects do not have to differentiate words with slightly different meanings. The test is designed as a broad measure of word knowledge and it is not intended to require the testees to differentiate between semantically related words or to show an awareness of shades of meaning.

2.3.2.1 Subjects

The subjects that took part in my survey were categorised into two groups according to their background culture and their specialisms. The first group was a miscellaneous group comprising postgraduate students from China studying a variety of disciplines at British universities. The second group comprised all the Hong Kong Chinese students studying a postgraduate taught course in the Engineering Department at the University of Warwick. Eighty four students participated in this study, of which sixty six were Mainland Chinese, and nineteen were Hong Kong Chinese. Of the Mainland Chinese students, twenty six were based at the University of Warwick, the others were at Aberdeen University, Aston University, Birmingham University, Bradford University, Coventry University, Liverpool University, London Imperial College, London University, Loughborough University, Nottingham University, Oxford University, Sheffield University, and Southampton University. Although all the students in this group came from Mainland China, they were a heterogeneous group in other aspects. In age, they ranged from twenty one to fifty six. Twenty nine were studying Engineering, seven were studying Physics, four were studying ELT, four were studying Business Administration, two were studying Law, two were studying Chemistry, and ten were studying Economics, Mathematics, Chemistry, Computer Science, Education, International Studies, British Studies, Social Studies and Psychology. The specialisms of four subjects were not known. The time they had spent in the U.K. varied, too. Two Engineering subjects had already been in Britain for sixty months, while two subjects had only just arrived.

2.3.2.2 Administration

The subjects did the test in their own rooms under the following conditions:

- no pre-teaching or preparation was involved.
- dictionary use was not allowed.
- subjects could take as much time as they needed.

Subjects were given separate scores for each of the four tests.

2.3.2.3 Results

Details were gathered regarding the students' length of stay in Britain, and whether or not they had attended a pre-departure course. Those students who had been in Britain longer, or who had attended a pre-departure course, did not gain significantly higher scores. Although the students from China had slightly better results, there was no significant difference between the two groups in all the four tests.

Table 2.3, 2.4 and 2.5 show the results of the 2,000, 3,000 and 5,000 word level tests respectively.

Table 2.3 Scores on the 2,000 word level		
	Mean	**SD**
Mainland Chinese	95.17	7.92
Hong Kong Chinese	88.71	8.98

Table 2.4 Scores on the 3,000 word level		
	Mean	**SD**
Mainland Chinese	92.88	10.35
Hong Kong Chinese	88.54	11.97

Table 2.5 Scores on the 5,000 word level		
	Mean	**SD**
Mainland Chinese	85.79	13.73
Hong Kong Chinese	82.04	17.21

The above Tables suggest that the subjects knew about 90% of words at the 2,000 and 3,000 word level, and between 80 and 85% of words at the 5,000 word level. This indicates that they knew about 4,200 words. This is lower than Gui's finding that Chinese university graduates had a vocabulary size of 6,000 words (Gui, 1987).

Table 2.6 shows the results of the University Word level.

Table 2.6 Scores on the University Word level		
	Mean	**SD**
Mainland Chinese	79.98	14.10
Hong Kong Chinese	74.79	16.70

For some students, the University Word level did not present great difficulties, this is probably because these students had been studying at higher degree levels for many years, their academic vocabulary was growing with the amount of reading they were exposed to. However, this was not the case with all the students. For example, one student who scored 100%, 94% and 90% for the 2,000, 3,000 and 5,000 word level respectively scored only 49% for the University word level test. However, this kind of low score (below 50%) represents only 7% of the students who participated in the study.

2.4 Summary

This chapter reviews different methods of measuring learners' vocabulary size in order to find a method that suits my own purposes of measuring my learners' general vocabulary knowledge. I used Nation's Vocabulary Levels' Test (1990). The results of this study show that students who participated knew approximately 4,200 English words.

Chapter Three

Some Preliminary Data

3.1 Introduction

As mentioned in Chapter Two, I was particularly interested in Lynn's findings that students most frequently annotated non-technical words, which suggests that such words were unfamiliar to them. Annotation could also suggest something different, however. For example, it could be used to mark a word that a student found interesting, rather than unfamiliar. To learn more about students' word knowledge, I decided to use a qualitative method - I conducted a small scale interview with some of the Engineering students.

Like other qualitative research methods, interviewing is a hypothesis generating approach. As stated by Seliger and Shohamy (1989), the ultimate goal of qualitative research is to discover phenomena, such as patterns of second language behaviour, not previously described, and to understand those phenomena from the perspective of participants in the activity. Nunan (1992) gives a clear account of how interviews are used in language research. 'Interviews can be placed on a continuum ranging from unstructured through semi-structured to structured.' Seliger and Shohamy (1989) liken qualitative research procedures to an upside-down pyramid, meaning that the investigation progresses from the general to the specific. Another of their images of qualitative research is that of a spiral where the researcher progresses from general to more specific data collection, but there is also a repetition of the cycles of observation and analysis. Since the procedures for qualitative research are open-ended, there are technically no prescribed procedures, but only general guidelines.

Interviewing is an important procedure in qualitative research. Interviews are personalised and therefore permit a level of flexibility that cannot be obtained by other procedures. The interviewer can probe for information and obtain data that often have not been foreseen. Much of the information obtained during an open/unstructured interview is incidental and reveals itself as the interview proceeds. According to Seliger and Shohamy (1989) and Nunan (1992), interviews can be differentiated by their degree of explicitness and structure, ranging from very open interviews to very structured ones. 'Open' interviews provide the interviewee with broad freedom of expression and greater depth, and one question leads to another without a pre-planned agenda of what will be asked. There is usually a topic for the interview but, by allowing the respondent maximum freedom of expression, ample and often unexpected information emerges. This kind of information would probably not be obtained if the interview was more structured. Other types of interviews are 'semi-open', 'semi-structured', and 'structured' interviews. As the names suggest they are increasingly structured and thereby increasingly limit the extent of free response.

However, the method of interviews raises some problems of data control. Being a participant in a conversation which later becomes the object of analysis makes the process of interviewing more complicated. Although the goal of most interviews is to observe the participants without distorting the process, two risks of distortion develop because of the interviewer's participatory role. The first risk develops at the time of the interview, when the interviewer's role influences its development. The second risk develops at the time of analysis: what is the role of interpretations and knowledge gained from participatory experience in the interview? Ideally one should not try to influence the interview towards any particular conclusion, and one should try to apply uniform standards of analysis to the interview - regardless of one's own role (or lack of it) in it. But it may be difficult to have a completely unbiased or objective interview with a uniform standard to the analysis, because social realities are sometimes constructed at least in part from individual efforts to make sense, therefore one person's definition of what is going on may differ from another's, making a search for consensual (and uniform) definitions a fruitless task.

Like all other objective interviews, the interview described in the following section tried to minimize the interviewer's sometimes distorting effect. However, this effect was not totally eliminated. Because of this, the data presented here is preliminary rather than conclusive. Two more factors also affected the status of the findings: only eight students took part in the interview, the number therefore, is very small; and secondly, because of the time constraint, the concept of subtechnical and technical vocabulary was not an easy one to convey to the participants. Therefore although every effort was made to give the participants some idea of these two types of vocabulary, their response still depended on their own understanding of these words.

Despite these disadvantages however, the data from the interview was interesting because the issues it raised about subtechnical and technical vocabulary were not in line with the claims of some prior researchers.

3.2 Interview

3.2.1 Redefining the subjects

As stated in Chapter One, I was interested in the vocabulary knowledge of students from Southeast Asia: China, Hong Kong, Taiwan, Thailand and Malaysia. The subjects of my first study (in Chapter Two) were not a very homogeneous group, however. Most of the students from China, for example, were studying for different research courses in different subject areas. This would have made the collection of further data rather difficult. I soon realized that my original plan to investigate the vocabulary learning needs of all Southeast Asian students was an undertaking that was too vast for the resources at my disposal, so I decided to restrict my investigation by narrowing down the research subjects to a more homogeneous group enrolled on a single course at the University of Warwick. I selected students from Hong Kong, Thailand, Taiwan and Malaysia who were taking the same Msc. Engineering program at the University of Warwick, because their reading materials were readily available.

Having redefined my research subjects, it is necessary to give more information about these students, and about the Msc. Engineering programme on which they were enrolled.

The Msc Engineering programme at Warwick University has three major course components: **Information Technology for Manufacture (IT), Engineering Business Management (EBM)** and **Manufacturing Systems Engineering (MSE).** It is a one-year taught course, which is jointly designed and implemented by the university and industry to provide advanced education and training for Science and Engineering graduates who wish to find work in the manufacturing industries. The programme is taught by a series of modules. Usually twenty six modules are available each year, although there are sometimes variations. Students make their selections taking into account their previous experience, current needs and future career prospective in consultation with the Director of Studies. Normally a module contains forty to fifty hours of lectures, tutorials and design periods. Each full module is equal to one unit of credit. Students are required to take modules equivalent to at least twelve units of credit together with an industrially orientated project resulting in a dissertation. Students are assessed by post module assignments and examinations, plus a dissertation on the project.

3.2.2 Procedure

At the time when the interviews were being conducted the students were very busy with assignments and other course demands; they needed to attend lectures five days a week, and also had some evening tutorials. For this reason, I tried to take as little of their time as possible. Each interview was limited to thirty minutes or so.

If I had had more time, I would have been able to probe the students' understanding of 'technical' and 'subtechnical' vocabulary more deeply, and I could have discussed these concepts with them in more detail, and given them more examples. In the time available, I was constrained to do the following: I explained that 'technical words' denoted the key technical concepts that they studied each day in their course; as opposed to 'other vocabulary'. The students were not asked explicitly about subtechnical vocabulary. So in this case, 'other vocabulary' included both subtechnical and General English words. However, as Chapter Two suggests, students at this level already know about 4,000 to 5,000 English words. 'Other vocabulary' was therefore likely to be interpreted as subtechnical vocabulary.

Eight students were interviewed. All of them were from Hong Kong, and were taking the same Msc. Engineering course. Of the eight students, six of them, Samuel, Raymond, Eric, Leo, Dicken and Michael, had studied a Manufacturing course either in a polytechnic or a similar college in Hong Kong. To quote the students, they had 'an Engineering background'. The other two students, Frank and Tony, had followed a different kind of course before coming to Warwick University. Frank had studied Textiles, Tony had studied Management Science.

An open interview procedure was used, because it was considered the best way to gather a pool of information which could be grouped, analysed and summarised. In the actual interview, I asked questions and the subjects responded in a one-to-one face-to-face situation. All the interviews were tape-recorded (see Appendix One for the transcription of the data).

3.2.3 Analysis

The analysis of qualitative data is a complex task. Categories were not identified in advance but emerged from the data themselves.

Below is a description of how the data were analysed.

- Transcriptions were made for each interview.
- The transcriptions were carefully reviewed and compared to see the commonalities, regularities or patterns occurring across the various interviews. Sections of data not related to language development were excluded (for example, Eric's comments on how difficult it was to find suitable accommodation on campus).
- From the remaining data two recurring patterns of responses were identified which were relevant to my investigation of vocabulary difficulties.

The two main issues relating to vocabulary learning that emerged in the interview were that modules related to first degree study were found easier than modules that had never been studied before; and technical vocabulary was reported to cause problems for the students.

Two categories of modules on the Msc. Engineering programme seemed to present difficulties for the subjects in my study who had followed different courses. These were modules that were specifically related to Engineering and modules that were specifically related to Business Management.

The Engineering category consisted of:

- Metallic Materials Selection (MMS)
- Automation and Robotics (AR)
- Programming Language (PL)
- Introduction to Manufacturing Systems (IMS)

and the Business Management category consisted of:

- Production Planning and Control Systems (PPCS)
- Human Factors in Industry (HFI)
- Financial Analysis and Control Systems (FACS)
- Manufacturing Strategies (MS)

A number of informants reported that one or the other category posed particular problems. For example Leo complained:

> I think it's difficult in the Business subject, ... due to haven't studied before ... if I haven't know the subject before I haven't any basic ideas, so sometimes I cannot hear what he's talking about.

Leo meant that subjects he had not studied before were difficult because he lacked background information.

Dicken had similar difficulty with modules in the Business category:

> For some topic it is quite difficult to learn some new vocabulary, ... maybe in some new topic, maybe Human Factors or something is not quite Engineering, for example, Finance, something else, other than, I'll be, even say, Finance or Human Factors quite a lot of new words.

The point Dicken was trying to make here was that he encountered a lot of new vocabulary in a totally new subject which was difficult for him to acquire.

Eric seemed to have the same problem as Dicken:

> I think generally it's OK but maybe I also have some vocabulary when I reading these references, especially for the Business reference, the module about the Business, ...

Michael also considered the Business Management modules more difficult than modules related to Engineering, because he had studied Engineering before:

> Actually the Management subjects will be more difficult, ... because most of the subject I had do more or less I study some, for a period of study ....

All the six students emphasized the importance of previous background knowledge. They felt that new subjects were difficult because they had not had a chance to study them before. Most of them did not consider those Engineering subjects such as 'Metallic Materials Selection' problematic because they had studied a similar course already in Hong Kong.

Both Frank who had studied Textiles and Tony who had studied Management Science felt that subjects such as 'Manufacturing Strategy', 'Human Factors in Industry', 'Information Systems Strategy', and 'Strategic Marketing' did not present any difficulties. This was Tony's view:

> ... Financial Analysis, I did some Financial before, so when I come to this module, I find it is quite easy, I don't need to read a lot of books, I just do the work.

On the other hand, both Frank and Tony found the Engineering subjects rather difficult.

Table 3.1 summarizes the students' perception of the difficulty of modules. The modules marked with a tick (√) are those modules that the subjects identified as being most difficult.

Table 3.1 Students' perception of difficult modules

	Samual	Raymond	Eric	Leo	Dicken	Michael	Frank	Tony
MMS		√						
AR								√
PL								√
IMS							√	
PPCS	√			√				
HFI					√			
FACS	√				√			
MS	√							
BM*			√			√		

BM* indicates the area of Business Management, but the specific module was not identified by the subject.

The second issue related to vocabulary learning was that students noted difficulties with technical words as opposed to other types of vocabulary. All the eight students identified vocabulary as a problem, although Samuel and Dicken did not specify what type of vocabulary was problematic for them.

Leo, Eric and Michael found that the specialist words in the modules related to Business Management caused the greatest difficulties, and two students, Tony and Michael, specifically mentioned that technical terms were a source of difficulty. For Tony, technical terms were a big problem. Below is an excerpt from what he said in the interview:

> ... because my background is Management Science, and I'm doing IT, which is a bit on the Engineering, something I uh ..., computer language, you know, the programming, so I find a little bit hard, problem. ... some of the technical like programming, computer programming as difficult ... It's mainly the terms, I think, because if I don't understand the terms, ... I can't do anything ...

Raymond also felt that both technical terminology and technical concepts were difficult for him:

> Metallic Materials Selection is a little bit different, because it's a very very technical side, yeah, it's very very technical and talking about a lot of Metallic theories ... The subject is quite difficult, the theories are also difficult, ... many terms. You need to have a clear idea about each terms ...

He concluded that he had difficulties with 'some technical terms in this module' ('Metallic Materials Selection').

All the subjects in this study had difficulties with the subject disciplines they had not encountered before and had fewer difficulties with familiar subjects. This was to be expected as the successful comprehension of text depends on the activation of appropriate content schemata or background knowledge. The students particularly mentioned that they had difficulties with listening when they attended lectures. They also confessed to difficulties in reading.

As mentioned earlier, one of the patterns emerging from the interview was that these students found technical vocabulary more difficult than non-technical words. This finding conflicts with Lynn's findings (1976) in his studies of students' annotated textbooks. Most students in his study found that non-technical words presented difficulties. Lynn called these non-technical words 'textbook English' words. Nation (1990) referred to these words as academic words. In this thesis I will call this type of word subtechnical vocabulary.

The finding that most students in my interview found technical vocabulary difficult also conflicts with received opinion in ESP. Because it is commonly believed that technical vocabulary is automatically learned during the subject course, technical vocabulary is not regarded as presenting great difficulties to ESP learners.

3.3 Summary of findings

This chapter gives a brief account of a series of interviews conducted in order to gain information about Southeast Asian Engineering students' vocabulary knowledge. The results of the interviews seem to suggest that some Hong Kong Engineering students consider technical vocabulary problematic. These results conflict with Lynn's findings that students did not find technical words difficult, although he used different methods in reaching this finding. The results of the interview also seem to challenge what is commonly believed by most ESP practitioners.

As mentioned in 3.2, there were constraints on the interviewing procedure. Although some effort was made to explain the concepts of 'technical' and 'subtechnical' vocabulary, I had no means of knowing exactly what the interviewees' understanding was. Secondly, the interview data cannot be interpreted in an absolute, definite manner. For this reason the data obtained here is considered tentative and hypothesis generating rather than conclusive.

Despite these limitations, the data elicited from the interview was interesting, and prompted me to collect more empirical evidence in order to investigate whether technical vocabulary did indeed cause the students great problems than subtechnical vocabulary. Chapter Four describes the development of an Engineering corpus, with which the main studies were to be carried out.

Chapter Four

The Development of the Engineering Corpora

4.1 Introduction

Chapter Two reports on the results of a study to determine the lexical knowledge of students from Southeast Asia. The aim of the study was to investigate the overall vocabulary knowledge of these students. The results showed that they were familiar with approximately 4,200 English words.

The next question that needed to be answered was the number of words such students actually need for their academic study purposes. This can be considered their vocabulary goal. Nation (1990: 11) proposed three ways to assess foreign language learners' vocabulary goals. One way would be to look at the vocabulary of native speakers of English and consider that as a goal for foreign language learners. The second way he proposed was to look at the results of frequency counts; and the third way he suggested was to rely on the practical experience of language teachers and researchers and decide how much vocabulary is needed for particular activities.

His first proposal to use native speakers' vocabulary as a goal is not very reliable. As discussed in Chapter Two, there is an enormous variation in the estimates of native speakers' vocabulary knowledge. For example, studies by Diller (1978) estimated a vocabulary size of 216,000 words for native speakers; whereas studies by Kirkpatrick (1907) indicated a vocabulary size of only 20,120 words. The estimate by Seashore and Eckerson (1946) reached yet another different result (of

150,000 words). These widely varying estimates make it difficult to use native speakers' vocabulary size as a goal for foreign language learners.

Nation's second proposed method of assessing a foreign language learner's vocabulary goal was to rely on existent frequency counts. This method could also be problematic. As stated in Chapter Two, frequency lists derived from corpora of General English do not reflect the vocabulary demands of subject specialisms.

Nation's third proposal was to rely on language teachers' experience. The weakness of this is self-evident. This way of assessing learners' needs is open to intuition and personal judgment. Teachers may have fixed ideas about what is appropriate to teach and what is not appropriate to teach, and these ideas may be influenced not by observation of needs, but by tradition of course design or pre-existent teaching materials. Also, English teachers may know little of the needs of the subject specialism.

The weaknesses of the above approaches made it difficult for me to use any of them to assess my learners' vocabulary goal. I decided to examine their reading materials. Reading materials here refer to the textbooks that students have to read for their assignments and project work. These materials will be described in more detail in section 4.4. The rationale behind examining these reading materials is that they can reveal naturally-occurring words, structures and patterns that we cannot always intuit.

I was particularly interested in the role of technical and subtechnical vocabulary in the texts the students were required to read. Chapter One examines definitions of both technical and subtechnical vocabulary. Of the two types of vocabulary, subtechnical vocabulary seems to be the more complex. Some ESP practitioners such as Higgins (1985) and Johns and Dudley-Evans (1985) define subtechnical vocabulary in terms of the classroom teaching situation; other writers such as McCarthy (1990, 1991) and Francis (1986) describe subtechnical words according to their discourse organizing function. A third group of writers identify subtechnical vocabulary by using a more objective corpus-based approach. This approach does not involve intuition and personal judgment. As mentioned above, this is the approach that I will adopt in this thesis to further investigate technical and subtechnical vocabulary. The criteria described in Chapter One for identifying technical and subtechnical words will be applied to the analysis of my corpus.

I shall refer to the original corpus on which Study One was based as the 'small corpus' and I shall refer to the larger corpus from which a list of words was drawn for Study Two as the 'large corpus'.

This chapter is divided into three parts: section 4.2 clarifies some terminology used in this chapter and hereafter; section 4.4 describes the development of the small corpus; and section 4.6 concerns the development of the large corpus.

4.2 Terminology

This section gives my own definition of technical and subtechnical vocabulary in this thesis. It also discusses other terminology used in this chapter and hereafter.

4.2.1 Definitions of technical and subtechnical vocabulary

Chapter One discusses studies and definitions of technical and subtechnical vocabulary. Yang's study seems to be particularly relevant to the present research, because he defines the two types of vocabulary on a corpora basis, which is objective and easy to replicate. As discussed in Chapter One, technical vocabulary in Yang's study is words that have 'a peak frequency of occurrence in one or several fields, but never occur, or have very low frequency of occurrence, in other fields' (Yang, 1986: 98). This definition generates two types of technical vocabulary - the first type would be words that occur with a high frequency in only one discipline and never occur in other disciplines; the second type would be words that occur in several disciplines but peak in one particular discipline, in which this word would be considered technical. The present research adopted Yang's definition. Therefore my technical vocabulary included the two types of words described above.

As discussed in Chapter One, Kennedy and Bolitho (1984) consider general English words that acquire a different meaning within an academic context to be 'subtechnical'. In this research however, these words were identified as 'technical', because they peaked in one particular module. Of course, some words that fit the frequency and range criteria of technical vocabulary might not be technical in meaning. This possibility will be examined in Chapter Five.

My definition of subtechnical vocabulary was as follows: words with an even distribution across modules (i.e. no peak occurrences), not falling within the most frequent 5,000 English words in Thorndike and Lorge (1944). This definition was different from Cowan's original definition of subtechnical vocabulary as words that are 'context independent and occur with high frequency across disciplines' (1974). The limitation of his definition, as discussed in Chapter One, was obvious. It did not for example, take into account the high frequency General English words. My definition of subtechnical vocabulary was also different from Yang's, who considered words with a high frequency across disciplines 'subtechnical'. His definition could include words with a peak occurrence in one particular discipline as 'subtechnical', while in effect these words fit the frequency and range criteria of technical vocabulary.

The cut-off point for technical and subtechnical vocabulary in this research was five, which means all technical and subtechnical words had to have a frequency of five or more than five. All words which had less than five occurrences were excluded from the word lists. The cut-off point of five was a practical decision, which allowed me to include a large number of words for analysis yet exclude low frequency words. It is possible that some words which had four occurrences or less were just as important as those that had five occurrences or more. However, as it was impossible for me to objectively evaluate the importance of the low frequency words in my corpus, I decided to analyse only those words which the students were bound to encounter a number of times in their reading.

4.2.2 Other definitions

Word: As discussed in Chapter One, I adopted Carroll, Davies and Richman's definition of word (1971), which is 'a string of graphic characters bounded left and right by space.' See Chapter One for the advantages and disadvantages of this definition.

Module: A 'module' is one of the units that together make up the complete postgraduate Engineering course. When I am discussing the behaviour of words, I use 'module' to refer to the *texts* which represent the module under analysis.

Text: A 'text' is a continuous piece of written language. In many places, *text* and *discourse* are used interchangeably. They are both used as a general term when discussing language patterns that are above sentence level.

4.3 Text processing

Two pieces of software were used for text processing: one for producing word frequency lists, the other for multi-word item processing. The software used for compiling word frequency lists was prepared in 1993 by Youxin Zheng, a postgraduate studying Engineering at Warwick University. This programme produced frequency lists that show the words which make up the texts in the corpus, together with their frequencies of occurrences. The word frequency list was produced by identifying each word form found in the text, counting identical forms and listing them with their frequencies. This programme also produced the most common totals calculated for word frequency lists, referred to as 'total tokens', and 'total types' (Barnbrook, 1996). In this case, a token is an individual occurrence of any word form, while each type represents a different word form (see Chapter One). As an example, my small corpus contained altogether 208,473 words, or tokens, representing 10,348 types.

The production of the word frequency list involved two main steps: splitting the text into words; and counting identical word forms. The first stage demands the definition of word boundaries in a form that the computer can recognize. In this case, the basic unit being identified and later counted was a string of characters separated by spaces or by other boundary characters. However, as mentioned in Chapter One, there are problems with this definition. Some multi-word units are in practice indivisible in their use in the English language, these units are conventionally written as phrases containing more than one word. Some of these units such as *of course*, are fixed phrases, which cannot be used with this meaning with any intervening words. Phrasal verbs such as *take up* and *put on* are also found to be used in an indivisible way sometimes. However, phrases such as *of course, take up* and *put on* consist of general English words which are included in Thorndike and Lorge's (1944) first 5,000 words, and were eliminated from the corpus for analysis. Multi-word technical terms were identified manually.

Once the text had been broken into appropriate units, the counting process involved a comparison between these units to ensure the count of the same items was incremented - if the sequence of characters was the same, the units were counted under the same item. To make this approach possible, upper and lower cases of the same letter were treated as the same item.

The problems of using this approach for text processing are discussed in Chapter One, among which one common limitation is: by breaking the text into individual word forms it removes the words from their original contexts. The effect of this is that word forms which can have more than one meaning are gathered together and counted as one single word.

The other program used to identify multi-word items was Microconcord (Johns and Scott, 1993). This programme allowed me to search for those multi-word technical terms identified manually. Since the number of this kind of items was not big, it was possible to first of all identify them manually. Microconcord then provides the tool to 'search through electronically-stored texts for all the occurrences of a word or combination of words' (Murison-Bowie, 1993).

4.4 The development of the small corpus

4.4.1 Introduction

Constructing a corpus involves the selection of texts which are representative and can support generalisations against which hypotheses can be tested. To achieve this aim, the first step is to define the whole of which the corpus is going to be a sample. For the purpose of this research, the whole body of text that the corpus was going to represent was the reading materials required for the Msc. Engineering course.

Other decisions concern a suitable size for the samples. Sinclair (1991: 19) warns, 'a corpus which does not reflect the size and shape of the documents from which it is drawn is in danger of being seen as a collection of fragments where only small-scale patterns are accessible.' A corpus made up of whole documents is open to a wider range of linguistic studies than a collection of short samples. The sampling technique of gathering the whole document is also reduced to a minimum, because the researchers do not have to worry about, for example, any differences between different parts of the document.

There are problems with including complete documents, too. As Sinclair (1991: 19) put it:

> The penalties to pay for including whole documents are that in the early stages of gathering, the coverage will not be as good as a collection of small samples and the peculiarities of an individual style or topic may occasionally show through into the generalities.

I decided to include representative samples from every text that was recommended on the Engineering programme. The samples were complete chapters, but not complete books. In some books each chapter deals with a related but separate topic, and this inevitably means that some content specific words important to my subjects were absent from the corpus. My corpus was a sample corpus, but the samples were balanced and representative. They constituted about 10% of the entire volume of recommended reading (the whole of which the corpus was designed to represent).

4.4.2 Selecting samples for the small corpus

As described in Chapter Three, the Msc. Engineering course was composed of several course components. The three main components were: **Information Technology for Manufacture (IT), Engineering Business Management (EBM)** and **Manufacturing Systems Engineering (MSE)**. All these course components were represented in modules. Students had to choose one of these components as their main component study. There were four other subordinate components: **CORE, SELECTIVE, IT/EBM** and **IT/MSE**. The **CORE** component did not specialise in one area, instead it was a mixture of all the three main components. All the students also had to study this component. The optional **SELECTIVE** component covered the three main components too, and students could choose the modules related to the study of their main component. **IT/EBM** was a mixture of **IT** and **EBM**, and **IT/MSE** was a mixture of **IT** and **MSE**, and only students studying relevant components could study them. Usually twenty six modules are available each year, although there are sometimes variations.

Texts for the corpus were chosen from textbooks used in the teaching of the modules. One book was chosen to represent each of the modules. Table 4.1 lists the course components, modules and the respective books named by the Engineering Department in connection with each module.

These books fell into two categories: those that were specially recommended by the subject lecturers, and those that were *not* specially recommended. Twenty four of the books used for the corpus were recommended books. Two other books were also used (see Table 4.1, marked with *) because the recommended books for these two modules overlapped with other modules.

Table 4.1 The Engineering course components, modules and the respective books

Course Component	Module	Book
CORE	Applied Statistical Methods (ASM)	Statistical Concepts and Methods
	Basic Computing (BC)	Computers and Common Sense
	Computer-aided Design and Computer-aided Manufacturing (CAD)	Computer-aided Design and Manufacture
	Financial Analysis and Control Systems (FACS)	Relevance Lost: the Rise and Fall of Management Accounting
	Quality Reliability Maintenance Systems (QRMS)	Practical Reliability Engineering
IT (Information Technology)	Artificial Intelligence (AI)	Artificial Intelligence
	Computer Engineering (CE)	Computer Networks
	Computer Language (CL)	Pascal: User Manual and Report
	Programming Language (PL)	The Complete FORTH
	Automation and Robotics (AR)	An Introduction to Robot Technology
EBM (Engineering Business Management)	Human Factors in Industry (HFII)	Managing Organizations
	Information Systems Strategy (ISS)	Strategic Planning for Information Systems
	Manufacturing Strategy (MS)	World Class Manufacturing
	Strategic Marketing (SM)	Basic Marketing: A Global Managerial Approach
MSE (Manufacturing Systems Engineering)	Industrial Engineering (IE)	Production and Operations Management
	Machine Tools (MT)	Modern Workshop Technology
	Metallic Materials Selection (MMS)	*Steel and its Heat Treatment
	Polymer Materials, Processes and Products (PMPP)	Plastics Engineering
	Logistics and Supply Chain	Manufacturing Planning and

	Management (LSCM)	Control Systems
IT/EBM	Information Technology Fundamentals (ITF) Production Planning and Control Systems (PPCS)	This is IT- A Manager's Guide to Information Technology *Operations Management: Theory and Problems
IT/MSE	Programmable Systems in Measurement and Control (PSMC)	Microcomputers in Engineering and Science
SELECTIVE	Project Planning Management and Control (PPMC) Financial Decision Making (FDM) Simulation of Production Systems (SPS) Introduction to Manufacturing Systems (IMS)	Project Management Demystified Investment Appraisal Simulations of Manufacturing Systems Materials and Processes in Manufacturing

Note: The second column lists names of the modules and their respective abbreviated names, for example, BC stands for the module 'Basic Computing', BM stands for 'Basic Marketing'. The abbreviated names will be used elsewhere in this thesis.

The criterion for the selection of the texts for the corpus was to choose as many recommended books as possible, in order to make the corpus as relevant as possible to the students' reading needs. Samples from the books were chosen from the opening chapters. It was difficult to decide how many pages from each book to include in the corpus, because some books (especially those in the IT component) have pages of diagrams, figures, graphs and charts. Modules in EBM do not include this type of illustration. Faced with this variety of book layout, I made the decision that all introductory chapters should be included in the corpus. If some of these chapters had diagrams and charts, then a following chapter would be included. Words included for each module varied to some extent, but the total number of tokens in the three main course components was similar. This can be seen from Table 4.2. As Table 4.2 shows, the length of the text for each module varied between 3,793 and 13,463 words. Although the length of each text varied, the length of the three main course components was very similar.

Table 4.2 The type/token figure of the small corpus			
Course Component	**Module**	**Token**	**Type**
CORE	QRMS	5,774	1,357
	FACS	4,749	1,220
	CAD	4,571	1,209
	ASM	3,856	1,087
	BC	3,793	1,104
IT	AR	5,697	1,225
	CE	12,286	2,184
	CL	7,026	1,128
	AI	6,803	1,325
	PL	9,017	1,167
EBM	MS	12,615	2,655
	HFII	9,534	2,389
	ISS	13,463	2,317
	SM	7,509	1,498
MSE	MMS	8,558	1,414
	LSCM	9,589	1,608
	PMPP	8,847	1,864
	IE	10,085	1,352
	MT	11,251	1,637
IT/EBM	ITF	8,452	1,965
	PPCS	9,059	2,247
IT/MSE	PSMC	10,150	1,595
SELECTIVE	SPS	4,046	1,043
	PPMC	8,689	1,837
	FDM	4,301	953
	IMS	11,081	2,405

Table 4.3 shows the length of each of the main course components.

Table 4.3 Type/token figure of the three main course components		
Course Component	Token	Type
IT	40,829	7,029
EBM	45,121	8,855
MSE	48,330	7,875

Using the program described in 4.3, the small corpus yielded a total of 10,348 types and 208,473 tokens.

4.5 Technical and subtechnical vocabulary in the small corpus

Section 4.2 describes my criteria for identifying technical and subtechnical words.

One category of my technical vocabulary was words that occurred more than five times with a low range across the corpus. The lowest range was one, which meant that the word occurred in only one module. This was found to be the common distribution of technical words in my corpus: 342 out of 668 technical words occurred with this distributional range. As an example, the word *polymer* occurred 106 times in only one module: 'Polymer Materials, Processes and Products'. Words with a higher frequency were also considered technical if they had a peak distribution in one or more than one module. 'Peak' is defined as the highest frequency point of a word, from a comparative perspective. There was always a big difference between the number of times a word occurred at its most frequent and the number of times it occurred in all the other texts. For example, the word *reliability* occurred in eight modules with a frequency of 154. It peaked in 'Quality Reliability Maintenance Systems' with a frequency of 140, and it occurred only twice in each of the seven other modules. Between the frequency of 2 and 140, there was a big gap. In this case, 140 was considered to be the peak frequency of this word. Another technical term *simulation* occurred forty times across five modules, with a distributional pattern of one, one, two, two, and thirty three. Thirty three was clearly a peak frequency.

In addition to single-word technical terms, I also identified a number of multi-word technical terms on the basis of their distributional patterns. For the purpose of this study, I defined a multi-word technical term as a string of words that always occurred together with a technical distribution pattern. Two categories of multi-word technical terms were identified: In one category the elements that made up the multi-word term were themselves technical, in that they had technical distribution patterns. The majority (95%) of multi-word technical terms identified were identified in this category. For example, *polyvinyl chloride* was considered a multi-word technical term because the two component words were both technical and co-occurred fifteen times in only one text. Similarly the multi-word technical term *polyphenylene oxide* occurred five times, and *shear strength* eleven times, with a distribution of one. Both elements in both terms only occurred together, never separately. In the other category of multi-word technical terms, some of the words in the strings were not technical. For example, although *mean time between failures* co-occurred across a low range of texts, these words did not have a technical distribution pattern separately. Only 5% of multi-word technical terms were of this type.

Technical words were identified according to the criteria outlined above. They were found to display two different distributional patterns, either a very low distributional pattern or a broader distribution which peaked in one module. Table 4.4 shows the low distribution of a sample of technical words in the module 'Polymer Materials, Processes and Products'.

Table 4.4 A sample of technical words from the module 'Polymer Materials, Processes and Products'

	Frequency	Range
polymer	106	1
molecule	21	1
thermoplastics	19	1
polyester	18	1
polyethylene	16	1
polyproplylene	15	1
polystyrene	15	1
polycarbonate	9	1
polyamide	7	1
polyurethane	6	1

The broader distributional pattern is shown in Table 4.5 for a sample of technical words in 'Quality Reliability Maintenance Systems'.

Table 4.5 A sample of technical words from the module 'Quality Reliability Maintenance Systems'

	Frequency	Range	Peak Frequency
reliability	154	8	140
probability	19	4	14
failure rate	17	1	17
maintainability	11	2	10
warranty	10	3	8
hazard	10	1	10
mean time between failures	10	2	9
repairable system	10	1	10
life cycle cost	5	1	5
rate of occurrence of failures	5	1	5

Words such as *reliability, probability, maintainability, warranty* and *mean time between failures* had a higher distribution than one, but their distribution peaked in one module only, and they only occurred once or twice in each of the other modules. For this reason they were classified as technical rather than subtechnical words. It can be assumed that some technical words that were important for the Msc. Engineering students did not occur frequently because the corpus was not big enough. A larger corpus would have enabled patterns to show more clearly.

Having identified technical words using the above criteria, the second category of words to identify from the corpus was the 5,000 most frequent words in the Thorndike and Lorge count (1944). These were considered General English words. Earlier tests (see Chapter Two) had suggested that the Chinese postgraduate students studying in the UK knew approximately 4,200 words. However, I decided to eliminate 5,000 rather than 4,200 general English words for two reasons: In the first place, this would be in line with earlier studies on the development of subtechnical vocabulary. Campion and Elley (1971) based their decision to exclude frequent words from their academic word list on findings that university students are familiar with the most frequent 5,000 English words. Praninskas (1972) followed the same approach in developing the American University List. Secondly, this research focuses on the study of subtechnical and technical vocabulary, and the fewer general English words included for analysis, the better.

The General English words were eliminated after technical words had been identified in the corpus; this was to avoid removing words such as *bit, bus, gate* and *memory*. Words of this kind were considered technical because they peaked in a particular module. They are discussed in Chapter Five together with other technical words identified in the corpus. The General English words eliminated from the corpus made up 4,971 types and 170,228 tokens.

Having identified in the corpus technical words and words that were frequent in General English texts, I was left with two remaining types of words: low frequency items (frequency of less than five) that occurred in only a small proportion of samples, and relatively high frequency words that occurred across a relatively wide range of samples. Low frequency words were ignored for the purposes of this research, but the second category constituted my subtechnical vocabulary. My cut-off point for subtechnical vocabulary was five occurrences, which means that words with a frequency of less than five were not counted.

346 types were identified as subtechnical in my small corpus, this accounted for 7,777 running words which made up 4% of the whole corpus. This figure is much lower than Inman's (1978) (See Chapter One) for subtechnical words. This is because my subtechnical words did not include General English words, which constituted approximately 82% of the running words in my corpus. Compared with Farrell's (1990) subtechnical words (see Chapter One), the number of subtechnical words identified in my corpus was also smaller. This may be because Farrell included low frequency words in his subtechnical word list. My subtechnical vocabulary did not include General English words as Inman's did, and did not include low frequency words as Farrell's did. The features of the subtechnical vocabulary in my corpus are discussed in Chapter Five.

4.6 The development of the large corpus

As described in 4.4, the small corpus was a collection of introductory chapters of 26 textbooks used for the teaching of the 26 Engineering modules. However, some samples included in the small corpus were not representative of the textbooks as a whole.

The following examples show the type of discourse of these samples.

> WHY WAS THE COMPUTER INVENTED?
>
> Many of the routine activities in today's society are performed by computers. For example, when we go on holiday our 'plane seats are often reserved by computers; the traffic in some major cities is, to a degree, controlled by computers; the egg which you might have had for breakfast may have been laid (no, not by a computer!) by a chicken whose life history is on record on a computer file; many of the bills we pay (rates, gas, electricity, telephone, insurance, etc.) are calculated and printed by a computer. Why? And how?
>
> It was outlined in the Introduction that there are three essential concepts that we need to examine in order to be able to think sensibly and talk intelligently about computing - What are computers? What can they do? How can we communicate with them? But first of all, what do the terms computer and computing mean?
>
> Obviously computing has something to do with reckoning or calculation, but man has been using his brain to do just that for centuries. The Egyptians built the pyramids;

whoever built Stonehenge left a calendar which can still accurately predict eclipses; the Romans designed and built long straight roads, aqueducts and heating systems; early explorers navigated the globe, and even radio and television were invented - all without computers. What is so special about them that we need computers today? It cannot simply be because they are calculating devices. We have many forms of such devices - the abacus (still used in the Far East), pocket and desk calculators, even supermarket check-out tills - all of which are cheaper and easier to use than computers. So, why was the computer invented?
(**Computers and Common Sense**, Hunt and Shelley 1988: 1)

Here is another example.

Marketing affects almost every aspect of your daily life.

When it's time to roll out of bed in the morning, does your General Electric alarm wake you with a buzzer - or playing your favorite radio station? Is the station playing rock, classical, or country music? Will you slip into your Levi's, your shirt from L. L. Bean, and your Hikes, or does the day call for your Brooks Brothers suit? Will breakfast be Kellogg's corn flakes - made with corn from America's heartland - or some "extra large" eggs and Hormel bacon cooked in a Panasonic microwave imported from Japan? Will it be Maxwell House coffee - grown in Colombia - or some Minute Maid orange juice? Maybe you're late and plan to get an Egg McMuffin at the McDonald's drive-thru. When you leave home, will it be in a Toyota, on a Huffy bike, or on the bus that the city bought from General Motors?

When you think about it, you can't get very far into a day without bumping into marketing - and what the whole marketing system does for you. It affects every aspect of our lives - often in ways we don't even consider.

In this chapter, you'll see what marketing is all about and why it's important to you. We'll also explore how marketing fits into our whole economic system
(**Basic Marketing: A Managerial Approach**, McCarthy and Perreault 1993: 1)

The last example:

I.1 HISTORICAL BACKGROUND
In the nineteenth century, the industrial revolution considerably enhanced man's physical power. In the present century, a second industrial revolution is taking place, with computers offering an enhancement of man's mental capabilities. It is quite unthinkable these days to undertake a major engineering project without the use of a computer. Since the late 1950s, the applications of computers and computing techniques in engineering disciplines of all types have increased dramatically because computers have become larger in memory capacity and faster in processing speed which mean that more complex problems can be tackled and that more calculations can be performed in a given time. More importantly, with the advent of microelectronics such as Very Large Scale Integration (VLSI) Technology, computer hardware is gradually becoming cheaper and cheaper every day and it is now within the financial reach of most industrial companies that wish to take advantage of its capabilities. Also due to VLSI, computer hardware is getting smaller in size. Because of this size reduction, its applications are increasingly being spread to other areas of industry which formerly could not use the traditional computers owing to their cumbersome size. As a result of these developments in computer science, computer-aided design and computer-aided manufacture (CAD/CAM) was conceived and is rapidly gaining acceptance in engineering industries for their ability to create major increases in productivity.
(**Computer-aided Design and Manufacture**, Groover and Zimmers 1984: 1)

The above three examples indicate that some of the introductory chapters are in some respects atypical for the purposes of studying technical and subtechnical vocabulary. Rather than introducing key technical terminology, these introductory chapters perform the 'social action' of acculturating students into the mode of enquiry of a new subject. The representativeness of the corpus could be improved by including more texts from later chapters in the same textbooks.

Technical vocabulary in my small corpus was identified as words with either low range, or a wider range with a peak distribution. A fairly large number of words in the small corpus displayed a wider distributional pattern. These words were not prototypical technical words, even though I classed them as technical because their distribution peaked in one text. The differences between technical and subtechnical words are less clear cut in the way both types of words have a similar distribution pattern. As discussed earlier, there are various reasons for the wider distribution of technical words. With a larger corpus, these reasons could be examined more fully, and a clearer and more satisfactory picture of the distinction between technical and subtechnical words could be attained. It would be interesting, for example, to see whether the distribution pattern of certain words (such as *hazard*) alter within a larger corpus. I therefore decided to enlarge the corpus before conducting Study Two because a larger corpus would enable me to more fully examine the distribution patterns of technical and subtechnical words.

The large corpus was still derived from twenty six sources for the twenty six modules. The same books were used for this purpose (see Table 4.1), but I added an extra chapter (or chapters) from each book. The length of texts this time varied between 8,452 and 16,690 running words.

Using the same software described in 4.3, the large corpus was found to contain 149,89 types, totalling 536,051 tokens.

4.7 Technical and subtechnical vocabulary in the large corpus

Section 4.5 describes the criteria for identifying technical and subtechnical words in the small corpus. The same criteria were followed when I identified these two categories of words in the large corpus.

Technical words were those with a high frequency and a low distribution. Some words with a wider distribution were also identified as technical if they displayed a peak frequency in one module. The frequency cut-off point for technical vocabulary was still five as in the small corpus, which means words that occurred less than five times were not considered in my analysis. With these criteria, altogether 1,358 technical words were identified, including both single-word technical items, and multi-word items. 623 technical words had a distribution of one: this was the commonest distribution for technical words.

As in the small corpus modules such as 'Polymer Materials, Processes and Products' still had a larger number of technical words with a distribution of one. For example, in the small corpus sixty one technical words were identified in this module, of which forty seven words had a distribution of one, while in the large corpus, 111 were identified and seventy two of them had a distribution of one. Twelve modules in the small corpus contained technical words that display the prototypical pattern of high frequency and a range of one. Technical words from these twelve modules showed the same distributional pattern in the large corpus. This is shown in Table 4.6.

Table 4.6 Increase of technical words in the large corpus

Module	**Total number of technical words**	**Number of technical words with a range of one**
	small corpus/large corpus	small corpus/large corpus
Polymer Materials, Processes and Products	60/111	44/72
Computer Language	56/93	37/46
Metallic Materials Selection	54/107	40/40
Computer Engineering	53/140	32/72
Machine Tools	46/85	32/36
Programmable Systems in Measurement and Control	31/56	10/12
Artificial Intelligence	28/51	18/26
Programming Language	27/68	11/23
Industrial Engineering	24/40	12/14
Manufacturing Strategy	23/45	10/17
Information Systems Strategy	21/42	13/13
Quality Reliability Maintenance Systems	18/58	10/36

Most modules displayed a significant increase of technical words, almost twice the amount was identified in the large corpus. In some modules there was a large increase in the number of technical words with a range of one. These modules were ‘Polymer Materials, Processes and Products’, ‘Computer Engineering’, ‘Programming Language’ and ‘Quality Reliability Maintenance Systems’. In other modules such as ‘Computer Language’, ‘Machine Tools’, ‘Programmable Systems in Measurement and Control’, ‘Artificial Intelligence’, ‘Industrial Engineering’ and ‘Manufacturing Strategy’, there was a fairly moderate increase.

Another fourteen modules that had fewer technical words in the small corpus also showed an increase in numbers. This can be seen from Table 4.7.

Table 4.7 Increase of technical words in the large corpus

Module	**Total number of technical words**	**Number of technical words with a range of one**
	small corpus/large corpus	small corpus/large corpus
Introduction to Manufacturing Systems	30/110	5/33
Logistics and Supply Chain Management	20/53	7/26
Computer-aided Design and Computer-aided Manufacturing	19/66	7/20
Financial Analysis and Control Systems	19/37	3/9
Production Planning and Control Systems	18/48	5/12
Automation and Robotics	17/77	7/43
Strategic Marketing	17/55	9/25
Applied Statistical Method	14/39	3/17
Financial Decision Making	14/31	6/11
Information Technology Fundamentals	14/53	3/10
Basic Computing	12/34	4/7
Human Factors in Industry	12/34	7/15
Project Planning Management and Control	11/26	5/12
Simulation of Production Systems	10/33	0/12

All of the modules had more technical words in the large corpus, except for two: 'Metallic Materials Selection' and 'Information Systems Strategy', which had the same amount of technical words as in the small corpus. One of the reasons could be that all the important technical words in these books were already introduced to the readers in the first chapters included in the small corpus.

I used the same definition of subtechnical words as in the small corpus: subtechnical words were words occurring with high frequency across modules, excluding the most frequent 5,000 English words. The cut-off point for a subtechnical item was also five.

Following the same procedure, I identified subtechnical words after technical and General English words had been identified. I used the same Thorndike and Lorge count (1944) for the elimination of the 5,000 most frequent words. Altogether 4,996 types making up 559,279 tokens were classed as General English words and were later eliminated, which accounted for approximately 84% of the total number of tokens in the large corpus.

Two types of words remained in the corpus after technical and General English words were removed. The first type was words with a high frequency and wide range, these words were the subtechnical vocabulary. The second category included words with a frequency of less than five. I disregarded these low frequency words in my analysis of the corpus and in subsequent studies, because they did not occur in the corpus with significant frequency to provide adequate information regarding their patterns of distribution and use.

Altogether 1,210 types were identified as subtechnical vocabulary in the large corpus, this accounted for 43,956 tokens and made up about 8.2% of the large corpus. This figure is still lower than Inman's (1978), because Inman did not eliminate General words from her subtechnical word list. This figure is also lower than Farrell's (1990), because my subtechnical vocabulary did not include low frequency words as his did (see Chapter One).

For the list of technical and subtechnical vocabulary, see Appendices Two and Three. The function and patterns of use of the subtechnical vocabulary in my corpus will be discussed in detail in Chapter Five.

4.8 Summary

This chapter describes the development of two Engineering corpora: the small corpus and the large corpus. The large corpus produced more data and enabled the examination of the distribution pattern of technical and subtechnical words to be more satisfactory. It also provided a wider choice of technical words to use in later experimental studies because most modules contained more technical words with a distribution of one.

Technical and subtechnical vocabulary in both corpora were identified on the basis of their frequency and range patterns. A word was identified as technical if it occurred five or more times in the corpus with a range of one, or a peak distribution. A word was classed as subtechnical if it occurred five or more times in the corpus, distributed across a wide range of modules. Technical and subtechnical words from the small corpus were used as sample words for the first study, whereas words from the large corpus were used as sample words for the second study of this research. Chapter Six describes these two studies.

Chapter Five

The Linguistic Analysis of the Large Corpus

Chapter Four describes the development of the two Engineering corpora. I identified two types of ESP vocabulary: technical and subtechnical vocabulary. This chapter concerns the analysis of these two types of words; it looks at the behaviour of these words in context in the corpus. The corpus referred to in the following analysis is the *large* corpus, the components of which are described in Chapter Four.

This chapter is divided into two parts: section 5.1 concerns the analysis of technical words, and section 5.2 concerns the analysis of subtechnical words.

5.1 Technical words

As described in Chapter Four, technical words were identified on the basis of their frequency and range patterns. A technical word in my corpus occurred with a high frequency and relatively low range. However, a word with a wider range was also considered technical if it had a peak distribution in one or more than one module. Using these criteria, I identified altogether 1,358 technical words, of which 623 had a range of one. Some technical words, however, occurred in more than one module, for example, *central processing unit* occurred in four IT modules. Sections 5.1.1 and 5.1.2 discuss these two types of technical words: technical words with a range of one and technical words with a wider range.

5.1.1 Technical words with a range of one

The number of technical words with a range of one varied to a great extent across modules. In some modules, a large number of technical words with a range of one could be identified: most of the IT modules and some of the MSE modules had a very large number of technical words with a range of one. For example, in the module 'Polymer Materials, Processes and Products', 64% of the total number of technical words had a range of one. In contrast, modules in the EBM component displayed a different characteristic, very few technical words with a range of one could be identified. Some of the MSE modules displayed similar characteristics to the EBM modules, while others did not. In this section technical words are examined on a module basis and modules are grouped (administratively) on the basis of the course components.

5.1.1.1 IT (Information Technology)

This component was represented by five modules: 'Automation and Robotics', 'Computer Engineering', 'Computer Language', 'Artificial Intelligence' and 'Programming Language'. Altogether 415 technical words were identified in the five modules, and 209 of them had a range of one, constituting about 50% of the total number. Table 5.1 shows the details of technical words in each module:

Table 5.1 Technical words identified in the IT component

Module	Number of technical words	Number of technical words with a range of one	Percentage of technical words with a range of one
Automation and Robotics	75	40	53
Computer Engineering	137	71	52
Computer Language	89	45	51
Artificial Intelligence	51	29	57
Programming Language	63	24	37
Total	**415**	**209**	

Technical words identified in the IT component tended to reveal the topic or the theme of modules, for example, *actuator* occurred thirty three times, *servocontrol* occurred nineteen times, and *articulated mechanical system* occurred nine times in 'Automation and Robotics'; In 'Computer Engineering', *integrated services digital network* occurred 128 times, *protocol* occurred sixty times, and *circuit switching* occurred eighteen times. Technical words in 'Computer Language' and 'Artificial Intelligence' were found to have the same characteristics - they revealed the topic or the theme of the particular modules they were identified.

As can be seen from Table 5.1, only 37% of the technical words in 'Programming Language' occurred with a range of one. This was lower compared to the other four modules. An example of a technical word with a range of one was *actuator* which occurred thirty three times. Other words such as *terminal*, *programmer* and *quotient* had a range of ten, seven and three respectively, and these kinds of distribution pattern were found to be much more common in this module. These words represented more general concepts than *FORTH operation* and could be used in other computer modules. The fact that there were more technical words representing general concepts in the module **Programming Language** suggested that this module was considerably less topic specific than the other four modules. Perhaps technical words particular to the topic or theme of this module were introduced in the later part of this textbook rather than in the introductory chapters.

The general range pattern of technical words in the IT component was that half of the technical words identified had a range of one. Only one module had a lower percentage of technical words with a range of one.

5.1.1.2 EBM (Engineering Business Management)

This component was represented by four modules: 'Manufacturing Strategy', 'Human Factors in Industry', 'Information Systems Strategy' and 'Strategic Marketing'. Altogether 172 technical words were identified, of which sixty six had a range of one, representing 38% of the total number of technical words (Table 5.2 shows the details). This percentage is lower than that for the IT component.

Table 5.2 Technical words identified in the EBM component

Module	Number of technical words	Number of technical words with a range of one	Percentage of technical words with a range of one
Manufacturing Strategy	43	17	40
Human Factors in Industry	34	13	38
Information Systems Strategy	42	11	26
Strategic Marketing	53	25	47
Total	**172**	**66**	

Compared to the IT component, EBM had fewer technical words and, of these, fewer with a range of one in each module. Most words had a range of more than one which suggested that these words were likely to appear in more than one specialism. For example, *retailer* was identified as a technical word in the module 'Human Factors in Industry', but it occurred in three other modules: 'Information Systems Strategy', 'Strategic Marketing' and 'Financial Analysis and Control Systems'. Other typical words and multi-word units in the EBM component were: *organizational behaviour*, *marketing strategy* and *recruitment*, which all had a range of more than one. These words were less technical in the sense that they did not belong to only one specific discipline.

5.1.1.3 MSE (Manufacturing Systems Engineering)

Represented by five modules: 'Metallic Materials Selection', 'Logistics and Supply Chain Management', 'Polymer Materials, Processes and Products', 'Industrial Engineering' and 'Machine Tools', this component had a different pattern from either IT or EBM. 386 technical words were identified, of which 175 had a range of one, representing about 44% of the total number of technical words. Table 5.3 summarizes this.

Table 5.3 Technical words identified in the MSE component

Module	Number of technical words	Number of technical words with a range of one	Percentage of technical words with a range of one
Metallic Materials Selection	101	30	30
Polymer Materials, Processes and Products	109	71	64
Machine Tools	85	34	40
Logistics and Supply Chain Management	51	25	49
Industrial Engineering	40	15	38
Total	**386**	**175**	

'Polymer Materials, Processes and Products' had the highest percentage, while 'Metallic Materials Selection' had the lowest percentage of technical words with a range of one. This suggests that the former had more topic specific words than the latter.

The number of technical words in MSE was smaller than in the IT component (which had 415 technical words in total) and higher than the EBM component (which had 172 in total). Taking the type/token figure of these components (see Table 4.3 in Chapter four) into consideration, we can tell that IT had the highest number of technical words in total. MSE also had fewer technical words with a range of one than the IT component, but had more technical words than the EBM component. Table 5.4 summarizes this.

Table 5.4 Technical words in the three components			
Component	**Number of technical words**	**Percentage of technical words with a range of one**	**Number of tokens**
IT	415	50	40,829
EBM	172	38	45,121
MSE	386	44	48,330

This could suggest that the technical vocabulary in the IT component was introduced in the early chapters of the textbooks, while the technical vocabulary in the EBM component was introduced later.

5.1.1.4 CORE

The CORE component was represented by a mixture of modules. It included modules from all the three main course components: 'Basic Computing' and 'Computer-aided Design and Computer-aided Manufacturing' for IT; 'Financial Analysis and Control Systems' for EBM; and 'Quality Reliability Maintenance Systems' for MSE. Besides these modules, this component included a module on statistics and statistical design called 'Applied Statistical Methods'.

220 technical words were identified, of which 87 had a range of one, representing about 38% of the total number of technical words. Table 5.5 shows the details of the CORE component. The first three modules had a smaller percentage of technical words with a range of one than the MSE related module. The two IT modules: 'Basic Computing' and 'Computer-aided Design and Computer-aided Manufacturing' had fewer technical words with a range of one. This was different

from modules in the IT component because most of them had a large number of technical words with a range of one. 'Basic Computing' and 'Computer-aided Design and Computer-aided Manufacturing' were more basic modules than modules in the IT component. The basic computing information presented in these two modules is needed by any kind of Engineer, and is therefore referred to in other modules that students study on the Engineering programme. 'Financial Analysis and Control Systems' was typical of modules in the EBM component in that few technical words had a range of one. It offers the same kind of basic information applicable to all types of Engineering situation.

Table 5.5 Technical words identified in the CORE component

Module	**Number of technical words**	**Number of technical words with a range of one**	**Percentage of technical words with a range of one**
Basic Computing	35	7	20
Computer-aided Design and Computer-aided Manufacturing	64	20	31
Financial Analysis and Control Systems	29	9	31
Quality Reliability Maintenance Systems	57	35	61
Applied Statistical Methods	35	16	46
Total	**220**	**87**	

'Quality Reliability Maintenance Systems' displayed similar characteristics to 'Polymer Materials, Processes and Products' in the MSE component: they both had more than 60% of technical words with a range of one. Although 'Applied Statistical Methods' had a higher percentage of technical words, it was similar to the first two modules in Table 5.5 in that the information in these modules were needed by any kind of Engineer and therefore referred to in other modules, too.

5.1.1.5 IT/EBM

Two modules represented this component: 'Information Technology Fundamentals' and 'Production Planning and Control Systems'. The first one was an IT module whereas the second one was an EBM module. Altogether 100 technical words were identified and 21 of them had a range of one.

Table 5.6 Technical words identified in the IT/EBM component

Module	Number of technical words	Number of technical words with a range of one	Percentage of technical words with a range of one
Information Technology Fundamentals	52	9	17
Production Planning and Control Systems	48	12	25
Total	**100**	**21**	

As in the two IT modules in the CORE component, few technical words with a range of one could be identified in 'Information Technology Fundamentals': the number of technical words with a range of one only constituted 17% of the total number of technical words. 'Production Planning and Control Systems' also displayed a typical feature of the EBM modules: the percentage of technical words with a range of one in all the EBM modules was less than half.

5.1.1.6 IT/MSE

This component was represented by one module: 'Programmable Systems in Measurement and Control'. Typical technical words and multi-word units in this module were: *boolean algebra, integrated circuits, binary digit* and *microcomputer*, and this suggested that it was actually an IT module.

Fifty five technical words were identified in this module, and eleven of them had a range of one. This constituted 20% of the total number of technical words. As in the **IT** modules in **CORE** and **IT/EBM**, relatively few technical words with a range of one could be identified in this module.

The modules in the **IT** component were found to have a relatively large number of technical words. This pattern indicated that modules in the **IT** component were more highly technical than the **IT** related modules in other components. This pattern was similar to that in two modules: 'Basic Computing' and 'Computer-aided Design and Computer-aided Manufacturing' in the **CORE** component. As mentioned earlier, the **IT** related modules in components other than **IT** deal with more basic knowledge that any Engineering student will need.

5.1.1.7 SELECTIVE

This component was represented by four modules: 'Simulation of Production Systems', 'Project Planning Management and Control', 'Financial Decision Making' and 'Introduction to Manufacturing Systems'. The first three modules belonged to **EBM**, they all displayed the typical features of the **EBM** modules: relatively few technical words had a range of one. The last module belonged to **MSE**. It seems that most technical words in the **IT** and **MSE** components had a range of one, which suggested that these modules were more highly technical. Technical words in the **EBM** component in contrast, had a range of more than one.

Table 5.7 Technical words identified in the SELECTIVE component			
Module	**Number of technical words**	**Number of technical words with a range of one**	**Percentage of technical words with a range of one**
Simulation of Production Systems	33	12	36
Project Planning Management and Control	26	12	46
Financial Decision Making	29	11	38
Introduction to Manufacturing Systems	103	28	27
Total	**191**	**63**	

5.1.1.8 Summary

The total number (632) of technical words with a range of one constituted 41% of the total number (1,539) of technical words identified in the corpus. However, the occurrence of technical words with a range of one differed widely across modules. Six modules had a high percentage (more than 50%) of technical words with a range of one. Four of the six modules: 'Automation and Robotics', 'Computer Engineering', 'Computer Language' and 'Artificial Intelligence' were from the IT component; the other two modules: 'Polymer Materials, Processes and Products' and 'Quality Reliability Maintenance System' were from the MSE and CORE components respectively. This suggested that these modules were highly technical or subject specific; technical words identified in these modules tend to reveal the theme of the particular modules. A lot of technical words in these modules are indexical of that discipline, i.e. unlikely to occur in other disciplines, for example, *polyamide* and *polycarbonate*. In contrast, some modules such as 'Information Technology Fundamentals' and 'Production Planning and Control Systems' had a lower percentage of technical words with a range of one. Technical words or multi-word units such as *organizational behaviour, marketing strategy* and *recruitment* in these modules tended to be used technically in other related fields. The features of these technical words are shared by words discussed in 5.1.2.

5.1.2 Technical words with a wider range

732 technical words identified in my corpus occurred in more than one module. As an example, *assembly language* occurred fourteen times in two modules: eight times in 'Computer-aided Design and Computer-aided Manufacture' and six times in 'Programmable Systems in Measurement and Control'. The first module belonged to the **CORE** component while the second one belonged to the **IT/MSE** component, but both of them were **IT** related modules. The word *byte* was identified as a technical word in three modules: 'Programming Language', 'Computer Engineering' and 'Information Technology Fundamentals'. It also occurred with a frequency of less than five in four other modules. Table 5.8 shows the frequency and range pattern of the word *byte*.

Table 5.8 Frequency and range pattern of *byte*

Module	**Frequency**
Programming Language	18
Computer Engineering	10
Information Technology Fundamentals	9
Basic Computing	3
Programmable Systems in Measurement and Control	2
Computer-aided Design and Computer-aided Manufacture	1
Computer Language	1
Total	**44**

The pattern of the wider range technical words was that they occurred in modules either within the same component or in modules across components but bound by the same specialism. For example, *byte* was identified as technical in three of the modules shown in Table 5.8, two of these belonged to IT and one belonged to the IT/EBM component, but all three modules were IT related. All the wider range technical words identified in my corpus had this feature.

5.1.3 General English words used in a technical sense

Nuttall (1982) believes that some everyday English words can cause serious problems to learners when they occur in a technical context. 'Some of the most dangerous misunderstandings arise when apparently everyday words are used in specialised senses by writers in specialised fields.' (Nuttall, 1982). This concern is emphasized by Kennedy and Bolitho (1984) as they write, ' ... words commonly met in *general* take a specialised meaning within a scientific or technical context'. Kennedy and Bolitho (1984), Nuttall (1982) and Trimble (1985) describe these words as subtechnical, while in this thesis these words were considered technical, because they had a peak frequency in one particular module, or sometimes in more than one module.

For the purposes of these studies, everyday words are defined as words that occur in Thorndike and Lorge's 5,000 most frequent words. An everyday word was identified as technical if it had a technical distribution pattern in certain modules in my corpus. This section discusses six examples of this kind of word: *argument, bit, bus, gate, grain* and *memory*. For the complete list of this kind of technical word, see Appendix Two.

1. *Argument*

Table 5.9 shows the distribution pattern of *argument*.

Table 5.9 Distribution pattern of *argument*	
Module	**Frequency**
* Artificial Intelligence	57
Human Factors in Industry	13
* Computer Language	7
* Programming Language	6
* Automation and Robotics	5
Information Systems Strategy	2
Logistics and Supply Chain Management	2
* Programmable Systems in Measurement and Control	2
Quality Reliability Maintenance Systems	2
Applied Statistical Methods	1
Computer Engineering	1
Financial Analysis and Control Systems	1
Machine Tool	1
Project Planning Management and Control	1
Total	**101**

Argument had a fairly wide range, but its peak distribution was obvious, too: it occurred with relatively low frequency in thirteen modules, whereas it had a peak frequency of fifty seven in one module 'Artificial Intelligence'. Thus according to my criteria it was a technical word.

Argument was used in a technical sense in five IT related modules: 'Artificial Intelligence', 'Computer Language', 'Programming Language', 'Automation and Robotics' and 'Programmable Systems in Measurement and Control'. These modules are asterisked in Table 5.9. Example 1 shows the technical sense of this word:

Example 1:
Let us consider facts when we say "John owns the book", we are declaring that a relationship , ownership exists between one individual object "John" and another individual object denoted by "the book". This relationship has a particular order: John owns the book, but the book does not own John. ... The statement "John owns the book" translates into a PROLOG fact as follows:

owns (john, book)

> The name of the relationship is called a predicate, and the objects inside the brackets are called arguments.
> **(Artificial intelligence,** Rich and Knight 1991: 30)

It can be seen that the word *argument* is an input parameter to a computer programme. The following concordance lines provide examples of *argument* used as a technical word in the above five modules:

the objects inside the brackets are called the *arguments*. So, the above example is a fact about the
These are (file) variables and represent the *arguments* and results of the computation. The block
PTY This is because {Squared} needs one *argument*, as we can see from the stack description
expressed in decibels) and the curve of the *arguments* (expressed in degrees) are traced se
erms of variables A, C and Z where Z is the *argument* of the function. 2.3.5 Subscripted variables

'Computer Engineering' was the only IT related module in which *argument* was used in a General English sense, as shown in Example 2:

> *Example* 2:
> Shannon's result was derived using information-theory *arguments* and has very general validity (**Computer Networks,** Tanenbaum 1996: 57).

Argument occurred in a relatively high frequency in 'Human Factors in Industry'. The following concordance lines show how *argument* occurred in this module.

oint it seems necessary to acknowledge the	*arguments*	against the use of appropriate theory
forward by Anthony (1986). Most of these	*arguments*	are summarized in Reading 2 of this bo
ome areas of the behavioural sciences, this	*argument*	has been well won; in others it has sca
used in the behavioural sciences. Strong	*arguments*	have been mounted against this approa

Although it occurred with a relatively higher frequency than in other non IT modules, *argument* was used here in a General English sense, and means 'set of statements in support of an opinion or proposed course of action'(Collins COBUILD English Language Dictionary, 1995).

Argument was used similarly in all the other non IT related modules: 'Information Systems Strategy', 'Logistics and Supply Chain Management', 'Quality Reliability

Maintenance Systems', 'Applied Statistical Methods', 'Financial Analysis and Control Systems', 'Machine Tool' and 'Project Planning Management and Control'. The following concordance lines provide examples from each of the modules respectively:

saw evidence that the competitive advantage	*argument*	is beginning to be used excessively
it should perhaps be on all of them. A similar	*argument*	holds for Chapter 20. In closing this
ten suffer high costs of failure under warranty.	*Argument*	and misunderstanding begin when we
shown in Table 2.10. Motivated by the above	*argument,*	we can now state the following gener
onal education and training) and despite some	*arguments*	in the data, Britain compares very bad
vibration have since been proposed, the simple	*argument*	relating to the effect of cutting speed o
believe that there is an official definition. The	*argument*	uses the concept of states just like the

It can be seen from the above examples that *argument* was used in a technical sense in all the IT related modules except 'Computer Engineering' and it was used in a non-technical sense in other modules.

2. *Bit*

Table 5.10 shows the distribution pattern of *bit.*

Table 5.10 Distribution pattern of *bit*

Module	**Frequency**
* Computer Engineering	143
* Basic Computing	61
* Programmable Systems in Measurement and Control	25
Information Technology Fundamentals	19
Project Planning Management and Control	19
* Computer-aided Design and Computer-aided Manufacturing	18
* Programming Language	17
* Artificial Intelligence	11
* Computer Language	6
* Automation and Robotics	4
Manufacturing Strategy	4
Introduction to Manufacturing Systems	3
Financial Analysis and Control Systems	1
Human Factors in Industry	1
Logistics and Supply Chain Management	1
Simulation of Production Systems	1
Strategic Marketing	1
Total	**385**

Like *argument*, *bit* also occurred across a wide range of modules, with an equally obvious peak distribution. The term *bit* means a digit in binary notation, i.e. 0 or 1 (the smallest unit in storage). It is explained as follows in one of the textbooks:

> *Example* 3:
> It is convenient for a computer to use a number system based upon the number two, rather than ten with which we are so much more familiar, because each of the signals within the circuit is confined to one of only two possible values: 0 and 1. It does not matter which symbol is assigned to each value; in fact it is impossible to tell which is used by internal circuitry of a microcomputer. However, it is common for low to be represented by the symbol 0 and high by 1 in input/output circuits.
> Each digit in a binary number is called a bit. The word is a rather neat contraction of the words "binary" and "digit", and the "bit" also provides us with a unit of information used in information-handling systems.
> **(Microcomputers in Engineering and Science,** Craine and Graham, 1985: 24)

Bit was used in its technical sense in eight IT related modules: 'Computer Engineering', 'Basic Computing', 'Programmable Systems in Measurement and Control', 'Information Technology Fundamentals', 'Computer-aided Design and Computer-aided Manufacturing', 'Programming Language', 'Computer Language' and 'Automation and Robotics'. These modules are asterisked in Table 5.10. The following concordance lines show the use of *bit* in these modules respectively:

frame is. SNA frames may be an arbitrary number of *bits*, whereas DECNET frames must be
1000. The 1's and 0's are known as binary digits, or *bits*. The numbers of symbols that can be
ut circuits. Each digit in a binary number is called a *bit*. This word is a rather neat contraction of
ractical purposes byte is always composed of 8 *bits*, a word differs in size (i.e. its length
microcomputer usually has a word length of 16 *bits* whereas the microcomputer has a word
byte may be the first byte of a larger item (i.e. a 16 *bit* variable). <name> refers to the next word,
data are represented as sequences of binary digits (*bits*). Higher-level languages allow one to
gives the binary and reflected binary codes for 3 *bits*. It can be seen that the transition from 3

This sense of *bit* was also used in one of the non IT modules: 'Simulations of Production Systems'. Example 4 shows how it was used in this module:

> *Example* 4:
> Most packages have a more elaborate method of generating uniform random numbers. For example, the function used in the ECSL package3 is:
>
> Xn = 3125 × Xn -1 modulo 2m
>
> where m is one less than the number of *bits* in a computer word.
> **(Simulations of Manufacturing Systems,** Carrie 1988: 34)

Although 'Artificial Intelligence' was also an IT module, *bit* was used in an everyday sense in this module. The occurrence of *bit* in this non-technical sense reflected a less formal style of writing, as Example 5 shows:

> *Example* 5:
> It tends to be a **bit** of a brain-twisting exercise to follow the recursion step by step as shown with a longer original version of it.
> **(Artificial Intelligence Rich and Knight, 1991: 13)**

Collins COBUILD English Language Dictionary (1995: 133) explains *bit* in this way:

> A **bit** of something is a small amount of it; an informal use. EG There'll be a bit of sunshine.

There are many other uses of *bit* in this informal way in the module as shown by the following concordance lines:

arer, let's rewrite the last line of the function. The	*bit*	after the space on the last line is what
revarious bits are removed from it and various other	*bits*	are added. But because we're finding
so on. Exactly that happens to our wm as various	*bits*	are removed from it and various oth
can return to top level and to you. To make this a	*bit*	clearer, let's rewrite the last line of t
call LISP a programming language is not merely a	*bit*	of wishful thinking, because - in the
down to arbitrary levels of complexity. The magic	*bit*	of LISP is that it automatically knows
the goal every time we need to look at this or that	*bit*	of it. So we fetch it once and bind the
or changing various elements of them, or joining	*bits*	of them onto other lists. At first sight
knows how to sort out such nests of lists - which	*bits*	to process first and what to do with th
us other bits are added. Because we're finding the	*bits*	we want to remove via assoc rather th

Bit was used in its everyday sense in all the six non IT modules: 'Project Planning Management and Control', 'Manufacturing Strategy', 'Financial Analysis and Control Systems', 'Human Factors in Industry', 'Logistics and Supply Chain Management' and 'Strategic Marketing'. The following concordance lines show the meaning of *bit* in these modules respectively:

are going to start at the start. You can skip	*bits*	if you like. To make life a little easier,
lar results that Japanese manufacturers did a	*bit*	earlier: product defects down from several
marketing prices supplied every conceivable	*bit*	of information for decision making and
n which each of the basic concepts is, with a	*bit*	of stretching, cutting and fitting made to
first glance the lot-for-lot technique seems a	*bit*	too simple-minded since it doesn't con
had a reputation or quality, even if it was a	*bit*	more expensive. Many of the Economy

The one occurrence of *bit* in module 'Introduction to Manufacturing Systems' had a different meaning from all the above as Example 6 shows:

> *Example* 6:
> Basically, machines are mechanized versions of such hand tools. Most tools are for cutting (drill *bits*, reamers, single-point turning tools, milling cutters, saw blades, broaches, and grinding wheels).
> **(Materials and Processes in Manufacturing,** DeGarmo, Black and Kohser 1988: 13)

In this particular context, *bit* was used as a particular part of a tool - drill, which is 'a machine that is used for making holes' (Collins COBUILD English Language Dictionary 1995: 432). As can be seen, *bit* was not used as an IT word in this context, yet it did not have a General English sense either.

One feature of *bit* was that it was frequently used with numbers such as 0, 1, 4, 5, 6, 12, 32, and most frequently with the number 8 (it had 18 occurrences) and 16 (it had 9 occurrences). Rather than being definitive, these numbers are usually defined by the system to group or hold data. The usage of *bit* with the above numbers is shown by the following concordance lines:

its set or used by the operation. The upper *8 bits*	are usually set to zero. 'addr' refers to a 16
000 and eventually reach 99999. Similarly, *8-bit*	binary counter will start at 00000000 and
terns which can be represented using *eight bits*,	but only patterns of eight bits, never
common 200 words could be encoded in an *8-bit*	byte as the number 0 through 199. Codes 2
with two's complement numbers. If an *8-bit*	counter counts back from 00000000 it wil
because it is such a useful size, a block of *eight bits*	is given the special name 'byte'.
two's complement notation and a block of *eight bits*	is 10000000 (decimal -128), while the
tions which can be selected or 'addressed'. *8-bit*	microcomputers usually have address bus
ented using eight bits, but only pattern of *eight bits*,	never seven or nine. Registers make up
0101010 Figure 1.9 Bit positions in an *eight-bit*	register The microcomputer's memory
than as assemblies of gates. Because an *eight-bit*	register consists of a set of eight latches, it
register Register 1 Register 2 Figure 1.10 *8-bit*	bus connecting two 8-bit registers called
number is a multiple of eight. A block of *eight bits*	represents a useful amount of information
to a 16 bit value, but with only the lower *8 bits*	set or used by the operation. The upper 8 bits
bit number operations. FORTH does have *8 bit*	store and fetch operations which are useful
for example, at the centre of each group of *eight bits.*	Systems like this are somewhat restricted
DECNET frames must be multiples of *8 bits.*	The network layers of the two architecture
bers +5, +15 and -15 be represented using *8-bit*	two's complement notation? 5. What is

ve address bus widths of at least 12 bits, and *16-bit*	address buses are virtually the rule. The
mbers', and are represented on the stack as *16 bit*	binary (beginners should look up the
imal number (two digits per byte) follows, a *16-bit*	integer follows, etc. Other possibilities for

neywell DDP-5 16 minicomputers with 1 2K *16-bit* words of memory. Later Honeywell DDP-
may be the first byte of a larger item (e.g. a *16 bit* variable). <name> refers to the next word
e are, like the majority of FORTH operations, *16 bit* number operations. FORTH does have 8 b
bits are usually set to zero. 'addr' refers to a *16 bit* value which represents the address of a
at as C.). (n addr.) 'plus-store'. Add n to the *16 bit* value stored at address, using C+ operatio
se. In the stack descriptions 'byte' refers to a *16 bit* value, but with only the lower 8 bits set

To summarize, *bit* was used as a technical word in the IT related modules ('Artificial Intelligence' being the only exception) but when *bit* was used in the non IT modules, it was used in a non-technical sense (the only exception is the module 'Simulation of Production Systems').

3. *Bus*

Table 5.10 shows the distribution of the word *bus*. Like *argument* and *bit*, *bus* was also used in a technical sense in the IT related modules: 'Programmable Systems in Measurement and Control', 'Computer Engineering', 'Basic Computing' and 'Computer-aided Design and Computer-aided Manufacturing', these modules are asterisked in Table 5.11.

Table 5.11 Distribution pattern of *bus*

Module	**Frequency**
* Programmable Systems in Measurement and Control	26
* Computer Engineering	14
Industrial Engineering	7
* Basic Computing	5
* Computer-aided Design and Computer-aided Manufacturing	2
Production Planning and Control Systems	2
Project Planning Management and Control	2
Strategic Marketing	1
Total	59

The following concordance lines show how *bus* was used in these modules respectively:

not necessarily the case. The width of the address *bus* determines the maximum number of
ese these are required to refrain from sending. A *bus* must have some arbitration mechanism to
eds to be taken out of the minicomputer, address *bus*, RAM, data bus, control bus, data bus,
any of the other CPUs through the common data *bus*. The performance of multiple-instruction

In contrast, in the non IT modules, *bus* was used in an everyday sense to mean 'a large vehicle which carries passengers from one place to another' (Collins COBUILD English Language Dictionary, 1995). Examples 7, 8, 9 and 10 show how *bus* occurred in four of the non IT modules: 'Industrial Engineering', 'Production Planning and Control Systems', 'Project Planning Management and Control' and 'Strategic Marketing':

Example 7:

An operating system is a configuration of resources combined for the provision of goods or services.

Bus and taxi services, motels and dentists, tailors and mines, fire services and refuse collectors, retail organizations, hospitals and builders are all operating systems .
(Production and Operations Management, Wild 1990: 3)

Example 8:

16. A Ridgeview City Transit study of 37 observations of peak load on a certain *bus* route has shown that the load is normally distributed with a mean demand of 114 passengers and $\Sigma(X - \bar{\chi})2 = 2{,}304$. If the transit operations office plans to supply two *buses*, each with a 62-person capacity, what percentage of the potential passengers will not be accommodated by this city service?
(Operations Management, Monks 1987: 67)

Example 9:

Press home these key dates. Let everyone know what they are and how long there is to the next one. There was a large sporting event where the project manager hired a bus and, in the little window at the front of the bus, showed the number of days to go before the big key date - the start of the event. You can erect a flip chart in the office foyer with a little sign saying 'Days to go', and mark the days left on a sheet under it.
(Project Management Demystified: Today's Tools and Techniques, Reiss 1992: 37)

Example 10:

... When you leave home, will it be in a Toyota, or on a Huffy bike, or on the *bus* that the city bought from General Motors?
When you think about it, you can't get very far into a day without bumping into marketing - and what the whole marketing system does for you.
(Basic Marketing: A Global Managerial Approach, McCarthy and Perreault 1993: 3)

Although *bus* was a content word in all the examples, it was not specific to the subject of the text. Instead, all the occurrences above were used as 'example' words to illustrate the author's main point. Example words are discussed in 5.1.4 and 5.2.2.

To sum up, the word *bus* was used as a technical word in all the IT modules in my corpus, whereas it was used in its everyday sense in the non IT modules.

4. *Gate*

Table 5.12 shows the distribution pattern of *gate*.

Table 5.12 Distribution pattern of *gate*

Module	**Frequency**
Programmable Systems in Measurement and Control	29
Basic Computing	17
Polymer Materials, Processes and Products	1
Total	47

Gate was used as a technical word in the module 'Programmable Systems in Measurement and Control', meaning a circuit that controls the flow of binary signals. This sense is illustrated in Example 11:

> *Example* 11:
> Modern computers and microcomputers use transistors as switches, but these transistors are connected to form logical building bricks known as *gates*. A logic *gate* is an electronic circuit with one or more inputs and one output. The voltages on the inputs and outputs of these *gates* can each take one of only two values, and the *gates* are each designed so that the relationships between this input and output levels are determined by a simple boolean expression.
> **(Microcomputers in Engineering and Science,** Craine and Graham 1985:22)

In this module *gate* co-occurred five times with *NAND* as in *NAND gate*.

Similarly *gate* was used as a technical word in 'Basic Computing' as shown by the following concordance lines:

of the original signal. By using the output from one *gate* as part of the input for another gate,
ough what are known as 'logic elements' or 'logic *gates*'. Logic gates Circuits are built up using
uit works for all). This can be done using two AND *gates*, one OR and one NOT, combined in the
of gates that we have not mentioned, and by using *gates* which can have more than two inputs,

One occurrence of *gate* was similar to that of *bit* in one of the non IT modules, which had a technical sense not related to IT, but related to the topic of the text in which it occurred. Example 12 shows how *gate* was used in 'Polymer Materials, Processes and Products':

Example 12:
The classic example of ESC is the brittle cracking of polyethylene washing-up bowls due to the residual stresses at the moulding *gate* (see injection moulding, Chapter 4) coupled with contact with the aqueous solution of washing-up liquid.
(Plastics Engineering, Crawford 1987: 27)

Like *bit*, *gate* was not an IT word in this context although most other occurrences of this word were IT related. *Gate* did not have a General English sense here either.

5. *Grain*

In a technical sense, *grain* means the average size of mineral crystals composing a rock. Table 5.13 shows the distribution pattern of *grain*.

Table 5.13 Distribution pattern of *grain*

Module	**Frequency**
Metallic Materials Selection	93
Introduction to Manufacturing Systems	52
Machine Tool	3
Manufacturing Strategy	1
Production Planning and Control Systems	1
Total	**150**

Grain was used in a technical sense in two of the MSE modules: 'Metallic Materials Selection', and 'Machine Tools'. It co-occurred very frequently with words such as *structure, diameter, boundary* and *size* in 'Metallic Materials Selection', as shown in Table 5.14.

Table 5.14 The co-occurrence of *grain* in 'Metallic Materials Selection'	
Combination of *grain*	**Frequency**
grain structure	2
grain diameter	5
grain boundaries	21
grain size	34

In 'Machine Tools', *grain* occurred together with the word *crystal*, as shown by the following concordance lines:

perature. It will be seen that the strained crystal *grains* have regained their normal state. When
temperature, the disposition of the new crystal *grains* shows up discontinuities in the material
graphic technique for examining the crystalline *grain* structure of a metal by polishing and

Similarly, *grain* was used in a technical sense in 'Introduction to Manufacturing System'. The following concordance lines show some of the occurrences:

tinuous segments of solid are known as crystals or *grains*, and the surfaces that divide them (i.e.,
ture is produced is one of nucleation and growth. *Grains* are the smallest structural units of
l processing procedures. Development of Metallic *Grains*. As meters solidify, a small particle of
n a metal is deformed a considerable amount, the *grains* become elongated in the direction of

In contrast, *grain* was used in a General English sense (as an example word) in 'Manufacturing Strategy' and 'Production Planning and Control Systems'. Example 13 shows how it occurred in my corpus:

> *Example* 13:
> Great Falls Export Company has 30 employees and handles 1,500 loads per year of *grain* from a North Dakota warehouse.
> (Operations Management, Monks 1987: 65)

Grain co-occurred with *boundary*, *structure* and *size*, as shown in Table 5.15.

Table 5.15 The co-occurrence of *grain* in 'Introduction to Manufacturing Systems''	
Combination of *grain*	**Frequency**
grain boundaries	7
grain structure	8
grain size	12

The pattern was obvious that *grain* was used as a technical word in all the MSE modules.

6. *Memory*

Memory in a technical sense, means part of a computer system where data and instructions are held. Table 5.16 shows the distribution pattern of *memory*.

Table 5.16 Distribution pattern of *memory*	
Module	**Frequency**
* Information Technology Fundamentals	92
* Basic Computing	65
* Computer-aided Design and Computer-aided Manufacturing	54
* Programming Language	35
* Programmable Systems in Measurement and Control	23
* Artificial Intelligence	19
* Computer Engineering	12
* Computer Language	4
Manufacturing Strategy	3
Production Planning and Control Systems	3
* Automation and Robotics	2
Human Factors in Industry	2
Simulation of Production Systems	2
Information Systems Strategy	1
Polymer Materials, Processes and Products	1
Project Planning Management and Control	1
Quality Reliability Maintenance Systems	1
Total	**320**

Memory was used in a technical sense in all the nine IT related modules. In Table 5.16 these modules are asterisked. The following concordance lines show the context of this word in these modules respectively:

of the computer, and together share its available *memory* addresses. Semiconductor devices are
which would send an electrical pulse through the *memory* unit when it sensed a 1 bit, and not
ly used today are intelligent terminals with local *memory* and a local microprocessor, so they are
value which represents the address of a byte in *memory*. The addressed byte may be the first
controlling the flow of information to and from *memory* and input/output devices. Some
es and the rings numeric names. Our working *memory* contains only four elements: the names
stems can be implemented by having a common *memory* shared by the processors. Usually, this
tor they also suggest an opportunity to conserve *memory* space and to introduce validity checks
role of the brain is taken by the computer. In its *memory* are stored: 1. a model of robot

Memory was also used in a technical sense in some of the non IT modules: 'Production Planning and Control Systems', 'Simulation of Production Systems', 'Information Systems Strategy' and 'Project Planning Management and Control'.

Memory as a non-technical word was only used in three modules: 'Human Factors in Industry', 'Polymer Materials, Processes and Products' and 'Quality Reliability Maintenance Systems'. The following concordance lines show the occurrences in these modules respectively:

ourage to do it that way, urging that without the *memory* hooks provided by alliteration, our staff
the presence of creep is recognized, and (b) the *memory* which the material has for its stress
dent, i.e. the coin and the die logically have no *memory*, so whatever has been thrown in the pas

Memory occurred three times in 'Manufacturing Strategy': in one occurrence it was used in a technical sense (as shown in Example 14) whereas in the other two it was used in a General English sense (as shown in Examples 15 and 16):

Example 14:
Compared with factory folks, office employees find it easier to pursue diversions, because their work and their inventory - a piece of paper or an entry into computer *memory* - is not very visible.
(**World Class Manufacturing**, Schonberger 1986: 53)

Example 15:
The yellow-light approach not only gives people the chance to explain real causes; it has them explain right when the event occurs, so there are no questions about bad *memories* and guessing.
(ibid: 20)

Example 16:
We see that the longest down/bad time in the twenty-four hours was between hour 18 and hour 21. That means there must be a three-hour buffer stock before the next process, Y. Y's down/bad duration, in turn, determines the buffer stock before Z. Actually, the buffer stock would be based on the longest down/bad time in recent *memory* - and that could be weeks!
(ibid: 70)

The pattern for *memory* was fairly obvious, too: in all the IT related modules and some of the non-IT modules it was used in a technical sense. In only three modules: 'Human Factors in Industry', 'Polymer Materials, Processes and Products' and 'Quality Reliability Maintenance Systems' was *memory* used in a non-technical sense.

To sum up this section, all the words discussed above were everyday words used in a technical sense. One pattern was that most of these words such as *bit, bus, gate* and *memory* were technical words in the IT component. This was probably because Computer Science literature has borrowed some of its terminology from General English. *Bit* and *gate* were used in a technical sense related to the MSE component. The word *grain* was used as a technical word in the MSE modules. Words like these were the ones that were considered difficult items by Kennedy and Bolitho (1984), Nuttall (1982) and Trimble (1985). Some of these words, *bus* for example, were not always used in a technical sense. *Bus* in the sense of 'vehicle' was used as an example to illustrate concepts. *Bus* and other similar words are discussed in more detail in 5.2.2.

5.1.4 Low range example words

Some words had a low range but were not technical; they were everyday words which were used in everyday contexts but in examples to illustrate concepts that the textbook writer was trying to explain. I call this type of words low range example words. Another type of words had a similar function, but occurred across a wide range of modules. Those words are discussed in section 5.2.2 of this chapter, because according to my frequency and range criteria, those words are classed as subtechnical rather than technical.

The corpus contained twenty three low range example words including multi-word units (see Appendix Two). Some examples in the corpus were lengthy and very much in detail, like the example using *tennis racket* in the module 'Strategic

Marketing'. The author of this textbook uses this example to illustrate the idea that 'marketing is more than selling and advertising'. He starts off with the function of a tennis racket - 'to hit the ball over the net', and then goes into detail about shapes, materials, weights, handle sizes, and type of strings before moving on to prices: 'You can buy a prestrung racket for less than $15. Or you can spend more than $250 just for a frame!' (**Basic Marketing,** McCarthy and Perreault 1993: 4). A list of eight items follows to show 'some of the many things a firm should do before and after it decides to produce tennis rackets'. The example takes up the space of one and half page.

In the same book *mousetrap* was also used as an example as Example 17 shows:

> *Example* 17:
> Although production is a necessary economic activity, some people overrate its importance in relation to marketing. Their attitude is reflected in the old saying: "Make a better mousetrap and the world will beat a path to your door."
> (**Basic Marketing: A Managerial Approach,** McCarthy and Perreault 1993: 5)

Examples are supposed to help the reader to understand the main ideas, on the assumption that the example is about a very familiar object, person or anecdote. For the *tennis racket* example to be effective, for example, the reader should have seen a tennis racket, a tennis ball and a net, he or she should know at least something about the shape of a racket, what materials it is usually made of, how heavy it normally is, what handle sizes are available on the market and so on. If this is not the case, he or she will not get much information from the example. It is the same with the *mousetrap* example: without the same cultural background as the writer, it is difficult to understand the 'better mousetrap' idea:

Example 18:
The "better mousetrap" idea probably wasn't true in Grandpa's time, and it certainly isn't true today. In modern economies, the grass grows high on the path to the Better Mousetrap Factory - if the new mousetrap is not properly marketed.
(ibid: 5)

Tennis racket and *mousetrap* are both topic specific words. They are not, however, related to the topic of the text ('marketing') as a whole. The writer uses these two words because they denote familiar (and, in the case 'mousetrap', somewhat amusing) examples rather than because these two words are particularly related to the concept of 'marketing'. Example 19 is another example of an example word unrelated to the topic as a whole:

Example 19:
Sometimes an item may be considered as both repairable and non-repairable. For example, a *missile* is a repairable system whilst it is in store and subjected to scheduled tests, but it becomes a non-repairable system when it is launched. Reliability analysis of such systems must take account of these separate states. Reparability might also be determined by other considerations.
(Practical Reliability Engineering, O'Connor 1991: 11)

In this example, the writer uses *missile* as an example to clarify the concept of 'repairable' and 'non-repairable' systems. However, the writer could also have used other words such as *vehicle* to illustrate the same concept, which he later did, as Example 20 shows:

Example 20:
For example, ... an engine or vehicle will be treated as repairable only up to a certain age.
(Practical Reliability Engineering, O'Connor 1991: 11)

Low range example words in my corpus were found to be used with other example words, for example, *dice* is used with *coins* and *roulette wheels* for the concept of 'probability'.

Example 21:
With coins, *dice* and roulette wheels we can predict the probability of the outcome from the nominal nature of the system. A coin has two sides, a die six faces, a roulette wheel equal numbers of reds and blacks. Assuming that the coin, die and wheel are fair, these outcomes are also unbiased, i.e. they are all equally probable, in other words they occur randomly. **(Practical Reliability Engineering, O'Connor 1991: 13)**

This is another example showing that low range example words, although topic specific, are not related to the topic of the texts they are in. An example word is usually used to help the explanation or development of the topic or concept of the text. As an example, the reason that the *tennis racket* example frequently occurred in the book **Basic Marketing** (McCarthy and Perreault, 1993) was not because the author wanted his readers to learn the concept *tennis racket*, but because he wanted to support his argument that 'marketing is more than selling and advertising'. However, if a reader does not know these words, it will be very difficult for him or her to get the meaning behind the example. As Bloor (1996) put it, an example is characteristically used by textbook writers and lecturers to introduce new concepts or difficult procedures by reference to something already known. As also noted by Bloor, an example can, however, prove a stumbling block to non-native or overseas students, who (for cultural or linguistic reasons) may not be familiar with the colloquial vocabulary or who do not recognize the rhetorical shift from proposition to analogy.

5.2 Subtechnical words

In this thesis, a word was identified as subtechnical if it occurred across a wide range of modules, whereas a technical word was one that had relatively low range. Some technical words occurred in only one module, some occurred in a wider range but with a peak distribution in one or more than one module. This was the distributional difference between subtechnical and technical words with a wider range.

In the analysis of technical words identified on the basis of frequency and range criteria (section 5.1) a lot of technical words were found to be topic specific, and to represent highly technical concepts within the given subject specialism. We would also expect subtechnical words to share characteristics other than patterns of frequency and range. Cowan (1974), Inman (1978) and Farrell (1990) claim that subtechnical vocabulary is 'context independent', and Marshall and Gilmore (1993) and McCarthy (1990, 1991) point out that subtechnical words are used to express relations between key concepts. Subtechnical vocabulary is nevertheless claimed to be 'elusive and confusing ' (Baker, 1988).

The following section investigates the meaning and function of subtechnical words by taking a closer look at the subtechnical words identified in my corpus on the basis of their patterns of frequency and range. Three kinds of words fell into the category of subtechnical vocabulary on the basis of their distribution. Of the total 1,210 subtechnical vocabulary identified, 1,041 words were procedural, 145 of them were wide range example words, and 23 words belonged in a 'rag bag' category (see Appendix Three).

5.2.1 Procedural vocabulary

Words in my procedural category are primarily metatextual. These words have a metadiscursive function and were used to 'talk about' other things. McCarthy (1991: 78) defines procedural vocabulary as words that 'enable us to do things with the content-bearing words or schematic vocabulary'. Robinson (1988) refers to procedural vocabulary as 'the simple lexis of paraphrase and explanation' and considers procedural words 'the main element in our interpretation and categorization of specific frames of reference'. He gave an example of two dictionary entries:

vermiculite *n* [U] a type of MICA that is a very light material made up of threadlike parts, that can be used for keeping heat inside buildings, growing seeds in, etc.

vermiform *adj* shaped rather like a worm
(Longman Dictionary of Contemporary English, 1978)

The procedural vocabulary in these examples is *type, material, made up, parts, used, keep, shaped* and *like*. These words and many other general words are highly useful not only in talking about specific words like *vermiculite*, but also in the cognitive process of categorizing and organizing features of meaning relative to other known entities. Thus, *vermiculite* is crucially a material, and is made up of a, b and c, is used for x, y and z; and it is a type of something, too. All these words are part of the procedural vocabulary used to talk about and conceptualize the relationships between items and fields. These words belong to the category of core vocabulary, and would be counted in my General English words.

Some of the procedural words in my data are of the type Francis (1986) calls 'anaphoric nouns'; some of the these words are 'Vocabulary Three' items, which are open lexical items and function as signposts in clause relations; and words used for hedges are also considered as procedural, because hedging is usually used to organize the discourse, engage the audience and signal the writer's attitude (Hyland, 1998). The following section discusses these categories.

Francis (1986) describes anaphoric nouns as being used to 'talk about the ongoing discourse', and to refer 'metadiscursively to a stretch of discourse preceding it' (Francis 1986: 3). In addition these nouns must 'face backward'. The noun 'must be presented as the *given* information in terms of which the *new* propositional content of the clause or sentence in which it occurs is formulated' (Francis, 1986: 4). In other words, these nouns must be presented as synonymous with the proposition(s) immediately preceding it. The use of an anaphoric noun usually sets up a two-part discourse relation, the first part is the proposition that is synonymous to the anaphoric noun, the second part is the clause that contains the anaphoric noun. In the discussion of anaphoric nouns I will use (1) for the first part, and (2) for the second part of the discourse.

Here are some examples of anaphoric nouns from my corpus:

Example 22:
(1) In this experiment a stress is applied to a plastic test-piece and the strain is recorded after a time, t (typically 100 seconds). The stress is then removed and the plastic allowed to recover, normally for a period of 4t. A larger stress is then applied to the same specimen and after recording the strain at time t, this stress is removed and the material allowed to recover.
(2) This *procedure* is repeated until sufficient points have been obtained for the isochronous graph to be plotted.
(Plastics Engineering, Crawford 1987: 47)

Example 23:
(1) As dislocations move, they are more likely to encounter and interact with other such dislocations, thereby producing resistance to further motion. Moreover, mechanisms exist to markedly increase the number of dislocations in a metal during deformation, the effect being an increased probability of interaction.
(2) These *phenomena* become significant when one considers mechanical working processes operating in the cold-working range.
(Materials and Processes in Manufacturing, DeGarmo, Black and Kohser 1988: 75)

In Example 22, (1) presents the *procedure*. (2) then introduces a new stage in the argument with the statement that 'This (what is now labelled) *procedure* is repeated .... And similarly in Example 23, (1) presents the *phenomena*, and (2) introduces a new stage in the argument.

As can be seen from Examples 22 and 23, the discourse relation is not achieved by the use of anaphoric nouns alone, but by the combination of anaphoric nouns and the definite reference item *this*. Other reference items found in my corpus were *the*, *these* and *same*. The following concordance lines show examples of nominal groups consisting of these three modifiers with the noun *characteristics*:

the females. Again this result reflects just the *characteristics* of that sample and does not
ristics. While there is a wide range of these *characteristics*, they share a common tendency to
e series of values obtained exhibits the same *characteristics* as our observed statistics on the

The modifier *such* was found either used alone or with *this* or *these* as in *analysis such as this* and *complications such as these*.

Example 24:
Here, individuals are asked to describe incidents relating to their work that they found particularly satisfying or dissatisfying. Their replies are then carefully analysed to uncover underlying themes and reactions. Finally, job satisfaction can also be assessed through interviews. Unfortunately *such procedures* are often long and costly; ...
(**Production and Operations Management,** Wild 1990: 66)

Example 25:
If we choose to purchase items rather than make them, what costs will necessarily be avoided and how do these compare with the costs which will be incurred?
An *analysis such as this* is often referred to as incremental cost analysis.
(ibid: 75)

Example 26:
The prospect of making an item on presently under-utilized equipment is a far more attractive one than if the present facilities are fully utilized and capital investment in equipment or recruitment of labour would be necessary. Nevertheless, even in the latter case it may be desirable to make items oneself if the quantity required is likely to increase considerably, or if the present manufacture of another item is due to cease, thus liberating capacity.
It is precisely because of *complications such as these* that ....
(ibid: 75)

A further group of modifiers was also found to form nominal groups with anaphoric nouns. The function of these modifiers is to complement and extend the organizational role of anaphoric nouns. These modifiers include items such *as further, similar, a second* and *final*. They do not add much information in terms of propositional meaning but serve to 'sequence' the various points or stages of the argument. Some of these modifiers form a nominal group as cataphoric references, which means they 'face forward', or allow the reader to 'anticipate' the information that follows next. Example 27 illustrates the use of *another* as an 'anticipatory' modifier.

Example 27:
(1) We could make **shorten** correspond even more closely to the verbal description of a loop by using the system function **and** instead of **cond**. **And** takes any number of arguments and evaluates them in order until one of them returns **nil**, or until there are no arguments left. For example,

(and 'cat 'dog nil 'mouse)

will return **nil**, but

(and 'cat 'dog 'mouse)

will return **mouse**. (2) So here's *another version* of the iterative **shorten**: ...
(Artificial Intelligence, Rich and Knight 1991: 42)

Here the nominal group *another version* has cataphoric reference, which is lexically realized in the rest of (2). However, *version* is an anaphoric noun which refers back to (1). The writer could not have introduced the concept of *another* without signalling to the reader what has been introduced before. The function of the nominal group is to maintain the ongoing discourse but at the same time signal to the reader that the writer may shift to another aspect of the topic. In Example 27, it effectively signals that the discussion of the previous *version* is over, and it is *another version* to be considered next.

It can be seen from Example 27 that the nominal group *another version* has cataphoric reference, although *version* is anaphoric. Thus nominal groups consisting of modifiers and anaphoric nouns do not always function as anaphoric devices. Example 28 illustrates this point:

> *Example* 28:
> (1) When metals are plastically deformed, they strain-harden (or work-harden); that is, they become harder and the yield stress is raised.
> (2) This is a progressive *phenomenon*.
> **(Materials and Processes in Manufacturing,** DeGarmo, Black and Kohser 1988: 45)

Example 27 differs from Example 28 in that the modifier *progressive* indicates more clearly what kind of *phenomenon* would be discussed or talked about next: it is progressive rather than radical for example; whereas *another* in *another version* only indicates to the reader that the topic of the *previous version* is closed. The modifier in Example 28 is an essential component of the sentence, because 'this is a phenomenon' would make no sense, or it would focus attention on *phenomenon*, which has no potential to move the discourse forward. *Phenomenon* in the context is used anaphorically as given information, but the modifier must be interpreted as new. Thus it is a cataphoric signal: the reader now expects *progressive* to be explained, as it is in the next sentence. Example 29 shows the immediate following context to Example 28:

> *Example* 29:
> As the load is increased to produce plastic deformation, an even greater load will be required to produce further flow. Various materials strain-harden at different rates.
> (ibid)

Anaphoric nouns exhibit some similarities with Winter's Vocabulary Three (Winter 1977). Vocabulary Three is a class of open system lexical items which facilitate a reader's interpretation of discourse (see Chapter One). These items have 'similar semantic properties to closed system items in sentence connection', and they constitute a special vocabulary of context for the clause relations of English; they are words which can function as special signposts of what a word means in sequence with its adjoining sentence. One of the most important signposting functions of Vocabulary Three items is that they indicate what kind of information is to be presented in the sentence or sentences to follow: they allow the reader to anticipate this information. In this sense they function most like the closed system items, the subordinators and sentence connectors (Vocabulary One and Two respectively) with which they exhibit paraphrase relations.

For example, the Vocabulary Three item *contrast* is very similar to the Vocabulary One subordinator *whereas* and the Vocabulary Two sentence connector *however*. All these words are characteristic signals of the clause relation 'comparative denial'. Winter (1977) defines clause relation as 'the cognitive process whereby we interpret the meaning of a sentence or group of sentences in the light of its adjoining sentence or group of sentences'. Hoey (1983: 19) categorises clause relations into two broad classes: Logical Sequence relations and Matching relations. Logical Sequence relations are relations between successive events or ideas (including Condition-Sequence and Cause-Consequence), whereas Matching relations are relations where statements are matched against each other in terms of degrees of identicality of description (including Contrast and Compatibility).

Examples 30 and 31 are Logical Sequence relations found in my corpus, this relation is signalled by Vocabulary Three items *subsequent* and *preceding* respectively:

Example 30:
The three main categories of hardware components are as *follows* and they will be described individually in *subsequent* sections of this chapter.
(**Computer-aided Design and Manufacture**, Groover and Zimmers 1984: 25)

Example 31:
Any name that begins with an uppercase letter is a variable. In the *preceding* data base two facts would match this question. The first matching fact is **likes (john, mary)**, ...
Another way to ask ...
(**Artificial Intelligence**, Rich and Knight 1991: 4)

In Example 30 the items *follows* and *subsequent* can be roughly paraphrased in the sense of *after* (Vocabulary One) and *hereafter* (Vocabulary Two). Example 31 is similar in the sense that *preceding* can be paraphrased with *before* either as a Vocabulary One subordinator or a Vocabulary Two connector. It allows the reader to anticipate that the writer will explain the preceding data, which he or she does next. The sequence does not reflect chronological order, but rather the physical order of the text on the page.

Example 32 shows an instance of Instrument-Achievement relation found in my corpus. Instrument-Achievement relation is a relation of instrument and result. The instrument is expressed by a Vocabulary Two item *thereby,* and Vocabulary One item *by-ing*, both of which can be used to convey the meaning of 'instrument', and 'what was the result?' can be used to convey the meaning of 'result'. This relation can be signalled by a Vocabulary Three item *technique*:

Example 32:
But regardless of the *technique* employed, one should always apply good judgment to see if the answer is reasonable.
(**Operations Management,** Monks 1987: 47)

Examples 33 and 34 are two Matching relations, indicating a comparing or contrasting relation. This relation is signalled by: *differentiate, synonyms* and *alternatively* respectively:

Example 33:
First, we need to *differentiate* between training and development which are often wrongly used as *synonyms*.
(**Managing Organizations,** Wilson and Rosenfeld 1990: 23)

Example 34:
At any one time one or more of the forces may be exerting particular pressure on the competing firms. The existing rivals may be competing viciously via a price war and/or aggressive advertising campaigns. *Alternatively*, competitors may be 'cooperating' to ward off an external threat.
(Strategic Planning for Information Systems, Ward and Griffiths 1990: 60)

Nouns that signal anaphoric devices and Vocabulary Three items that signpost clause relations are important in the organization of arguments. It needs to be noted that anaphoric nouns and Vocabulary Three items overlap in many cases and therefore are not mutually exclusive. For example, the word *technique* is used as an anaphoric noun, it is also a Vocabulary Three item that signposts an Instrument-Achievement clause relation (see Example 32).

Apart from anaphoric nouns and Vocabulary Three items, other aspects of metadiscourse such as hedging was also found to help organising the discourse or the author's stance towards its content. Hedging refers to words or phrases 'whose job it is to make things fuzzier (Lakoff, 1972), implying that the writer is less than fully committed to the certainty of the referential information given. Hedges reflect a relation between writers and readers, the use of hedges is an important means of both supporting the writer's position and building writer-reader relationships (Hyland, 1994). Hedges are items such as *may, might, could, possible, perhaps, probability, believe, assume*, which mark the writer's reluctance to present or evaluate propositional information categorically (Holmes, 1988; Hyland 1996 a.b). These forms imply that statements contain personal beliefs based on plausible reasoning, rather than the definite knowledge that the propositions conveyed is undoubtedly true. By expressing the tentativeness, the writer is avoiding personal accountability for statements.

Hedges found in my corpus tend to reveal the author's opinions through the use of words such as *uncertainty*, *approximate* and *estimation*. The principal role of these words in the text is to indicate the author's less-than-fully-committed attitude towards the certainty of the information given. That is, they tend to express the author's tentativeness. Halliday (1970) regarded this kind of tentativeness 'an intrusion' by the author, because the use of hedges allows the authors to downplay their statements and anticipate audience response by adjusting the degree of certainty they give to their claims.

Example 35:
Examination of photomicrographs of chips shows that chips of this type could be formed *approximately* by the type of deformation illustrated in the model.
(**Modern Workshop Technology**, Baker 1966: 1)

In this example, the author uses the word *approximately* to indicate that the statement being made is not an established truth or well known fact, this enables him to anticipate possible oppositions to the claims made.

The claims made in the following examples are all presented in a tentative manner as above to allow possible refutation by the reader. In other words, the use of *potentially, primarily* and *overall* in the following examples enables authors to 'construe a situation in terms of variations from how the discourse community conventionally structures the world' (Hyland, 1998: 451).

Example 36:
Such concepts, their meaning and interpretation, can *potentially* be an area of interdisciplinary discussion.
(**Managing Organizations,** Wilson and Rosenfeld 1990: 33)

Example 37:
At low temperatures, resistivity becomes *primarily* a function of crystal imperfections.
(**Materials and Processes in Manufacturing,** DeGarmo, Black and Kohser 1988: 24)

Example 38:
On the other hand, daily processing of part of the records could lead to even greater *overall* computational cost than weekly regeneration.
(**Manufacturing Planning and Control Systems**, Vollmann, Berry and Whybank 1992: 40)

Example 39:
It is crucial to appreciate that the number 67.4 is *fundamentally* different from the individual characters '6', '7', and '4', which happen to be stored in that sequence.
(**This is IT: A Manager's Guide to Information Technology**, Eaton, Smithers and Curran 1988: 39)

My data suggests that metadiscourse is a salient feature of academic texts sampled for my corpus, as Table 5.17 shows.

Table 5.17 Frequency and range of the procedural vocabulary analysed in section 5.2.1

	Frequency	Range
overall	176	23
characteristics	163	23
procedure	129	18
procedures	129	18
differentiate	101	20
technique	99	17
fundamentally	93	18
primarily	70	17
potentially	66	17
complications	59	14
version	44	14
approximately	38	15
alternatively	23	14
phenomena	23	11
analysis	16	7

5.2.2 Wide range example words

Example words such as *airline* were similar to the word *mousetrap* discussed in section 5.1.4 of this chapter. Although they were different in terms of frequency and range, they had the same function in texts as examples. As discussed earlier, words such as *mousetrap* were used in a lengthy anecdotal or storytelling way, that is why they occurred with a high frequency within one particular module, whereas example words such as *airline* were only used as a short example and they occurred across a wide range of modules. Table 5.18 shows the distribution pattern of the word *airline*:

Table 5.18 Distribution pattern of *airline*

Module	**Frequency**
Information Systems Strategy	7
Production Planning and Control Systems	7
Strategic Marketing	3
Computer Engineering	2
Human Factors in Industry	2
Information Technology Fundamentals	2
Project Planning Management and Control	2
Quality Reliability Maintenance Systems	2
Basic Computing	1
Computer-aided Design and Computer-aided Manufacturing	1
Industrial Engineering	1
Manufacturing Strategy	1
Total	**31**

As can be seen, *airline* occurred with a fairly low frequency in each module, which means it was used by each author briefly once or twice. Example 40 illustrates:

> *Example* 40:
> Industry specific protocols, such as for banking and *airline* reservation, allow computers from different companies to access each other's data bases when that is needed.
> (**Computer Networks,** Tanenbaum 1996: 21)

Television, aeroplane, automobiles, credit and *cheque* were similarly used within examples (see Examples 41-44):

> *Example* 41:
> A second advantage of digital transmission is that voice, data, music, or even images, such as *television*, facsimile, or video telephone, can all be multiplexed (mixed together to make more efficient use of the equipment.
> (**Computer Networks,** Tanenbaum 1996: 20)

> *Example* 42:
> For example, many people have a good conceptual, if not detailed, knowledge of how, say a car or aeroplane functions and of the role they play in transportation systems: yet in their own way they are no more complex than a computer.
> (**This is IT: A Manager's Guide to Information Technology,** Eaton, Smithers and Curran 1988: 29)

Example 43:
No one disputes the need for articles to be reliable. The average consumer is acutely aware of the problems of less than perfect reliability in domestic products such as TV sets and *automobiles.*
(Practical Reliability Engineering, O'Connor 1991: 30)

Example 44:
In the US, Merrill Lynch launched its cash management account back in 1978. This combines traditionally separate banking products such as line of *credit*, *cheque*, investment and equity accounts into a single monthly statement, with idle funds being swept automatically into a high interest account.
(**Strategic Planning for Information Systems**, Ward and Griffiths 1990: 57)

Like some of the low range example words, wide range example words were also used with other example words. In Example 45, a wide range example word *ambulance* is used with *taxi* and *bus service* as an example of a 'transporting system'.

Example 45:
Thus transport and service systems are dependent on customers not only taking their output and in some cases specifying what that output shall be, but also for the supply of a major physical input(s) to the function without which the function would not be achieved. For example, in transport, a taxi, *ambulance* or bus service moves customers or something supplied by them, e.g. pieces of luggage.
(Project Management Demystified: Today's Tools and Techniques, Reiss 1992: 11)

It can be seen that low range example words (such as *racket* and *mousetrap)* and wide range example words (such as *airline* and *ambulance)* had the same function in the context: to illustrate concepts that the textbook writer is trying to explain. They are widely used by writers because these words are supposed to be familiar to all readers, and require the activation of shared background knowledge. All the readers are supposed to know what a *bus* is, but knowing this entails knowing lots more about systems and expectations in society.

The distinction between low and wide range example words is a product of the corpus make-up and size. A different corpus would yield different distribution patterns for these words.

5.2.3 Other wide range words

I use 'other wide range words' to refer to a 'rag bag' category of items that are not example words, not procedural, not General English words (not within the Thorndike and Lorge 1944 5,000 word count), and not technical words. Twenty four words were found to fall into this category:

> academic, accumulator, albeit, chart, economist, hopefully, investor, lowercase, marital, merchants, monitor, monitors, nineteenth, pinion, producer, retailer, rivet, satellite, semicolon, tidy, traditional, transistor, xerox, zero

The word *albeit* is a logical connective, and is strictly speaking a grammatical word, but its occurrences in the Thorndike and Lorge count (1944) are rare, and therefore it was not removed from my corpus along with other logical connectives such as *although.*

Hopefully, tidy and *traditional* might also be considered General English words. However, as noted elsewhere, a General English word is defined as a word that is within the first 5,000 words in the Thorndike and Lorge count (1944). These three words have surprisingly low frequency in the count - they occurred eight, four, and four times respectively.

Chart, economist, investor, merchants, monitor, monitors, producer, retailer, satellite and *transistor* are a group of wide range words which are topic specific, and related to the topic of the text as part of the central information content of the text. For example, *investor* and *merchants* were used to explain the concept of 'trading'.

> *Example* 46:
> In some respects, this is not a new phenomenon. If we go back five hundred years to the publication of perhaps the original accounting book by Fra Pacioli, we can ask what kinds of events were occurring in fifteen century. Venice that led to a demand for accounting information. Undoubtedly *merchants* were trading goods with other countries. Consider a group of *investors* who acquired goods produced in northern Italy and chartered an expedition to sell them in India.
> **(Relevance Lost: The Rise and Fall of Management Accounting,** Johnson and Kaplkan 1987:14)

Likewise *satellite* is used to explain the concept of 'telecommunication'.

> *Example* 47:
> A second possibility is a satellite or ground radio system. Each IMP has an antenna through which it can send and receive. All IMPs can hear the output from the *satellite*, and in some cases they can also hear the transmission of their fellow IMPs to the satellite as well.
> **(Computer Networks,** Tanenbaum 1996: 26)

Billion, lowercase, marital, nineteenth, pinion, rivet, semicolon, xerox and *zero* are a group of 'odd' words, which occur only once in each of the texts in which they occur. These words all have the same frequency and range pattern: they occur five times in five texts.

5.2.4 Summary of section

Section 5.2 of this chapter discusses the subtechnical words identified in my corpus based on the patterns of frequency and range. A small number of words (12%) classed as subtechnical in my corpus is example words, such as *airline* and *bus*. Example words are used in the corpus with a wide range because different writers use the same examples. Twenty four words are considered to be in a 'rag bag' category, which constitute 2% of the subtechnical vocabulary identified. A majority of 86% of the subtechnical vocabulary has a procedural or metadiscursive function in the organization of arguments. Some of these words are anaphoric nouns, some of them are Vocabulary Three items, and some of them are used in hedged propositions. All these words are important in reading and text structure. Hedges, for example, reflect the relationship between a writer and readers. These are important means not only to support the writer's position but also to build writer-reader relationships. The use of hedges suggests that academic discourse is not simply objective or informational. Nor is it true that academic texts are written in an impersonal style. Knowing the use of hedges enables readers to predict lines of thought and interrogate authors on their positions regarding propositions.

Subtechnical vocabulary occurs with a wide range across modules or texts. Barber (1962), Inman (1978) and Farrell (1990) found similar subtechnical words, for example:

> approximate, continuously, conventional, deflection, distortion, estimation, evaluate, evaluation, explained, fig, greater, horizontal, illustrated, indication, maintained, rate, summarised, version, vertical

Although the studies conducted by Barber, Inman and Farrell are also corpus based and use the criteria of frequency and range, Barber's corpus is in the discipline of Electronic Engineering, Biochemistry and Instrumental Optics; Inman's study is based on ten disciplines: Biology, Mathematics, Physics, Chemistry and Chemical Engineering, Geology, Mining Engineering, Electrical Engineering, Civil Engineering, Mechanical Engineering and Metallurgical Engineering; and Farrell's study is in the area of Electronics (including basic electricity).

5.3 Summary of chapter

This chapter examines the characteristics of technical and subtechnical vocabulary beyond their frequency and range patterns. Technical vocabulary with a range of one was found to be key technical terminology in the sense that it represents the topicality of the module; whereas technical words with a wider range were found to represent notions relevant to several disciplines. My subtechnical vocabulary exhibits a number of characteristic features. 12% of the words are used within examples, and 2% is categorized into a 'rag bag' which consists of a mixture of words with different features, but the majority of my subtechnical vocabulary (86%) is found to have a procedural, discourse organizing function. Some of the procedural vocabulary that fall into my subtechnical list also occur in lists drawn up by writers who did not identify them on the basis of frequency and range. This suggests that the criteria that used in this research are a valid means of grouping words which shared semantic or metadiscursive properties.

Chapter Six

Study One and Study Two: Students' Receptive Knowledge of Technical and Subtechnical Vocabulary

6.1 Introduction

Chapter Three describes a small scale interview to investigate postgraduate Engineering students' vocabulary learning problems. Some students particularly mentioned that technical vocabulary presented problems in their reading of the textbooks assigned by their subject lecturers. They attributed this to a lack of background information. Different types of technical vocabulary were identified as not being recognized: some of the technical vocabulary was IT related, and some was EBM related. The students believed that they could not recognize the technical vocabulary because they had not studied a similar subject before (see Chapter Three).

This information conflicts with the received wisdom of ESP, which identifies subtechnical rather than technical vocabulary as problematic for ESP learners. ESP learners are believed to have acquired a broad base of General English words during their secondary education, and it is claimed that they can learn technical terms directly from their subject tutors and textbooks (Nation 1990: 19).

ESP practitioners commonly claim that teachers should not attempt to teach technical vocabulary. Higgins (1985) states firmly that 'It is not the job of the English teacher to teach technical vocabulary.' Moody (1975) expresses the same thought: 'One cannot and should not expect the ESP teacher to deal with levels

1-4, since it is quite unrealistic to expect that he will know enough to deal with them. If he does try to deal with these levels, he will arouse amusement, if not contempt, in his students: they know far more about such things than he does.' (Moody posits six levels of vocabulary, levels 1-4 relate to specialist vocabulary.) Robinson (1980) supports this view, by recommending that coursebook writers do not need to concentrate on specialist vocabulary, as students will absorb it from their main course studies. Trimble (1985: 128) also states explicitly that 'Technical vocabulary by itself does not pose enough of a problem for the majority of non-native students to need special attention in the classroom.' Like Moody, he considers it pointless for a teacher not trained in science to 'teach' technical vocabulary to students who have already learned or are learning this highly specialised lexis in their subject matter courses. All these claims suggest that it is a firmly established fact that technical vocabulary does not present any problems for ESP learners, as learners will automatically learn these technical words and concepts from their tutors, coursebooks and lectures.

The interviews described in Chapter Three, however, contained students' comments which were not entirely in line with the received opinion. I decided to investigate students' knowledge of technical and subtechnical vocabulary in order to have a better knowledge of whether technical vocabulary had indeed been acquired more readily than subtechnical vocabulary. Two studies were conducted for this purpose.

6.2 Study One

6.2.1 Introduction

Many vocabulary tests were reviewed in Chapter Two. From the range of tests reviewed in that chapter I selected a method which enabled me to investigate students' receptive knowledge of technical vocabulary and subtechnical vocabulary. I decided to use a vocabulary test similar to that of Meara and Jones (1987) (see Chapter Two).

6.2.2 Procedure

The process of designing the test can be divided into three stages: deciding on the test format to be used; choosing the words to be tested; and implementing the test (including planning and administration).

Several test methods were considered before I decided the right method for my purpose. One of them was open-ended questions. In the case of this test, I considered presenting a list of words by asking 'what does X mean?' The difficulty lies in evaluating the answer. Evaluations would be particularly difficult with the highly technical terms, as in many cases an Engineering specialist would need to be consulted to interpret the answers. A further problem with this kind of test is that lower level learners would be disadvantaged because it requires productive skills, such as defining and explaining.

A multiple choice test was also considered. Gui (1987) conducted a study using this format to test Chinese learners' vocabulary size. For each English word, he presented the learner with a choice of four equivalents or definitions in Chinese. This method did not prove successful, as Gui lacked, first of all, a good criterion for deciding what a word was, and his choice of words included proper nouns and different inflected forms of words. The method was not feasible for my studies because my subjects had learned their specialism through the medium of English. For this reason they were more likely to know words in English than in their first language (Cantonese). Another reason that this method was unsuitable was that it would be particularly difficult to translate subtechnical words, many of which do not correspond on a one-to-one basis with Chinese words.

The alternative to L1 definitions was to offer multiple choice definitions in English. One very obvious disadvantage of this was that a learner can fail by not understanding the multiple choice alternatives rather than because of ignorance of the target word. On the other hand, it is also possible for a learner to answer correctly without knowing the target word because he has all the necessary knowledge to exclude the distracters, despite the fact that he lacks knowledge about the item being tested.

The third format that I considered was words in context. In this kind of test a reading passage is presented to the testee, with the target words either underlined or italicised. The testee is asked whether he or she knows a target word, or is asked to paraphrase it, or give a synonym.

The length and selection of paragraphs are important in this case. Several reliable but relevant items need to be selected within a relatively short passage. The present vocabulary test would have to include a lot of paragraphs if this test method was used, in which case it would take the subject too much time to be feasible and exhaustion might be a strong factor affecting subjects' performance. A further disadvantage of this type of test is that in order to know the meaning of a word in its context, the context itself has to be understood, therefore it involves reading ability to quite a high degree.

The fourth format that I considered was the cloze test. In this type of test a reading paragraph is presented to the testee, with gaps left for completion by the testee. The problem with this kind of test was that it involves testing reading ability to a high level. A student has to understand the context before he or she can complete the gap with a correct word. Another problem was the same as that for 'words in context', many passages would need to be used to test knowledge of an adequate sample of technical and subtechnical words.

A different test format was the yes/no format, in which the question is simply set as 'Do you know X?' Subjects are expected to read a list of words and tick the ones they know. The problem with this test format is that it relies on the testees' honesty. If subjects claim they recognise a word, we do not know whether they are telling the truth. It is possible they are ticking more words than they really know just to get a better score. It is also possible that they believe they know certain words when as a matter of fact they do not. To enhance the reliability of the test results, imaginary words can be added to the authentic word list. This method was first used by Meara and Jones (1987) in testing non-native speaking students' vocabulary size (see Chapter Two).

When imaginary words are used they function as a detector in that the number of imaginary words subjects claim to recognize determines the reliability of their scores. If they claim to know a high proportion of imaginary words, this casts doubt on the reliability of their answers regarding the authentic words; if they claim they know none of the imaginary words, then the results are highly reliable. Scores can be adjusted by a formula, according to the proportion of imaginary words the subject claims to know (Meara and Jones, 1987).

The level of honesty was expected to be higher in my study than in the circumstances where Meara and Jones' test is normally used, because their test serves as a placement test while the studies reported in this thesis were conducted for research purposes only. The students were well aware of the fact that their performance on the test would not affect their academic success in any way.

A yes/no test has obvious advantages over the open-ended question, multiple choice, words in context and cloze test formats. It shares the same qualities as open-ended and yes/no tests in that each word appears in isolation, but it does not present the difficulty of evaluating the answers as an open-ended question does, neither does it provide confusing distracters as a multiple choice test does. It is better than words in context and cloze tests because these two tests involve too much testing of reading abilities, which was not the purpose of this present test. A yes/no test does not take long to write, because the process simply consists of selecting word items, and it takes only a short length of time for the subject to complete the test. I therefore decided to adopt the yes/no format for the test in this study.

I chose the sample test items from my small corpus by applying the same frequency and range policies described in Chapter One of this thesis. The test items included both technical and subtechnical items. Technical words were selected according to the sub-specialisms of each experimental subject. Technical words had, by definition, a low distribution, and it would have been unfair to test subjects on their knowledge of words that did not occur in the course modules they had opted to study. Ten technical words were selected to represent each module, and each subject was tested on words from each of the twelve modules which constituted the Msc. course requirement. Like any sampling, the ideal situation would be to present the subjects with a list which includes all the words. However, in cases where this is not possible, the best alternative would be to choose the ten most representative words.

Once again, words for this part of the test were selected for their typicality, or in other words their high frequency in a low range of texts. It was apparent, however, that several modules did not provide sufficient instances of typical technical words. For this reason the technical words selected to represent 'Simulation of Production Systems' and 'Basic Computing', for example, had a higher distribution and a lower frequency in individual samples than the technical words for 'Polymer Materials, Processes and Products' and 'Steel and its Heat Treatment'.

Subtechnical words occurred across a wide range of module samples, so the same 120 words were used with every subject in this part of the test. The words which were selected from the subtechnical list for use as test items were those with typical subtechnical characteristics: they all had a range of at least five and a frequency of at least ten. They displayed markedly procedural characteristics.

A decision was made not to include the most frequent fifty six words of the subtechnical word list because it was considered important that both technical and subtechnical words occurred with the same degree of frequency within the corpus as a whole, although the subtechnical and technical words chosen as test items had clearly distinct distributional patterns across the corpus.

The next stage involved making up imaginary words. It was considered important that the imaginary words should be plausible, and that they should obey the phonological, graphological and morphological rules of English. The list of imaginary words was created in consultation with a native speaker of English, who confirmed their plausibility. Some of the invented words were certainly more plausible than others. For example, some invented words such as *windstore* or *rowball* could be figurative coinages, while words such as *integratory* could be a question of arbitrary usage, whereas other invented words such as *ultivistic* seem much less likely coinages. However, the results show that this kind of inconsistency does not present a problem because there were very few false alarms.

The following procedure was involved in making these words: in some cases one morpheme in a real English word was changed, for example, *-pect* in the word *respect* was retained, and *res-* was changed into *dis-* to form **dispect*. In the same way I made up **plausive*, with *plau-* retained for the word *plausible*, and *-sible* changed into *-sive*. In other cases, I used new combinations of real English morphemes, for example, I took *peli-* from the word *pelican* and *-ographic* from the word *geographic*, and combined the two together to have **peliographic*. For the technical section, attention was paid to the length and complexity of the non-words, so that they appeared to be similar to the authentic words. For example, in the module 'Metallic Materials Selection', six of the ten real words ended with the phonemes *-ite*: *austenite, martensite, pearlite, ferrite, cementite* and *bainite*, so the made-up words were **aurelite, *antite*, and **lucanite*.

It was important for the imaginary multi-word technical terms to be non-existent English phrases. One way to do this was to make one word in the multi-word technical term an imaginary word, for example, in **peliographic robot* and **quantiple separator*, the first word in both multi-word technical terms were invented. In this way, the multi-word technical terms were bound to be non-existent (see Appendix Four).

6.2.3 The scoring of the test

The test consisted of two parts, a technical section and a subtechnical section. The technical section included sample words from the twenty six modules, with ten real technical words and five imaginary words in each module, while the subtechnical section consisted of a list of subtechnical words; 120 real subtechnical words were used together with sixty imaginary words. The same formula was used to penalize subjects who claimed they recognized imaginary words.

The design of the test involved deciding on the best way to calculate the subject's score taking into account the number of modules each student has taken and the number of imaginary words he or she claims to recognize. An illustration of the test format is shown in Table 6.1.

Table 6.1 Example format of the test

Please tick (√) the words or phrases you know the meaning of, for example, *experiment*, or *information technology*.

torce	lubricant	tangential
shear	sedulant	velocity
rake	clearance	wedge
entrave	incudency	tungsten
deformation	orthogonal	circumhinge

Table 6.1 shows that the test yields two types of response: 1) Real Yes (RY) which means a subject claimed that he or she knew a word which was a real English word, in which case he or she would gain a point, while 2) Imaginary Yes (IY) means a subject claimed that he or she knew a word which was not a real word but an imaginary word, in which case he or she would be penalised.

Ideally, a subject who knew all these words would have responded YES to all the real words, and should leave the imaginary words unmarked. However, in practice, there might be some IY responses, indicating that the subject claimed to know some imaginary words. These figures allow a check to be made on how truthful the subject is. For example, if a subject claimed to know all the real words and rejected all the imaginary words, this would indicate that the subject does know the real words. However, if a subject claimed to know all the real words, but also claimed to know all or some of the imaginary words, then it would be obvious that he or she is not very honest and not very sure of the words. Suppose, for example, that the subject claimed to know 50% of the real words, but also claimed to know 20% of the imaginary words. The IY score indicates that his or her 50% of the RY score would need to be adjusted downwards. To do this, Meara and Jones (1987) and Meara and Buxton (1987) suggested the signal detection theory proposed by Kling and Riggs (1971), and later developed into a formula by Anderson and Freebody (1983):

$$P(k) = \frac{P(h) - P(fa)}{1 - P(fa)}$$

P(h) (the probability of making a 'hit') is the proportion of real words that the subjects recognise (RY). P(fa) (the probability of a 'false alarm') is the proportion of imaginary words the subject claims to know (IY). The formula adjusts the RY score downwards if IY is large. P(k) in signal detection theory represents the likelihood of a real target actually being acknowledged: in the present study it indicates how many of the target words the subject can be deemed to know. For a more detailed description of the formula, see Meara and Jones (1987) and Meara and Buxton (1987).

Two steps were involved for the administration of the test: a trial on two native English speakers, and administering the test to the non-native speaker Engineering students. The test was piloted on two native speaker specialists. One of them was a teaching assistant in the Engineering Department, whose main research area was Civil Engineering. She completed twenty one modules of the test, the score being 96%. For the subtechnical section, the score was 100%. The other specialist involved was a postgraduate computer scientist. He completed six modules, the score being 100%. He did equally well for the subtechnical section of the test. There were no false alarms in either section for either of the two subjects. The results of the two native subjects indicated that their knowledge of the subject specialism and the subtechnical words enabled them to have good scores for both sets of words.

Fifty Msc. Engineering students took the test. Twenty were from Hong Kong, and thirty were from Thailand, Taiwan and Malaysia. Although these students might have had different language exposure and learning experience before they came to the UK, the reliability of the test was not affected because none of the students had become an expert yet in any one particular module, and all of them were studying the same course when the test was taking place. It should be noted that this kind of mix of background is typical of this programme, and every year the intake to the programme is much the same.

The test was administered to them at the end of the academic year. The subjects did the test in their own rooms. No dictionaries were allowed. On average it took ten minutes to complete the test.

6.2.4 Results

The formula proposed by Meara and Jones (1987) was used to adjust the scores, but in fact little adjustment proved necessary, as there were remarkably few instances of 'false alarms' in the data: 6% for subtechnical vocabulary and 7.7% for technical vocabulary. Table 6.2 presents the results of the vocabulary tests.

Table 6.2 Comparison of subtechnical and technical scores

	Mean score	SD
subtechnical words	94.29	8.42
technical words	76.40	13.52

It can be seen that subjects knew many more subtechnical than technical words. This difference was highly significant ($p<.01$). The results of this study will be discussed in section 6.3.4 together with those of Study Two.

6.3 Study Two

The findings of Study One suggested that the postgraduate Engineers had better receptive knowledge of subtechnical vocabulary than that of technical vocabulary. 76.40% of the technical words were recognized, which suggested that 25% of the technical words were not acquired by the end of the course in the manner suggested by some EAP practitioners (see Chapter One). These findings therefore, challenge the common belief that technical vocabulary is naturally acquired by students from exposure to the course. These findings raised several questions, regarding the research design of Study One.

6.3.1 The choice of test items: involvement of the subject lecturers

The first issue that needs to be considered is how well technical items were represented in the first study. The working definition for these two categories of words was already described and discussed in Chapter Four (see 4.2.1). Subtechnical words were defined as those high frequency words which occurred across disciplines, while technical words were those words with high frequency but very low range. However, some words with higher distributions were also identified as being technical, because they had a peak frequency in a particular module.

After completing Study One I developed the large corpus (see Chapter Four) which enabled me to make a wider choice of words for Study Two. I also needed to ensure that the technical words I chose for the test were truly representative of the technical words encountered in the Msc. Engineering programme, as suggested by their frequency and range patterns in the corpus. In order to have a better knowledge of what the Engineering lecturers would consider a representative technical word, I contacted twenty one subject lecturers in the Engineering Department. These subject lecturers covered the teaching of the twenty six modules under analysis.

Each subject lecturer was presented with a list of technical words identified from the module (or modules) he or she had been teaching. These technical words had been identified according to the same criteria of high frequency and low range; the lecturers were then asked to tick the ten technical words that they themselves regarded as 'technical'. All the lecturers had been teaching on the same course for many years, and their experience could be used as a reliable way to identify words that were most central to the module topic. The following excerpt shows the wording of my request:

'Please tick ten technical terms from each Table that you think are typical of, and best represent the module' (see Appendix Five for the complete letter of request to a subject lecturer).

Most of the typical words identified by the subject lecturers had a relatively low frequency of occurrence in the module. Table 6.3 shows the frequency patterns of the words identified by the subject lecturers.

Table 6.3 Frequency patterns of the words identified by the subject lecturers

Frequency	Number of technical words	Number of words with distribution of one	Percentage in total (%)
000-025	148	85	59
025-050	68	13	19
050-100	26	3	12
100-150	5	1	0
150-200	7	0	0
200-250	4	0	0
300-350	1	0	0
350-400	1	0	0
Total	**260**	**101**	

Of the 260 words identified by the subject lecturers, 148 words have a frequency of twenty five occurrences or lower in my corpus. These words constituted 59% of the total number of technical words identified. Table 6.4 shows the range patterns of the words chosen by the subject lecturers. Of the total number of 260 words identified by the subject lecturers, 107 had a range of one, this constituted 41%. The words that had a range of one to five made up 87% of the total number of words. This shows that most of the words recognized by the Engineering specialists as central to their module also had a fairly low range distribution. The subject lecturers' decisions regarding what typified a technical word agreed with my own although the subject lecturers were selecting words on the basis of meaning rather than frequency and range distribution.

Table 6.4 Range patterns of technical words identified by the subject lecturers

Range	Number of technical words	Percentage in total (%)
01	107	41
02	48	18
03	30	12
04	25	10
05	15	6
06	8	3
07	8	3
08	2	1
09	5	2
10	7	3
11	1	0
13	2	1
15	2	1
Total	**260**	

6.3.2 The vocabulary test

The nature of the yes/no test and reason for using it have already been discussed in section 6.2. One of the advantages of the yes/no format is that it takes relatively little time to administer. A disadvantage of this format, the dependence on the experimental subjects' honesty, can be overcome by applying a formula so that dishonest subjects can be penalized and their scores can be adjusted downwards.

The same amount of words were included in the test: ten technical words for each module. Each student would thus be tested on 120 words if he had been following twelve modules as is required by the Engineering Department. As in Study One, the technical words to be tested were module dependent, in other words, the set of technical words varied according to the study plan of the testee. 120 subtechnical words were also included. Subjects were tested on the same set of subtechnical words from the large corpus because subtechnical words occur across a range of modules, and should therefore be equally well known to all testees, regardless of the modules they had studied.

The difference between the set of technical words in Study Two and Study One was that having chosen words for each module in the large corpus using the criteria of frequency and distribution, the Engineering lecturers were also consulted. The effect of consulting the Engineering lecturers was to ensure that the technical terms chosen for the test not only had a recognized frequency and range pattern, but were also central to the topic of the module(s) in which they occurred. Of the 260 technical vocabulary included, fifty seven of them overlapped with those in Study One.

Like technical words, the selection of the 120 subtechnical words for the present study was also based on frequency patterns. The subtechnical words chosen had a similar frequency of occurrence as the set of technical words. This was considered important in choosing the test items. The results of the test would be biased if the frequency of occurrences of the subtechnical and technical words was substantially different, because if one type of words is more frequent in the corpus, it is likely to be known better by the testees, therefore the score will be higher. The comparison of the two types of words would therefore be based on different grounds.

Table 6.5 shows the comparison of frequencies for both sets of words. These are expressed as percentages in Table 6.6. Both technical and subtechnical words for Study Two fall at the low frequency end of the corpus. Although a rough measure,

Table 6.6 shows that the frequency for both sets of words was not substantially different.

Table 6.5 Frequency patterns of technical and subtechnical words for Study Two

Frequency	Number of technical words	Number of subtechnical words
000-025	148	69
025-050	68	34
050-100	26	13
100-150	5	0
150-200	7	4
200-250	4	0
300-350	1	0
350-400	1	0
Total	**260**	**120**

Table 6.6 Percentage of technical and subtechnical words for Study Two

Frequency	Technical words	Subtechnical words
000-025	57%	58%
025-050	26%	28%
050-100	10%	11%
100-150	1%	0
150-200	2%	3%
200-250	1%	0
300-350	1%	0
350-400	1%	0
Total	**260**	**120**

The same set of imaginary words was used in Study Two as in the previous study. All these words were plausible and obeyed the phonological and morphological rules of English. Consultation with a native English specialist confirmed their plausibility (see 6.2). Five imaginary words were placed in each module, making up a total of fifteen words for each module.

The test still consisted of two parts, a technical and a subtechnical section. For the test format, see section 6.2. Section One of the test included sample words from the twenty six modules, with ten real technical words and five imaginary words in each module, while Section Two consisted of a list of subtechnical words; 120 real subtechnical words were used together with sixty imaginary words. The same formula was used as in Study One to penalize subjects who claimed they recognized imaginary words.

Sixty one non-native subjects took the test. Twenty six were from Hong Kong, and thirty five from Thailand, Taiwan and Malaysia. All of them were studying the same Msc. Engineering course. As in the previous study, this study was also conducted at the end of the one year taught course. The subjects did the test in their own rooms. No dictionaries were allowed.

6.3.3 Results

Like Study One, few instances of 'false alarms' were found in the data: 6.6% for subtechnical vocabulary and 7.9% for technical vocabulary. Table 6.7 presents the results of the Vocabulary Test for Study Two.

Table 6.7 Comparison of subtechnical and technical scores

	Mean Score	**SD**
subtechnical words	97.72	6.65
technical words	83.32	10.93

There was a highly significant difference ($t=8.79$, $p>0.001$) between scores on technical and subtechnical words.

6.3.4 Discussion

The results of both studies suggested that the Southeast Asian postgraduate Engineering students had poorer receptive knowledge of technical vocabulary than subtechnical vocabulary. As mentioned in 6.3.2, of the 260 technical vocabulary included in Study Two, fifty seven overlapped with those in Study One. Of the 406 different technical words that appeared in both tests, only 117 of them were recognized. These results were especially surprising because most technical words included in my studies were identified by subject lecturers as being important and therefore representing the topicality of the modules.

I contacted twenty one subject lecturers for comments on the results of Study One. They were asked first of all, to suggest any reasons why participants failed to learn certain technical words; secondly, whether these words were important for understanding concepts of a particular module and for post module work. Finally they were encouraged to make any comments regarding the results of the study. Eleven responses were received regarding eleven modules: 'Industrial Engineering', 'Information Systems Strategy', 'Metallic Materials Selection', 'Automation and Robotics', 'Basic Computing', 'Computer Language', Information Technology Fundamentals', 'Applied Statistical Method', 'Manufacturing Strategy', 'Financial Analysis and Control Systems' and 'Quality Reliability Maintenance Systems'. The subject lecturers commented that most technical terms included in the test were important.

The following are relevant examples of the subject lecturers' comments regarding 'Industrial Engineering', 'Information Systems Strategy', 'Metallic Materials Selection', 'Automation and Robotics' and 'Basic Computing'.

Seventeen students studied 'Industrial Engineering'. Six of them did not recognize the word *remuneration*, while two of them did not recognize *resource utilization*. The subject lecturer commented, 'The (two) words/phrases are important in the context of the module'.

For 'Metallic Materials Selection', six words: *martensite, pearlite, ferrite, carbide, cementite* and *bainite* were not recognized by four, three, two, two, four and three students respectively, out of a total number of fourteen students. The subject lecturer commented: 'Not knowing these six (words) would result in poor post module work'.

Thirty students studied the module 'Automation and Robotics'. Twenty four of them did not recognize the word *resolver,* and twenty one of them did not recognize *incremental encoder;* nineteen did not recognize *articulated mechanical systems*, and nine, five, four and three students did not recognize *end effector, vector, actuator* and *gripper* respectively. The subject lecturer commented: 'Those words (the words listed above) which the highest number of students failed to recognize are important technical words'.

Regarding the fact that two of the six students did not recognize the word *central processing unit* from 'Basic Computing', the subject lecturer commented, 'everyone on the course should know by the end of the course what *C.P.U.*(*central processing unit*) is'.

All these comments show that the technical words included in my studies were regarded as important by subject lecturers. In an attempt to explain why these words were not recognized, the lecturer who was teaching 'Programming Language' suggested that the students 'did not pay attention during the module', 'they did not ask for these words to be explained if they did not understand' or 'they were not motivated'.

These comments may contribute to an understanding of why technical vocabulary was not learned during study of the subject disciplines, and are interesting areas to research. There may, however, be another factor contributing to the students' problems with technical words. The lecturer teaching 'Quality Reliability Maintenance Systems' commented, 'I would normally use alternative words than the ones mentioned during the module'. What he meant was that for the same technical concept, different terminology was used in teaching and in the textbook. This comment was in response to the finding that out of a total number of forty five students, sixteen did not recognize *repairable system*, ten did not recognize *life cycle cost* and *rate of occurrence of failures*, and three did not recognize *mean time between failures* and *probability*. The subject lecturer did not identify what alternative terms were used in place of the terms mentioned above.

The use of alternative terms does not help the students to learn key technical vocabulary, instead it may cause a certain amount of confusion, depending on the degree of difference between the terms used by the subject lecturer and the terms used by textbook writers. Even if the terms used by the two parties are not substantially different, the students may still be uncertain whether the terms encountered in the textbooks are the same ones taught by the subject lecturers. It might help the students if the subject lecturers equated terminology used in the lecture theatre with that used in the textbook.

Another possible reason why technical vocabulary was not recognized is related to the range distribution of these words. Altogether 283 technical words (out of the total of 406) were not recognized in Study One and Study Two. 178 of these had a range of one, this constitutes approximately 60% of the total number of words not recognized. The pattern that emerged from both studies is that technical terms in the **EBM** component were better recognized than those in the **IT** or some of the **MSE** components.

Table 6.8 The frequency and range pattern of the word *supplier*

	Frequency
Information Systems Strategy	28
Manufacturing Strategy	27
Strategic Marketing	9
Industrial Engineering	7
Simulation of Production Systems	7
Information Technology Fundamentals	5
Financial Decision Making	4
Project Planning Management and Control	4
Production Planning and Control Systems	3
Financial Analysis and Control Systems	3
Logistics and Supply Chain Management	2
Quality Reliability Maintenance Systems	2
Metallic Materials Selection	1
Total	**102**

As described in Chapter Six, modules in the **EBM** component had more words with a wider range, which means that many words did not occur in only one module, but in two or more than two modules. *Supplier*, for example, occurred 102 times across thirteen modules, it had a peak distribution in two modules:

twenty eight occurrences in 'Information Systems Strategy', and twenty seven occurrences in 'Manufacturing Strategy'; the frequency in other modules varied, as shown in Table 6.8.

Other words such as *behaviour* and *personnel* were also identified as technical because of their peak frequency distributional patterns. *Behaviour* had a frequency of 171 across fourteen modules, with peak frequencies of seventy six, forty nine and fourteen in three modules. Similarly, the word *personnel* had a frequency of fifty three, with two peak frequencies of thirty and eleven in two modules and relatively lower frequencies in thirteen other modules.

These words had an extremely wide range. This kind of word, although identified as technical in the modules in which they peaked, was not chosen to be included in my studies. Words with a lower or more moderate range met my criteria for selection of the test items, because they were considered to represent the modules better. For example *assembly line* was a better choice than *behaviour* or *personnel*, because it had a range of six, with a peak frequency distribution in 'Human Factors in Industry'. Therefore this word was chosen to represent this module. Only one student out of thirty three did not recognize this word. Although with a much lower range than *supplier*, *behaviour* and *personnel*, the range pattern of this word enabled a student to encounter it in the study of several modules.

In contrast, twenty four out of thirty students did not recognize the word *resolver* and twenty one out of thirty did not recognize *incremental encoder* for the module

'Automation and Robotics'. Both these words had a range of one. This seems to suggest that the lower the range the more likely it will be that students will fail to recognize it.

As described in Chapter Five, the range patterns of modules that belonged to the MSE component displayed two characteristics. Three modules: 'Metallic Materials Selection', 'Polymer Materials, Processes and Products' and 'Machine Tools' had a range pattern similar to the IT component. These modules contained highly subject specific words such as *polycarbonate*, *polystyrene*, *polypropylene*, *polyethylene*, *polyester*, *polyamide* and *polyurethane*, which were found only to occur in 'Polymer Materials, Processes and Products'. This means that students did not have much chance to encounter these words outside this module.

Two modules in the MSE component 'Logistics and Supply Chain Management' and 'Industrial Engineering' had range patterns more similar to the EBM modules. *Inventory control* and *shop floor control* in 'Logistics and Supply Chain Management', for example, tended to have a wider range. The first of these occurred in five different modules, and the second occurred in three modules. Forty six students recognized these two words while only one student failed to recognize them.

It seems that low range affects students' receptive knowledge of technical vocabulary. As mentioned in Chapter Five, the Msc. Engineering course was represented by different modules, and each module was represented in most cases by recommended textbooks. The ideal situation is that students on this course will read the textbooks assigned by their lecturers. If they learn the concepts expressed in the textbooks they will also learn the important technical words that appear in the books. If, however, a student does not read the only book in which certain technical terms occur, then he or she will have to learn these words from face to face teaching in lectures and seminars, and possibly in handouts (assuming that 'alternative terms' are not used as described in 6.3.4), otherwise he or she will not have the chance to encounter these technical terms, and naturally will not learn them. Neither will he or she learn the concepts associated with the terminology.

6.4 Summary of findings

This chapter describes two studies, the results of which are mutually supportive and suggest that the Southeast Asian postgraduate Engineering students had better receptive knowledge of subtechnical than technical vocabulary. These findings do not agree with the common belief among ESP practitioners that technical vocabulary is learned through studying the subject specialism. Possible reasons for not learning the technical vocabulary were, for example, that some subject lecturers used different terms from those used by textbook writers, which might cause confusion to the students; another reason was that most technical words included in my studies had a range of one, which means that these words occurred in only one textbook sampled in my corpus. Chapter Seven investigates other possible reasons why technical vocabulary had not been learned by students at the end of their course.

Chapter Seven

Study Three : The Lexical Familiarization of Technical Words

7.1 Introduction

Chapter One discusses the technicality of technical terms. Technicality suggests the complexity of the formation of technical vocabulary. The use of technical terms may make the reading of technical textbooks difficult. Therefore it is important for textbook writers to explain, define or use any means to facilitate the understanding of technical vocabulary, for example by the use of lexical familiarization strategies (Bramki and Williams, 1984).

Chapter Six describes two studies which set out to examine postgraduate Engineering students' receptive knowledge of technical and subtechnical words. The results of both studies support each other and suggest that students' receptive knowledge of technical vocabulary was not as good as their knowledge of subtechnical vocabulary. The results show that a substantial proportion of technical words was not recognized. There are a number of possible reasons why these technical words were not recognized. Some of the technical words in the test were different from the words used by subject lecturers to refer to the same concepts. Also most technical words included in the test had a fairly low range; if the students had not read the limited number of texts in which these words occurred, they would have little chance of encountering these words.

This chapter examines whether the technical words that were not recognized by the students in the two studies were explained adequately in the textbooks when they were first introduced. My purpose in examining lexical familiarization in this chapter is to investigate whether lexical familiarization had an effect on students' receptive knowledge of technical words, as measured in my two studies reported earlier (see Chapter Six). I wanted to see whether lexically familiarized technical words were better recognized than those that were not familiarized in the textbook. Thus I needed to, first of all, identify which technical words were lexically familiarized, and which ones were not. The classification of the lexically familiarized technical words into specific categories is not my primary concern in this chapter.

7.2 Criteria for identifying occurrences of lexical familiarization

Bramki and Williams (1984) identified six categories of lexical familiarization: 'exemplification', 'explanation', 'definition', 'stipulation', 'synonymy' and 'illustration' (see 1.4). My analysis of lexical familiarization of the technical words was based on this categorization, with some modifications, because some of their instances of lexical familiarization were found to be inconsistent. For example, *is known as* and *refer to* are the signalling devices for both 'explanation' and 'definition'. Their criteria make it difficult to distinguish familiarization expressed through 'explanation' and 'definition'.

Williams (1980) points out that lexical familiarization is more effective if textbooks use a recognizable and consistent typographic system that indicates to the reader which lexical items are being introduced. Several typographic devices, for example italicisation, underlining or bold type can be employed for this purpose. One of the most recent studies to examine lexical familiarization is by Handee (1996). Handee identified technical words as being lexically familiarized when they were italicised, this was because italicization was the signalling device used in the textbook she used as her corpus. For this thesis, I will count as instances of lexical familiarization those explanations accompanying words in italics, bold typefaces and also inverted commas. These are the most typical typefaces for the introduction of new words used in the Engineering textbooks in my corpus. Of the twenty six books, two books do not use a different system for technical words. **Steel and its Heat Treatment** (Thelning, 1984) only uses the italic typeface for the word *Figure*, this can be seen from Example 1:

> *Example* 1:
> ***Figure* 1.14 illustrates a simple model of how we can picture the transformation of γ-iron during martensite formation.**
> **(Steel and its Heat Treatment, Thelning, 1984: 15)**

And **Computer-aided Design and Manufacture** (Groover and Zimmers, 1984) only uses the bold typeface for subtitles such as **Historical Background**. I did not examine instances of lexical familiarization in these two books, because I did not have any consistent way of recognizing when the author intended to familiarize a technical word.

Strategic Planning for Information Systems (Ward and Griffiths, 1990) uses a different typeface from other books. Instead of using a different typeface for the newly-introduced technical word, the authors of this book use italic typeface for the clause which either defines or explains the technical word. Example 2 and 3 illustrate this:

Example 2:
Strategy can be defined as:

An integrated set of actions aimed at increasing the long-term well-being and strength of the enterprise relative to competitors.
(**Strategic Planning for Information Systems**, Ward and Griffiths 1990: 52)

Example 3:
A business unit can be defined as '*a unit that sells a distinct set of products or services, serves a specific set of customers and competes with a well-defined set of competitors*'.
(**Strategic Planning for Information Systems**, Ward and Griffiths 1990: 64)

In these examples, *strategy* and *business unit* were the terms introduced, and the parts in italics were instances of lexical familiarization.

In **The Complete FORTH** (Winfield, 1983) the author uses two different typefaces for the newly-introduced technical word. Inverted commas and italics are both used, this can be seen from Examples 4 and 5:

Example 4:
In reverse Polish each arithmetic operator comes after the numbers upon which it operates (termed 'operands')
(**The Complete FORTH**, Winfield, 1983: 3)

Example 5:
{<} replaces these two numbers by a single number called a *flag*, which may only have one of two values, true or false.
(ibid: 31)

In some cases, although a technical word is italicised, it is not lexically familiarized by the writer. Example 6 shows:

Example 6:
The previous program outline is more graphically expressed in a *syntax diagram*.
(**Pascal: User Manual and Report**, Jensen and Wirth 1985: 3)

This is the second occurrence of *syntax diagram* and it is in italics. The first occurrence is a subtitle: **The Syntax Diagram**, which is in the bold typeface. In most of the other occurrences of *syntax diagram* it is used as part of the caption for figures, for example, *Figure 1.e Syntax diagram for UnsignedNumber.* This use occurs thirty four times.

7.3 Categories of lexical familiarization

The following is a description of the seven categories of lexical familiarization as they occurred in my corpus.

7.3.1 Definition

Definition familiarization in my corpus was achieved by: 'term + class + characteristic', often realized as 'A is a B which C'. For example:

> *Example* 7:
> A *sample* from a statistical population is the set of measurements that are actually collected in the course of an investigation.
> **(Statistical Concepts and Methods,** Bhattacharyya and Johnson 1977: 8)

In some cases, the 'term' did not come first, for example:

> *Example* 8:
> During the item's life the instantaneous probability of the first and only failure is called the *hazard rate.*
> **(Practical Reliability Engineering,** O'Conner, 1991: 4)

> *Example* 9:
> The number of independent movements the object can make with respect to the coordinate set R is called its number of *degrees of freedom.* (DOF)
> **(An Introduction to Robot Technology,** Coiffet and Chirouze 1983: 33)

In both examples 8 and 9, the 'terms' came after the 'class'. Swales (1971) considers these two types of formula to be equivalent in function. However, Francis (1986) points out the difference in terms of the *given/new* information. In cases where the 'term' comes first, the 'term' has been already established as given information prior to the definition, whereas in cases where the 'term' comes last, the term is usually treated as new information following the already given information.

Some definition familiarizations were achieved by means of the definite pronoun *this, these* or sometimes *such* in my corpus, for example:

> *Example* 10:
> A more compact way of representing binary numbers is to use the base 16. These are called **hexadecimal** numbers (often abbreviated to **hex** for short) and their range of digits is from 0 to 15.
> **(Microcomputers in Engineering and Science**, Craine and Graham 1985: 15)

> *Example* 11:
> It is desirable to consider first the simplest type of cutting operation, in which the cutting edge is straight, parallel to the original plane surface of the workpiece, and perpendicular to the direction of cutting, and in which the length of the cutting edge is greater than the width of the chip removed. Such an operation has been described as *orthogonal* cutting.
> **(Modern Workshop Technology**, Baker 1966: 1)

> *Example* 12:
> Once we have collected some facts together, we can ask questions about them. Such a collection is called a *data base*.
> **(Artificial Intelligence**, Rich and Knight 1991: 3)

In examples 10 and 12, *are/is called* was used as the signalling device, while in example 11, *described as* was used. The reader needs to interpret these as an indicator of definition. As can be seen from the above examples, *these, such an operation* or *such a collection* performs the function of anaphoric (backward pointing) reference which refers to a nominal group, a clause, a sentence or a paragraph mentioned earlier. For example, in the first instance above, *these* refers to the nominal group *base 16*, while in the second instance, *such an operation* refers back to the whole preceding sentence. The second instance illustrates the writer's concern with the term *orthogonal cutting*. In order to understand this term, the reader needs to know the role of the anaphoric noun *(such an) operation,* which refers back to the whole preceding sentence. Anaphoric nouns together with the definite pronoun *this, these* or *such* were found to be frequently used by writers to familiarize technical vocabulary in the category of definition.

Definition was the most common form of lexical familiarization, as can be seen from Table 7.1.

7.3.2 Explanation

The 'explanation' type of lexical familiarization explains the meaning of a newly-introduced term in terms of an equivalent or opposite meaning by means of a sequence of words (a phrase, a sentence or several sentences) (Bramki and Williams, 1984). In my corpus most of the explanations were achieved by direct means. The following are some examples from the corpus:

> *Example* 13:
> **Possession utility** means obtaining a product and having the right to use or consume it.
> (**Basic Marketing: a Global Managerial Approach**, McCarthy and Perreault 1993: 6)

> *Example* 14:
> **Micro-marketing** is the performance of activities that seek to accomplish an organization's objectives by anticipating customer or client needs and directing a flow of need-satisfying goods and services from producer to customer or client.
> (ibid: 8)

> *Example* 15:
> Carrying out detailed component plans is sometimes facilitated by including **safety stocks** and/or **safety lead times** in the MRP records. ... Safety lead time is a procedure whereby shop orders or purchase orders are released and scheduled to arrive one or more periods before necessary to satisfy the gross requirements.
> (**Manufacturing Planning and Control Systems**, Vollmann, Berry and Whybank 1992: 33)

In the third example, *safety lead times* is in bold typeface when it appears the first time, the second time it is in the normal typeface.

Some explanations were achieved by contrasting the newly-introduced term with a term that the writer has already explained, for example:

Example 16:
Target marketing says that a marketing mix is tailored to fit some specific target customers. In contrast, **mass marketing** - the typical production-oriented approach - vaguely aims at 'everyone' with the same marketing mix.
(**Basic Marketing: a Global Managerial Approach**, McCarthy and Perreault 1993: 45)

Example 17:
This goal leads to designs in which the system consists of many processors located close together, called a **local network**, to contrast it with the far-flung **long haul network**.
(**Computer Networks**, Tanenbaum 1996: 4)

Some explanations are achieved by operational description, for example:

Example 18:
Ductility. A load-bearing device or component must not distort so much under the action of the service stresses that its function is impaired, nor must it fail by rupture, though local yielding may be tolerable.
(**Plastics Engineering**, Crawford 1987: 24)

7.3.3 Exemplification

Exemplification is achieved by providing an instance or instances of what a newly-introduced term refers to (Bramki and Williams, 1984). Such instances can be concrete items, situations or numbers.

In my corpus, five instances of exemplification were found. The following are two examples:

Example 19:
An *algorithm* or computer program consists of two essential parts, a description of *actions* that are to be performed, and a description of the *data*, that are manipulated by these actions.
(**Pascal: User Manual and Report**, Jensen and Wirth 1985: 1)

Example 20:
The highest ranking term in the hierarchy is *production system*. A production system includes people, money, equipment, materials and supplies, markets, management, and the manufacturing system. In fact, all aspects of commerce (manufacturing, sales, advertising, profit, and distribution) are involved.
(**Materials and Processes in Manufacturing**, DeGarmo, Black and Kohser 1988: 8)

7.3.4 Synonymy

Flowerdew (1992) defines synonymy as the pairing of a lexical item (including multi-word units, phrasal verbs and set phrases) with an item having the same or similar meaning. Only one instance of familiarization by synonymy was found in my corpus:

Example 21:
Next, find a nice family and move the needed machines and work stations into cell 2. Then create cell 3, and so on. Engineers sometimes call this approach *group technology*, although many prefer to use the more descriptive term *cellular manufacturing*.
(**World Class Manufacturing**, Schonberger 1986: 10)

7.3.5 Derivation

According to Flowerdew (1992), lexical familiarization by derivation occurs when the derivation of a word or more often a word part (often from Latin or Greek) is cited to explain the meaning of the term. There are two instances of familiarization by derivation in my corpus:

Example 22:
The other way to do the same thing repeatedly is very much simpler, and is called *iteration*. You've heard of the word 're-iteration', to do with repeating words or phrases? Well, iteration is the same thing applied to a program.
(**Artificial Intelligence**, Rich and Knight 1991: 41)

Example 23:
Ordinary arithmetic notation is often called 'infix' because the operators (+, -, *, / etc.) are often fixed in-between the numbers.
(**The Complete FORTH**, Winfield 1983: 3)

7.3.6 Analogy

According to Bloor (1996), an analogy is characteristically used by textbook writers to introduce new concepts or difficult procedures by reference to something already known. Two instances of analogy were used by writers in my corpus:

Example 24:
In *Japanese Manufacturing Techniques* I told about working for the fastest bricklayer in North Dakota and about how he yelled at me if I didn't place bricks so that he could reach and find them without looking. That concept - exact placement of all the parts to eliminate search - has enabled the world's motorcycle manufacturers and some tractor producers to change models after each unit. That is called *mixed-model production* ...
(**World Class Manufacturing**, Schonberger 1986: 8, 9)

Example 25:
In programming there are two ways in which you can do the same thing repeatedly. One is called *recursion*, and the other is called *iteration*. Beginners often find the distinction difficult quite to grasp, so I'll do my best to explain it clearly. Recursion is what **shorten** did: It actually contained a line of LISP code that made **shorten** happen again inside itself - a part of the execution of **shorten** involved executing **shorten** again. A visual analogy is the effect you get when standing between two parallel mirrors: endless copies of your reflection stretching away to infinity....
(**Artificial Intelligence**, Rich and Knight 1991: 38)

In the first example, the writer explained mixed-model production by referring to a personal experience; while in the second example, the writer explains recursion by giving examples and a visual analogy.

7.3.7 Complex lexical familiarization

I include further a type of familiarization which I call complex lexical familiarization, because it involves a combination or mixture of devices: direct explanation + derivation. There was only one instance of complex lexical familiarization in my corpus:

Example 26:
... *marketing* is both a set of activities performed by organizations and a social process. In other words, marketing exists at both the micro and macro levels. ... The term "marketing" comes from the word market - which is a group of sellers and buyers who are willing to exchange goods and/or services for something of value.
(**Basic Marketing: A Global Managerial Approach**, McCarthy and Perreault 1993: 8, 14)

7.4 Results

In the following discussion I will only refer to the familiarization of technical words which occurred in my corpus five or more times.

Using the criteria presented above, I examined the 406 different technical words that were used in the two studies. 289 of these had not been recognized by one or more students while 117 of them had all been recognized (see Appendix Six and Seven for these words). I did not examine lexical familiarization in the two books mentioned above in which no consistent typeface is used to distinguish the newly-introduced technical words.

Of the 406 technical words examined, 111 instances of lexical familiarization were found. These instances can be divided into seven categories: 'definition', 'explanation', 'exemplification', 'synonymy', 'derivation', 'analogy' and 'complex' lexical familiarization. Table 7.1 shows the frequency of use of the different categories of lexical familiarization for the technical words included in the two studies reported in Chapter Six.

Table 7.1 Categories of lexical familiarization for the technical words included in Study One and Study Two

Category	Number	Percentage (%)
definition	61	56
explanation	39	35
exemplification	5	5
synonymy	1	1
derivation	2	1
analogy	2	1
complex	1	1
Total	**111**	**100**

Table 7.2 shows the frequency of use of the different categories of lexical familiarization for the technical words that were not recognized in both studies.

Table 7.2 Categories of lexical familiarization for the technical words that were not recognized in Study One and Study Two	
Category	**Number**
definition	23
explanation	22
exemplification	2
analogy	1
derivation	1
Total	**50**

Of the 289 words which subjects did not recognize, only 50 (17%) were lexically familiarized.

Table 7.3 shows the frequency of use of the different categories of lexical familiarization for the technical words that were recognized in both studies.

Table 7.3 Categories of lexical familiarization for the technical words that were recognized in Study One and Study Two	
Category	**Number**
definition	38
explanation	17
exemplification	3
analogy	1
synonym	1
complex	1
Total	**61**

Of the 117 words which all subjects recognized, 61 (52%) were lexically familiarized.

7.5 Summary of findings

Of the 406 technical words examined, only 111 words were lexically familiarized, which constitutes 27% of the total number of technical words examined. This suggests that only a small percentage of technical words are lexically familiarized in my corpus. This finding is surprising, as other writers believe that technical vocabulary is familiarized in the early part of textbooks (Bramki and Williams 1984, Trimble 1985). In my corpus, however, there is evidence that lexical familiarization is not frequent in the introductory chapters, because only 27% of the technical words were lexically familiarized. One of the reasons why technical words are not adequately explained by textbook writers may be that writers expect their readers to have acquired background information before they read the book. Example 27 from my corpus shows how the authors of one book perceive their readers:

Example 27:
Much of the following text assumes that you, the reader, have a minimal grasp of computer terminology and a "feeling" for the structure of a program. This purpose of this section is to spark your intuition.
(Pascal: User Manual and Report, Jensen and Wirth 1987: 1)

With the perception that their readers are already familiar with the topics they write about, few of the technical words in this book are familiarized. As Example 6 shows, *syntax diagram* is a newly introduced technical word, yet it is not lexically familiarized.

At this point some of my findings are different from that of Bramki and Williams (1984) and Handee (1996). Both reported that technical terms in titles, headings, and subheadings are usually familiarized. In my corpus, however, many are not. Example 6 is one of the instances: *syntax diagram* is used as a subheading, but it is not familiarized.

It seems that other writers have similar assumptions that the readers might already be familiar with the subject. This could be the reason why many technical words in these books are treated as given information. Here is an example from a different book:

> *Example* 28:
> Growth is not the problem. The problem is the more-of-the-same approach to growth. A restaurant is a little *job shop*, to use the manufacturing term.
> (**World Class Manufacturing**, Schonberger 1986: 5)

This is the first occurrence of *job shop* and it is in italics. Thirteen occurrences of *job shop* are found in the corpus, the following excerpt contains five of the occurrences:

> *Example* 29:
> It will not work if it becomes a big *job shop* - where a job (platter) has to traverse vast distances from one shop to another, waiting for one thing or another at most of the shops. ... Over years we came to believe that stop-and-go production was the fate of the *job shop*. We also believed that *job shops* were the fate of industry, because customers are fickle; they want the variety that *job shops* can provide. *Job shops* people looked enviously at the flow shops, where work just flows down a production line or through pipes continuously.
> (ibid: 5, 6)

Example 29 shows that none of the instances provide contextual help for the reader, simply because the author considers it unnecessary, and he assumes that this term is already familiar to his readers.

Similarly the multi-word technical term *undeformed chip thickness* is also expected to be familiar to the readers already. Example 30 provides the context in which this word first occurs:

> *Example* 30:
> Cutting force does not in general increase linearly with *undeformed chip thickness*. Fig. 1.28 shows that the specific cutting pressure decreases as the undeformed chip thickness increases.
> (**Modern Workshop Technology**, Baker 1966: 33)

As can be seen, the first occurrence of this word is in the italic typeface. It occurs eleven times in the book, the rest of the occurrences are in normal typeface, and it is not familiarized.

In some cases, the writers do not provide adequate contextual aid for the readers. This can be seen from Example 31:

> *Example* 31:
> This goal leads to designs in which the system consists of many processors located close together, called a **local network**, to contrast it with the far-flung **long haul network**.
> (Computer Networks, Tanenbaum 1996: 4)

In order to understand *local network*, a reader needs to know first of all, what a *processor* is, and also how *close* the writer means by *close*. To make it more difficult, a new term *long haul network* is introduced without any familiarization, as the immediate following context shows:

> *Example* 32:
> In addition to a favorable price/performance ratio, local networks have other advantages over a single centralized system. For one thing, they are more reliable, since a single hardware or software failure in a well-designed network will only bring down one processor, and not affect the others.
> (ibid)

The contrast functions as a kind of familiarization. However, if the readers do not have any knowledge of *long haul network*, it is possible that they may find it difficult to understand the new technical word *local network*.

Table 7.2 shows that fifty technical words were lexically familiarized among the 289 words that were not recognized in my studies, this only constitutes 17% of the total number of technical words that were not recognised. And from Table 7.3 we can see sixty one technical words were lexically familiarized of the total 117 words that were recognized, and this comprises 52% of the total technical words that were recognized. Table 7.4 summarizes this finding.

Table 7.4 Summary of lexical familiarization of the technical words included in Study One and Study Two			
		Occurrences of technical words lexically familiarized	**Percentage (%)**
Number of technical words not recognized	289	50	17
Number of technical words recognized	117	61	52

This seems to suggest that there is a greater possibility for a technical word to be recognized if it is lexically familiarized. Most of the words not recognized are not lexically familiarized.

As a summary of this section, my findings seem to conflict with Bramki and Williams' claims that lexical familiarization tends to be more frequent in the early part of textbooks. Of the 406 technical words examined, only 111 of them were found to be lexically familiarized. Table 7.4 of this section seems to suggest that there is more of a tendency for a technical word to be recognized if it has been lexically familiarized. Perhaps one reason why technical words were not recognized in my studies was because most of them (83%) were not lexically familiarized. However, unless other variables such as memory, subject understanding and other aspects of reading skills are examined, the study described in this chapter can only be considered a tentative attempt to investigate the effect of lexical familiarization on the receptive knowledge of technical vocabulary.

Chapter Eight

Pedagogical Implications

8.1 Introduction

As stated in Chapter One of this thesis, my first research question concerns the features of technical and subtechnical vocabulary identified on the basis of frequency and range patterns. Chapter Five examines these features beyond the frequency and range differences of these two types of vocabulary. It was found that although identified according to frequency and range, the majority of the technical vocabulary in my corpus was topic specific and represented the topics of the subject specialism. Technical vocabulary identified using frequency and range criteria proved to contain many words that subject tutors judged to be highly subject specific (see Chapter Six). Although words in my subtechnical list exhibited a number of features, most of the subtechnical vocabulary in my corpus had a procedural or discourse organizing function. Some of the words in my subtechnical vocabulary list also occur in lists of subtechnical words drawn up by writers using criteria other than frequency and range. This suggests that frequency and range criteria arrived at a valid way of grouping these words into semantic/functional categories. This is in relation to my second research question.

The question of whether Southeast Asian postgraduate Engineering students had better receptive knowledge of technical or subtechnical vocabulary was my third research question stated in Chapter One. The results of both studies described in

Chapter Six suggested that the Southeast Asian postgraduate engineers were deficient in technical rather than subtechnical vocabulary.

Chapter Seven discusses the extent to which the technical vocabulary not recognized by my subjects was explained and defined in the textbooks they were recommended to read. The findings of this study suggested that most technical vocabulary was not lexically familiarized in my corpus.

This chapter is concerned with the pedagogical implications of these findings.

8.2 Findings regarding technical and subtechnical vocabulary

Three types of technical vocabulary were found in my corpus. The first type of technical vocabulary (essential to key topics) is the key terminology that the Southeast Asian postgraduate Engineering students should learn through studying the course, because technical vocabulary is particularly important for students who need to read textbooks efficiently. A pedagogical list can be drawn up by using the frequency and range criteria. Section 8.3 discusses this in more detail, and section 8.4 gives practical suggestions of how ESP teachers and subject lecturers can use this list.

The second type of words (used in explanations, descriptions) is called low range example words. These words such as *racket* and *mousetrap* have the same characteristics as technical words in that they have a high frequency and a low range of occurrences. Such words are used in very detailed and lengthy examples, and if the student reader is not familiar with these words, he or she may have difficulties understanding several paragraphs of text, especially in situations where the author and the reader do not share the same cultural background. Such words are difficult to predict, as they may have low frequency in other corpora and may not occur in, for example, in a 5,000 word list. However, in a list of high frequency and low range words they are easy to distinguish from 'true' technical terms, because such words do not reflect the topic of the text. It is therefore relatively easy to extract them and pre-teach these words to the learners who use these textbooks.

The third type of technical vocabulary, General English words used in a technical sense, may cause difficulties to learners. This is also pointed out by Kennedy and Bolitho (1984) and Trimble (1985), who call these words subtechnical vocabulary. These words are more predictable than the low range example words, because they relate to the topic of the module. Such words are also fairly easy to extract by using the frequency and range criteria, and can therefore be pre-taught to the learners using these textbooks.

As mentioned earlier, my subtechnical vocabulary includes several different categories of words. Most subtechnical vocabulary is procedural or metadiscursive. One type of words that has this discourse organizing function is a group of nouns that plays the anaphoric role. Another type is Vocabulary Three items, which signal clause relations. As noted in Chapter Five, these two sets of words are not mutually exclusive. Hedges were also found to signal text organization, evaluate its contents and convey attitudes to readers. Another category of words that falls into my subtechnical vocabulary list is a group of wide range example words.

An anaphoric noun refers back to a stretch of discourse preceding it and is at the same time synonymous to that part of the discourse. Anaphoric nouns were found to be used frequently by writers and are therefore important for the students who use the textbooks. In particular, anaphoric nouns together with *this*, *these* and *such* are used frequently by textbook writers to familiarize technical vocabulary in the category of definition (see Chapter Seven). The anaphoric reference was found to be a nominal group, a clause, a sentence or a paragraph that is mentioned earlier. Understanding backpointing reference will help the students to understand the textbooks. Anaphoric nouns are easy to identify because they belong to an extensive but closed set.

Vocabulary Three items signal clause relations. *Contrast*, for example, has the semantic function 'to contrast' like the Vocabulary One subordinator *whereas* and the Vocabulary Two sentence connector *however*. All these words are used as signals for the clause relation 'comparative denial'. Like anaphoric nouns, Vocabulary Three items also occur frequently in the Engineering textbooks, and therefore students who use the textbooks need to have knowledge of these items. Students need to know that a clause relation may be signalled in the discourse by the use of a Vocabulary Three item. Knowing this helps the reader to predict the oncoming information. It is useful to teach Vocabulary Three items so that the students will know that the word *contrast*, for example, signals the same kind of relation as *whereas* or *however*. In some cases, the students may have a better knowledge of Vocabulary One and Two items than Vocabulary Three. Vocabulary One and Two belong in the most frequent 5,000 English words and therefore should have already been learned during their secondary education. It may help students to recognize the semantic closeness of the Vocabulary Three items to Vocabulary One and Two.

The analysis of hedges indicated the existence of hedging used in the Engineering writing. This suggests a need for greater and more systematic attention to be given to this important device.

Wide range example words have the same function as the low range example words discussed above which fall into my technical vocabulary list. The differences between these two types of example words are their frequency and range patterns (see Chapter Five). These words are not subtechnical in the sense of being procedural or metadiscursive, but they have a wide range, which suggests that writers of different textbooks use the same kind of examples to illustrate concepts. Since these example words are common across textbooks, it is useful for students to know them.

8.3 The application of the frequency and range criteria

Frequency and range have proved to be good criteria for identifying subtechnical and technical vocabulary (see Chapter Six). The frequency and range criteria can be used for many purposes in the teaching of vocabulary to specified groups of students. Having found that the receptive knowledge of technical vocabulary of the

Southeast Asian postgraduate Engineering students is relatively poor compared to their subtechnical vocabulary knowledge, the criteria can be used, for example, to draw up a comprehensive list of technical vocabulary from a larger corpus of all the twenty six textbooks used for the teaching of modules.

8.4 The teaching of technical vocabulary

My third research question stated in Chapter One is whether the Southeast Asian postgraduate Engineering students have better receptive knowledge of subtechnical or technical vocabulary. The results of the studies described in Chapter Seven suggest that these students need help with technical vocabulary rather than subtechnical vocabulary. The findings challenge received opinion regarding the type of vocabulary that should be taught on an ESP programme. They suggest that the terminology traditionally dealt with solely in the subject specialism should be taught or at least reinforced in some way in the ESP classroom. This will not be welcome news for the majority of ESP practitioners. In particular, these findings raise questions of how these students can best be helped to acquire the technical vocabulary they need.

The finding that the students need help with technical vocabulary poses questions for ESP teachers. It is uncertain whether first as to whether they have the capacity and ability to help students to cope with technical vocabulary, as they regard themselves as linguists rather than terminologists. I will start the discussion by considering ESP courses in general, in particular the role of ESP teachers, focusing on some problems that may arise when helping students to cope with technical vocabulary. I will then discuss the pedagogical implications for Warwick engineers and propose some possible solutions.

8.4.1 The role of the ESP teacher

The ESP teacher's role is one of many parts. Swales (1985) prefers to use the term 'ESP practitioner' rather than 'ESP teacher' to reflect the many roles that ESP teachers play. In addition to the normal functions of a classroom teacher, ESP teachers will have to deal with needs analysis, syllabus design, materials writing and evaluation. I do not intend to go into further details of these aspects, instead I will consider the knowledge framework that ESP teachers have to cope with.

As well as teaching the language of ESP, ESP teachers need also to struggle with the subject matter which, in many cases, may be beyond their previous experience. They may find themselves having to refer to texts whose contents they know little or nothing about. For this reason it is possible they may feel a sense of inadequacy and doubt their ability to cope. Hutchinson and Waters (1987: 162) summarise some of the causes of this inadequacy:

> I) there is a tradition in education of separating the Humanities and the Sciences. Languages have usually been allocated to the Humanities camp. The result has been that English teacher often receive little or no education in the Sciences.
> ii) Many ESP teachers are reluctant settlers in the new territory. They would prefer to be teaching Literature and Social English in the comfortable environs of ELT, but have been obliged by economic pressure to emigrate. This does not engender a great desire to learn about the new area.

Sheerin (1981) also points out that lack of confidence is one of the 'real difficulties' and many ESP teachers feel alienated by the subject matter they are expected to teach due to lack of specialist knowledge. Strevens (1988) puts it this way: there is a 'gap between the learner's knowledge of the special subject and the teacher's ignorance of it'.

Some ESP teachers feel that they do not need to learn specialist knowledge. One of the suggestions Strevens gave for ESP teachers was to become familiar with the language of the subject. This suggestion is not as easy as it sounds, however. In almost all ESP teaching situations, there are three parties involved: the ESP teacher, the subject teacher and the students. The subject teacher and the student may have different views regarding the ESP teacher's engagement in their specialism. First of all, if the students expect the ESP teacher to be an authority, they may find it difficult to accept a teacher who is forced to admit ignorance of their specialism. On the other hand, if students expect their ESP teachers only to teach, for example, English grammar or vocabulary, they may well be disconcerted when the ESP teacher appears to be teaching their specialism. Subject teachers may also object to this.

Some ESP teachers do manage to gain some knowledge of the specialism. Adams Smith (1980: 38), for example, writes, 'some of the finest ESP teachers I have met could probably pass their students' subject area examinations very creditably'. Many teachers, however, teach students at a range of levels, and from a range of subject areas. Abbot (1983: 35) asks, 'How many disciplines can the EST teacher be expected to acquire a 'layman's outline knowledge' of? To what level should the teacher strive in each of these? Would such a level of knowledge be achievable? Yet mightn't a little knowledge be a dangerous thing?' Greenall (1981: 24) raises similar questions, 'Can the EST teacher be expected to absorb the values and symbols of science without the thoroughgoing training which scientists have?' Selinker (1979) points out that for ESP teachers to understand a scientific text properly, they would need to know the concepts and presuppositions involved. Ignorance of these means that the text as a whole is not understood.

It seems that ESP teachers are not capable of coping with the specialism single-handedly, and what they may need to do is to collaborate in some way with subject specialists.

There are various kinds of collaboration, the most typical of which is team teaching, which usually involves two teachers (the ESP teacher and the subject teacher) being present in the classroom together. A pioneering study is by Johns and Dudley-Evans (1985), who consider that knowledge of both language and subject specialism is necessary for academic success:

> An overseas student's failure to keep pace with his course or with his research is rarely attributable to 'knowledge of the subject' or 'knowledge of the language' alone: most often these factors are inextricably intertwined. If their work is separate, it is difficult for the subject teacher, and even more so for the language teacher, to take account of that intertwining. In the triangle of which the three angles are the student, the subject teacher and the language teacher, each needs a certain type of assistance and feedback from the other two (1988).

Other similar studies have been undertaken by Chamberlain and Baumgardner (1988), Jackson and Price (1981) and Guezguez (1986). All these examples of team teaching involved two teachers, with the ESP teacher acting 'as a constant check to see that any questions or problems of structure, vocabulary, or pronunciation were given immediate attention' (Chamberlain and Baumgardner, 1988: 107); while the subject teacher was present to cope with issues of content.

This kind of collaboration, however, has constraints from the practical point of view. The most obvious one is timetabling arrangements. In many situations the three parties may not be available at the same time. Sometimes this type of collaboration is not possible because of the ESP teacher's heavy teaching load. Some subject teachers may also fear being observed by a language specialist. Adams Smith (1980: 76) points out that 'there are massive difficulties in the way of establishing real team teaching between disciplines as dissimilar as English and biochemistry or mechanical engineering, among them being near-total mutual incomprehension of purpose, subject matter and pedagogical approach, together with conflicting schedules, different commitments to research and problems of basic attitudes'.

Another type of collaboration is called 'subject-language integration' by Dudley-Evans (1983) to refer to a situation where there is normally only one teacher present in the classroom, who is involved with both language and content. The materials being taught are, however derived from some earlier collaboration between the ESP teacher and the subject teacher. The assignments are jointly planned and then marked by both the ESP teacher and the subject teacher.

All the above examples of collaborative work between ESP teachers and subject teachers are concerned with broad language skills such as listening and writing (for example, Dudley-Evans, 1983), or lecture comprehension and writing examination answers (for example, Johns and Dudley-Evans, 1985). They demand the full commitment of the ESP teacher and the subject teacher. However, practical constraints mean that such close collaboration is not possible all the time.

8.4.2 Proposals for providing English Language support for postgraduate Engineering students

The Southeast Asian postgraduate Engineering students that participated in my studies had not received much ESP instruction at Warwick University. Some of these students had attended a one-week pre-sessional course at Warwick, which is conducted in September each year. Shortly after the students arrive at the university, they are placed on this course. The timing of the course has several bad effects. In the first place, the course is too short, and this poses difficulties for both the ESP teacher and the students. Preparing ESP materials that are useful and creative is a difficult task. Materials preparation for such a short course is more difficult. The time constraint makes it more important for the materials to have a clear and coherent unit structure. If it is tightly structured, it will produce a monotonous patterns of lessons, otherwise, not much input can be provided, and as a result not much learning takes place. The effect of this is that both teachers and students will not have a sense of progress and achievement.

This course is conducted shortly after the students arrive at Warwick, a time when they are still very new to the university and need to absorb and get used to a new culture and a new environment. Some students have not left their family and friends before, and therefore need time to adjust mentally and emotionally. Other students, for example, need time to find the right accommodation. All these factors make it difficult to concentrate. Ideally a pre-sessional ESP programme should last longer than this, to allow time for students to settle down and become motivated and focused. A longer pre-sessional programme is available for Warwick students from all disciplines. In three phases, it can be attended for one, two or three of the summer months. Participants must pay, however, to attend the programme.

In the teaching of the one-week course, collaboration between subject lecturers in the Engineering Department and ESP teachers is desirable. Rather than focusing on formal classroom teaching as Johns and Dudley-Evans (1985) did, I suggest contact between the ESP teachers and the subject lecturers in the joint preparation of a technical vocabulary list. In other words, ESP teachers and subject lecturers can work together to create a lexical syllabus. Syllabus here refers to a plan of work to be taught on a particular course.

This syllabus can be produced according to frequency and range which have proved to be good criteria for identifying technical vocabulary. A list of technical vocabulary identified using frequency and range criteria proved to contain many words that subject lecturers judged to be highly topic specific (see Chapter Six). By using these criteria, a more comprehensive pedagogical list can be developed with a larger corpus including all the twenty six textbooks used for the teaching of the modules. Subject lecturers could be consulted regarding the representativeness of this list.

Such a list would be useful to both subject lecturers and ESP teachers. The subject lecturer can refer to this list to make sure the technical terminology is covered in lectures during the course of the academic year; and to check students' understanding of these technical words. The ESP teacher can provide multiple exposures to these words.

8.5 The implications of lexical familiarization

The findings of the study described in Chapter Eight suggest that most technical terminology is not familiarized in the early parts of the Engineering textbooks included in my corpus. This conflicts with Bramki and Williams' belief (1984) that technical terms are lexically familiarized in the early parts of textbooks.

Educational textbook writers write for students at different stages of specialization. These differences may be reflected in the way technical terms are familiarized. Lexical familiarization is predominantly found in textbooks addressed to comparative new-comers to the subject field. Huckin (1981), for example, says, 'Those who write to non-specialists must be sure to explicate the most important concepts in their texts, by using examples, definitions, analogies or other forms of illustrations.' Bramki and Williams' (1984) claim that lexical familiarization occurs in the early parts of textbooks is based on the study of a textbook called **Introductory Economics**, which is clearly targetted at a not-yet specialist readership. In contrast, writers writing for more specialized readers may not make the effort to lexically familiarize newly introduced technical terms, because they believe that their readers are already familiar with most technical terms in the texts.

The Engineering textbooks included in my corpus are for postgraduate level engineers, who should have completed an undergraduate degree course, and therefore may be expected to have mastered most basic technical terminology in the relevant disciplines. However, postgraduate programmes often draw students from quite a wide range of educational background, and their knowledge of technical vocabulary is not therefore on the same level. Some students are still new-comers to some modules, although not to all modules. It seems in this case, that there is a gap between writers' perception of their readers' technical vocabulary knowledge and the readers' actual knowledge.

My finding that technical terms are not lexically familiarized in the postgraduate Engineering textbooks means that ESP teachers of students needing to read these books should not spend much time in the classroom helping their students to acquire strategies for capitalizing on lexical familiarizations. Rather, it might be helpful if subject lecturers spent more time in lectures explaining the meanings of the technical terms. Tutorials or workshops may be helpful to find out which technical words the students are familiar with, and which words are problematic and need to be taught either in class or on an individual basis. Another alternative might be to introduce some good accessible technical dictionaries.

Another aspect of the findings shows that among the limited instances of lexically familiarized technical items, the most frequent lexical familiarization is achieved by definition and direct explanation (see Chapter Seven). While it is not necessary to spend too much time on teaching strategies to cope with lexical familiarization, it is necessary to teach those related to definition and direct explanation. For example, the students should be taught that when a definition occurs, three elements occur together: the term, the class and the characteristic. They should also be taught the lexical signalling of the familiarization, in other words, the words used to signal definition and direct explanation; and the change of definition components, in cases where the term does not always come first.

It is useful for ESP teachers to raise students' awareness of different typefaces. The study described in Chapter Seven and other studies, for example, by Handee (1996) show that signalling of lexical familiarization is accompanied by typographic clues such as italics and bold typefaces. These typographic differences can be relatively easily recognized by students whose first language is written in roman script; this difference, however, might be more difficult to recognize for students whose first language is, for example, Chinese. It may be useful to train such students to be more sensitive to typographic changes.

8.6 Summary

This chapter summarises the findings of this research and the pedagogical implications for the teaching of vocabulary to the postgraduate Engineering students. Frequency and range have proved to be good criteria in this research, in identifying technical and subtechnical vocabulary in the Engineering textbooks. In the light of the findings of the two studies (Chapter Six) that these students had poorer receptive knowledge of technical vocabulary than of subtechnical vocabulary, the frequency and range criteria can be used to create a lexical syllabus for the joint use of the ESP teacher and the subject lecturer. Other pedagogical lists such as the low range and wide range example words, and words with a General English meaning and a technical sense can also be produced. Knowledge of these words should make reading of the Engineering textbooks more efficient.

Chapter Nine

Conclusion

9.1 Overview of the present research

This research addresses four research questions: what are the characteristics of subtechnical and technical vocabulary in my corpus according to frequency and range; how valid are the frequency and range criteria in the study of subtechnical and technical vocabulary; whether it is subtechnical or technical vocabulary that ESP students have better receptive knowledge of; and finally what role lexical familiarization plays in students' recognition of technical vocabulary.

To answer the first question, the definition of subtechnical and technical vocabulary is first examined (Chapter One). Previous studies have offered a wide range of conclusions concerning subtechnical vocabulary. Some studies conclude that subtechnical vocabulary consists of difficult items for learners (Higgins, 1985), while others conclude that subtechnical vocabulary consists of General English words used in a technical sense (Trimble, 1985); some studies focus on the discourse organizing or metadiscourse functions (McCarthy, 1990, 1991; Hyland, 1998), while other studies focus on the frequency and range patterns of this type of vocabulary (Inman, 1978; Farrell, 1990). The present research closely examines these definitions and arrives at an overall description of subtechnical vocabulary. To approach this task, Chapter Five analyses the distribution of words in different texts. The analysis shows that subtechnical and technical vocabulary play different roles in the texts apart from differences of frequency and range.

To answer the second question, two Engineering corpora were developed (Chapter Four) and analysed (Chapter Five). The analysis shows that although identified on the basis of frequency and range, the subtechnical and technical vocabulary in my corpora have the features of those words in lists drawn up by writers who used other criteria. Therefore frequency and range criteria are found to be good criteria in identifying these two types of vocabulary. These criteria can be used for various purposes in the teaching of vocabulary for the postgraduate engineers (see Chapter Eight).

Chapter Six answers my third research question by investigating my subjects' receptive knowledge of subtechnical and technical vocabulary. The results of the studies indicate that there is a significant difference in the subjects' receptive knowledge of technical and subtechnical vocabulary. This finding challenges the popular notion that subtechnical vocabulary presents difficulties for learners while technical vocabulary can be acquired through the study of subject disciplines. In this case, the students had not learned some of the technical vocabulary by the end of the course.

Although the present research began as an investigation of the vocabulary needs of the Southeast Asian postgraduate Engineering students, the final analysis presents a description of the frequency and range criteria as a tool for the investigation of subtechnical and technical vocabulary which can be applied to all kinds of academic text. The texts used in this study are for a particular group of learners, but other texts can also be used for other specified group of learners in other areas. In the concluding chapter of this thesis, I discuss the limitations of this research (9.2) and several specific applications of the frequency and range criteria, regarding subtechnical and technical vocabulary (9.3).

9.2 Limitations of the research

The study described in Chapter Two found that the Southeast Asian postgraduates knew approximately 4,200 English words. My decisions regarding the number of General English words to be eliminated before identifying subtechnical vocabulary were based on that study.

However, my decisions did not reflect the precise picture of these students' vocabulary size. As described in Chapter Two, Mainland Chinese postgraduates' mean score for the 5,000 word level was 85.79, which suggests these students knew about 4,290 words instead of 4,200 words. The Hong Kong participants' mean score for the 5,000 word level was 82.04, which indicates that the Hong Kong postgraduates knew only about 4,102 words instead of 4,200 English words. Therefore it would have been more precise if the most frequent 4,200 words were eliminated from my corpus. Removing 5,000 rather than 4,200 words from my corpus means that 800 words that my subjects might not have learned were excluded from my studies. These words may be important for the reading of the Engineering textbooks.

In relation to my first research question, the large corpus used for the analysis of features of subtechnical and technical vocabulary is not a collection of the complete twenty six books used for the teaching of the Engineering modules. The fact that only the early parts of the textbooks were included in the corpus excludes the possibility of more technical terminology for analysis. For example, some low frequency technical vocabulary such as *experimental design* or *polymerisation* might have had a higher frequency of occurrence within a larger corpus. The exclusion of these words as technical is related to my frequency cut-off point. My frequency cut-off point for technical vocabulary is five, which means technical words with a frequency of less than five were not counted. This criterion excluded the possibility of words such as those listed above being classed as technical.

The findings regarding my third research question suggest that the postgraduate Engineering students had imperfect receptive knowledge of technical vocabulary. In the light of these findings, we can conclude that technical words were not learned from the course in the manner suggested by Robinson (1991) and many other ESP practitioners.

These studies were undertaken at the end of the students' course studies. It would also have been interesting to test students' receptive knowledge of technical and subtechnical vocabulary at the beginning of the course, or shortly after they arrived at Warwick University.

It is possible that the students' knowledge of technical words, although remaining poor, increased during the year in line with Robinson's claim that technical words tend to be learned through course work. Alternatively the students' knowledge of subtechnical and technical vocabulary may have been as good at the beginning as at the end of the course. In this case neither subtechnical nor technical vocabulary would have been learned through course work. A third possibility would be that the students' receptive knowledge of subtechnical and technical vocabulary was better at the beginning than at the end of the course. From the practical point of view, however, this possibility is least likely. However, without a test of these students' vocabulary knowledge at the beginning of the course, the extent to which new vocabulary was acquired during the year is not known.

The vocabulary test described in Chapter Seven was a receptive one, and words were presented without any surrounding context. Thus the results did not reflect subjects' ability to produce or interpret the two types of words in context. It may be argued that a contextual element is particularly important in the case of subtechnical items, because many have little or no meaning outside the text, serving a discourse organizing 'procedural' function. It is likely, however, that a contextualised test would strongly confirm my findings, because context would aid recognition of procedural words more than it would aid the recognition of topic specific words with extra-linguistic meaning.

9.3 Suggestions for future research

Chapter Eight indicates future research based on a larger corpus including all the twenty six textbooks used for the teaching of all the Engineering modules. This corpus could be a valuable resource for many purposes.

In the first place, it could be used to develop the syllabus proposed in section 8.4. From this larger corpus, a more comprehensive list of technical vocabulary could be developed by using frequency and range criteria. Secondly, one of the findings of this research is that technical terminology is not lexically familiarized in the early parts of the textbooks. This could mean that textbook writers assume that they are writing for a highly specialist readership, but does not preclude the possibility that technical terms may be lexically familiarized in other parts of the textbooks, for example in the middle or at the end, rather than in the introductory chapters. Only with a larger corpus can we investigate whether or not lexical familiarization occurs in the middle or other parts of the Engineering textbooks. The findings of this research can be used to inform both the subject lecturers and the ESP teachers so that suitable teaching approaches can be employed. In the case that technical terms are familiarized at the middle parts, subject lecturers need to make sure that the students pay special attention to those parts, where technical terms are explicitly and intentionally explained and defined by the textbook writers and this should help the students understand the terms better. As far as ESP teaching is concerned the findings concerning lexical familiarization within a larger corpus can inform the ESP teachers about the categories of lexical familiarization that are used in the books. ESP teachers can find out whether there are more categories of lexical familiarization in the Engineering textbooks than those discovered by Bramki and Williams (1984) in their Economics textbooks, and whether definition and direct explanation are the two most commonly used techniques in familiarizing the technical terms (see Chapter Seven). With this information, the ESP teachers can better decide which categories deserve class time.

In Chapter Five I mention that example words may cause misunderstandings if the writer and reader do not share the same cultural knowledge. These misunderstandings can be prevented if these words are pre-taught in an ESP course. With a larger corpus a more comprehensive list of example words can be produced.

My findings regarding my third research question suggest that subtechnical vocabulary presented fewer difficulties to my subjects than technical vocabulary did. My subjects did not appear to have acquired essential technical vocabulary through course work, in the manner proposed by Kennedy and Bolitho (1984), Trimble (1985), Nation (1990), and Robinson (1991), even though they were at the end of their study period and were about to return to their countries of origin as qualified engineers.

However, my findings do not necessarily reflect the vocabulary knowledge of the entire overseas student population. They probably have greater relevance for postgraduates than for undergraduates, because most postgraduate students already have a certain amount of experience of reading academic texts in English, and will have encountered subtechnical words across a range of fields in their undergraduate studies. This will give them an advantage in any subtechnical vocabulary test which undergraduate students will not share. Moreover, many postgraduate courses, like the Warwick Msc. in Engineering, enroll students with backgrounds in a variety of subfields, and require participants to take modules in subfields that are new to them, as well as in subfields they have studied before. This means that at the start of the course all students will be familiar with the terminology of some modules but not of others. Whereas undergraduate students may be treated to a slow introduction to the terminology of their subject area, Masters degree students have only a few months to learn terminology which some of their fellow students already know.

My findings regarding my third research question suggest that the Southeast Asian postgraduate Engineering students have better knowledge of subtechnical vocabulary than technical vocabulary when they complete their course. Future experiments of this type will need to investigate students' vocabulary knowledge at both the beginning and the end of their course of study, in order to ascertain the type and quantity of vocabulary acquired while studying in Britain. It may be that undergraduates' knowledge of the two types of vocabulary increases as an even pace over the extended degree period, while students on twelve-month postgraduate courses arrive with a good background in subtechnical vocabulary and acquire fewer new words while they are in Britain. If further studies confirm these findings, investigations will be needed as to why otherwise successful students (in terms of passing the degree courses and satisfying the Engineering department's academic requirements) fail to learn terminology that is essential to a full understanding of their subject specialism. It is not known, for example, whether they lack the listening skills necessary to absorb oral definitions given during subject lectures. Answers to these questions may inform decisions on ESP syllabus design in the future.

9.4 Conclusion

The findings of this thesis suggest that a technical word list common to different Engineering disciplines would be useful in learning situations where students of different fields of Engineering have to follow the same English programme. The lists provided in this study are incomplete because they are based on a relatively small text population. They do, however, provide insight into the features of the two categories of academic lexis: technical and subtechnical vocabulary. Frequency and range criteria as a tool for the investigation of these two types of vocabulary should also continue to prove useful in other related ESP situations, and it is hoped that they will provide a basis for much future research.

Bibliography

Abbot, G. 1983. 'Training teachers of EST'. *English for Specific Purposes*. 2 (1): 33-36.

Adams Smith, D. 1980. 'Co-operative teaching: bridging the gap between E and SP' in: British Council, 1980. **Team Teaching in ESP**. ELT Document 106. London: British Council English Teaching Information Centre.

Adams Smith, D. 1983. 'ESP teacher-training needs in the Middle East'. *ESP Journal* 2 (1): 37-38.

Anderson, R. and Freebody, P. 1981. 'Vocabulary knowledge' in: John, T. (ed.) **Comprehension and Teaching: Research Reviews**. Delaware: International Reading Association.

Anderson, C. and Freebody, P. 1983. 'Reading comprehension and the assessment and acquisition of word knowledge' in: *Advances in Reading Language Research* 2: 231-256.

Baker, M. 1988. 'Sub-technical vocabulary and the ESP teacher: an analysis of some rhetorical items in medical journal articles'. *Reading in a Foreign Language* 4 (2).

Barber, C.L. 1962. 'Some measurable characteristics of modern scientific prose' in: Swales, J.(ed.) **Episodes in ESP**. Oxford: Pergamon.

Barnbrook, G. 1996 **Language and Computers: A Practical Introduction to the Computer Analysis of Language**. Edinburgh: Edinburgh University Press.

Bialystok, E. and Sharwood Smith, M. 1985. 'Interlanguage is not a state of mind: an evaluation of the construct for second-language acquisition'. **Applied Linguistics** 6: 101-117.

Biber, D. 1993. 'Using register-diversified corpora for general language studies'. *Association for Computational Linguistics*.

Biber, D., Conrad, S. and Reppen, R. 1994. 'Corpus-based approaches to issues in Applied Linguistics'. *Applied Linguistics* 15 (2): 169-189.

Bloor, M. 1996. 'The English of Computer Science: Linguistic Issues and Learner's Problems' in: Hewings, M. and Dudley-Evans, T. (eds.) **Evaluation and Course Design in EAP**. Prentice Hall Macmillan in Association with the British Council. 180-191.

Blum-Kulka, S. and Levenston, E.A. 1983. 'Universals of lexical simplification' in: Faerch, C. and Kasper, G. (eds.) **Strategies in Interlanguage Communication.** 119-139. London: Longman.

Bramki, D. and Williams, R. 1984. 'Lexical familiarisation in economics text, and its pedagogic implications in reading comprehension'. *Reading in a Foreign Language* 2 (1): 169-181.

British Council, 1980. **Team Teaching in ESP**. ELT Document 106. London: British Council English Teaching Information Centre.

British Council, 1981. **The ESP Teacher: Role, Development and Prospects.** ELT Document 112. London: British Council English Teaching Information Centre.

Brown, D. F. 1980. 'Eight Cs and a G.' *RELC Journal Supplement 3*. Singapore: Regional English Language Centre. 1-17.

Campion, M. E. and Elley, W.B. 1971. **An Academic Vocabulary List.** Wellington: NZCER.

Carroll, J. B., Davies, P. and Richman, B. 1971. **Word Frequency Book**. New York: American Heritage.

Carter, R. 1987. **Vocabulary: Applied Linguistic Perspectives**. Allen & Unwin Ltd.

Carter, R. and McCarthy, M. (eds.) 1988. **Vocabulary and Language Teaching.** London: Longman.

Carton, A. 1971. 'Inferencing: a process in using and learning language' in: Pimsleur, P. and Quinn, T. (eds.) **The Psychology of Second Language Learning**. Cambridge: Cambridge University Press. 45-48.

Chamberlain, D. and Baumgardner, R.J. (eds.) 1988. **ESP in the Classroom: Practice and Evaluation**. ELT Document 128. Modern English Publications in association with the British Council.

Chapman, L.J. (ed.) 1980. **The Reader and the Text**. London: Heinmann.

Clark, D. and Nation, P. 1980. 'Guessing the meanings of words from context: strategies and techniques'. *System* 8 (3): 211-220.

Coleman, H. (ed.) 1989. Working with Language: A Multiplainary Consideration of Language Use in Work Contexts. Contributions to the Sociology of Language 52. Berlin: Mouton de Gruyter.

Collins COBUILD English Language Dictionary. 1995. London: HarperCollins.

Cornu, A.M., Vanparijs, J., Delahaye, M., and Baten, L. (eds.) 1986. **Beads or Bracelet: How Shall we Approach LSP?** Selected papers from the 5th European Symposium on LSP, Leuven, Belgium, 26-30 August 1985. Leuven: Oxford University Press.

Cowan, J. 1974. 'Lexical and syntactic research for the design of EFL reading materials'. *TESOL Quarterly* 8 (4): 391.

Crismore, A., Markkannen, R., and Steffensen, M. 1993. 'Metadiscourse in persuasive writing: a study of texts written by American and Finnish unversity students'. *Written Communication* 10(1): 39-71.

Crombie, W. 1985. **Process and Relation in Discourse and Language Learning.** Oxford: Oxford University Press.

Cronbach, L.J. 1942. 'An analysis of techniques for diagnostic vocabulary testing'. *Journal of Educational Research* 36: 206-217.

Crothers, E. and Suppes, P. 1967. **Experiments in second-language learning.** New York: Academic Press.

Dale, E. 1965. 'Vocabulary measurement: techniques and major findings'. *Elementary English,* 42: 895-901.

Daries, A. Criper, C. and Howatt, A.R.P.(eds.) 1984. **Interlanguage.** Edinburgh: Edinburgh University Press.

De Hann, P. 1992. 'The optimum corpus sample size?' in: Leitner G. (ed.) **New Directions in English Language Corpora.** Berlin-New York: Mouton de Gruyter.

Diller, K. 1978. **The Language Teaching Controversy.** Newbury House.

Dodd, S. 1993. 'Lexis - an uncharted land' in: Pemberton, R. and Tsang, E. S. C. (eds.) **Studies in Lexis.** Language Centre, Hong Kong University of Science and Technology. 1993: 35-47.

Dudley-Evans, T. 1983. 'An experiment in the team teaching of English for occupational purposes' in: Dudley-Evans, T., 1983. **Papers on Team Teaching and Syllabus Design.** Occasional Papers 27, Singapore: SEAMEO Regional English Language Centre. 35-41

Dudley-Evans, T., 1983. **Papers on Team Teaching and Syllabus Design.** Occasional Papers 27, Singapore: SEAMEO Regional English Language Centre

Dudley-Evans, T., 1987. 'Introduction to Dudley-Evans' in: Dudley-Evans, T. (ed.), 1987. 'Genre Analysis and ESP'. *ELR Journal* 1. Birmingham: English Language Research, University of Birmingham.

Dudley-Evans, T. (ed.), 1987. 'Genre Analysis and ESP'. *ELR Journal* 1. Birmingham: English Language Research, University of Birmingham.

Dunmore, D. 1989. 'Using Contextual Clues to Infer Word Meaning: Evaluation of Current Exercise Types'. *Reading in a Foreign Language*, 6 (1): 337-347.

Elliman, J. 1980. 'EST TTICorvallis 1979'. *AL Manakh* 4 (1): 33-47.

Engels, L.K. 1968. 'The fallacy of word counts'. *IRAL* 111 (2).

Faerch, C. and Kasper, G. (eds.) 1983. **Strategies in Interlanguage Communication.** 119-139. London: Longman.

Faerch, C., Haastrup, K. and Phillipson, R. 1984. 'Learner language and language learning'. Kopenhamn: Gyldendals Sprogbibliotek and Cleveden: Multilingual Matters.

Farrell, P. 1990. 'Vocabulary in ESP: a lexical analysis of the English of Electronics and a study of semi-technical vocabulary'. *CLCS Occasional Paper* 25: 1-81.

Fillmore, C. 1992. Corpus linguistics or Computer-aided armchair linguistics. in Svartvik, J. (ed.) 1992a. **Directions in Corpus Linguistics.** Proceedings of the Nobel Symposium 82, Stockholm 1991. Berlin: Mouton de Gruyter. 35-60.

Flowerdew, J. 1992. 'Definitions in science lectures.' *Applied Linguistics* 13 (2): 202-221.

Flowerdew, L. and A.K.K. Tong, (eds.) **Entering Text.** 183-96. Language Centre. Hong Kong University of Science and Technology.

Francis, G. 1986. **Anaphoric Nouns.** University of Birmingham: English Language Research.

Francis, W.N. 1982. 'Problems of assembling and computerizing large corpora' in: Johansson, S. (ed.) 1982. **Computer Corpora in English Language Research.** Bergen: Norweigian Computing Centre for the Humanities. 7-24.

Francis, W.N. 1992. 'Language corpora B.C.' in: Svartvik, J. (ed.) 1992a. **Directions in Corpus Linguistics.** Proceedings of the Nobel Symposium 82, Stockholm 1991. Berlin: Mouton de Gruyter. 17-32.

Francis, W.N. and Kucera, H. 1989. **Manual of Information to Accompany a Standard Corpus of Present-day Edited American English, for Use with Digital Computers.** Department of Linguistics, Brown University.

Gairns, R. and Redman, S. 1986. **Working with Words: A Guide to Teaching and Learning Vocabulary.** Cambridge: Cambridge University Press.

Gershman, S.J. 1970. 'Foreign language vocabulary learning under seven conditions'. Ph.D thesis. Columbia University.

Ghadessy, M. 1979. 'Frequency counts, word lists, and materials preparation: a new approach'. *English Teaching Forum* 17 (1): 24-27.

Greenall, G. 1981. 'The EST teacher: a negative view' in: British Council, 1981. **The EST Teacher: Role, Development and Prospects**. ELT Document 112. London: British Council English Teaching Information Centre.

Guezguez, S. 1986. 'Team teaching: one way of motivating EFL students'. *English Teaching Forum* 24 (2): 42-43.

Gui, S. 1987. 'A survey of the size of vocabulary of Chinese students'. *Language Learning and Communication* 1 (2): 121-232.

Halliday, M.A.K. 1970. 'Functional diversity in language as seen from a consideration of modality and mode in English'. *Foundations of Language* 6: 322-361.

Halliday, M.A.K. and Martin, J.R. 1993. **Writing Science. Literary and Discourse Power**. London: The Palmer Press.

Handee, S. 1996. 'Lexical familiarization in Computer Science text: an preliminary investigation'. MA Dissertation, University of Warwick.

Higgins, J. 1985. 'Hard facts (Notes on teaching English to science students' in: Swales, J. (ed.) **Episodes in ESP**. Oxford: Pergamon.

Hoedt, J. and Turner, R. (eds.) 1981. **New Bearings in LSP**. Copenhagen: Copenhagen School of Economies.

Hoey, M.P. 1983. **On the Surface of Discourse**. London: Allen & Unwin Ltd..

Hoey, M.P. 1991. **Patterns of Lexis in Text**. Oxford University Press.

Hoey, M.P. (ed.) 1993. **Data, Description, Discourse: Papers on the English Language in Honour of John McH Sinclair on his sixtieth birthday**. London: Harper Collins.

Hoffman, L. 1981. 'The linguistic analysis and teaching of LSP in the German Democratic Republic' in: Hoedt, J. and Turner, R. (eds.) 1981. **New Bearings in LSP**. Copenhagen: Copenhagen School of Economies.

Hoffman, L. 1986. 'Material design in function of defined objectives' in: Cornu, A. M., Vanparijs, J., Delahaye, M., and Baten, L. (eds.) 1986. **Beads or Bracelet: How Shall we Approach LSP?** Selected papers from the 5th European Symposium on LSP, Leuven, Belgium, 26-30 August 1985. Leuven: Oxford University Press. 45-60.

Hofland, K and Johansson, S. 1982. **Word Frequencies in British and American English.** Harlow: Longman.

Holley, F.M. 1973. 'A study of vocabulary learning in context: the effect of new-word density in German reading materials'. *Foreign Language Annals* 6: 339-347.

Holley, F.M and King, J.K. 1971. 'Vocabulary glosses in foreign language reading materials'. *Language Learning* 21: 213-219.

Holmes, J. 1988. 'Doubt and Certainty in ESL textbooks'. *Applied Linguistics* 91: 20-44.

Huckin, T. 1981. 'Readability from a cognitive perspective' in: Anderson, P., C. Miller and J. Brookman (eds.) **New Essays in Technical and Scientific Communication: Theory, Research and Criticism.** Farmingdale, New York: Baywood Press.

Hutchinson, T. and Waters, A. 1987. **English for Specific Purposes: A Learner-centred Approach.** Cambridge: Cambridge University Press.

Hyland, K. 1994. 'Hedging in academic writing and EAP textbooks'. *English for Specific Purposes* 13 (3): 239-256.

Hyland, K. 1996a. 'Talking to the academy: forms of hedging in science research articles'. *Written Communication* 13(2): 251-281.

Hyland, K. 1996b 'Writing without conviction? Hedging in science research articles'. *Applied Linguistics* 17(4): 433-454.

Hyland, K. 1998. 'Persuasion and context: the pragmatics of academic metadiscourse'. *Journal of Pragmatics* 30: 437-455.

Ilson, R. (ed.) 1985. **Dictionaries, Lexicography and Language Learning.** Oxford: Pergamon.

Inman, M. 1978. 'Lexical analysis of scientific and technical prose' in: Todd-Trimble, M., Trimble, L. and Drobnic, K. (eds.) **English for Specific Purposes: Science and Technology.** Oregon: Oregon State University Press.

Jackson, M. and Price, J. 1981. 'A way forward: a fusion of two cultures' in: British Council, 1981. **The ESP Teacher: Role, Development and Prospects**. ELT Document 112. London: British Council English Teaching Information Centre.

Johansson, S. 1975. 'Some aspects of the vocabulary of the learned and scientific English'. *Gothenburg Studies in English 42*.

Johansson, S. (ed.) 1982. **Computer Corpora in English Language Research**. Bergen: Norweigian Computing Centre for the Humanities.

Johns, T. and Dudley-Evans, A. 1985. 'An experiment in team-teaching of overseas postgraduate students of Transportation and Plant Biology' in: Swales, J. (ed.) **Episodes in ESP**. Oxford: Pergamon.

Judd, E.L. 1978. 'Vocabulary teaching and TESOL: a need for re-evaluation of existing assumptions'. *TESOL Quarterly* 12 (1): 71-76.

Kennedy, C. and Bolitho, R. 1984. **English for Specific Purposes**. London: Macmillan.

Kennedy, G. 1992. 'Preferred ways of putting things with implications for language teaching' in: Svartvik, J. (ed.) 1992a. **Directions in Corpus Linguistics**. Proceedings of the Nobel Symposium 82, Stockholm 1991. Berlin: Mouton de Gruyter. 335-373.

King, P. 1989. 'The uncommon core: some discourse features of student writing'. *System* 17 (1): 13-20.

Kinsella, V. (ed.) 1980 **Language Teaching Surveys I**. Cambridge: Cambridge University Press.

Kirkpatrick, E.A. 1907. 'Vocabulary test'. *Popular Science Monthly*. 70: 157-164.

Kling, J. and Riggs, L. 1971. **Woodworth and Schlossberg's Experimental Psychology**. Third Edition. London: Methuen and Co.

Konecni, E. 1978. 'Scientific and technical rhetoric' in: Trimble, L. and Trimble, M. T. (eds.) **English for Specific Purposes: Science and Technology**. Oregon: Oregon State University Press.

Kucera, H. and Francis, W. 1967. **Computational Analysis of Present Day American English**. Providence, Rhode Island: Brown University Press.

Lado, R. Baldwin, B. and Lobo, F. 1967. 'Massive vocabulary expansion in a foreign language beyond the basic course: the effects of stimuli, timing and order of presentation'. Washing D.C.: U.S. Department of Health, Education, and Welfare, Project No. 5-1095.

Lakoff, G. 1972. 'Hedges: A study in meaning criteria and the logic of fuzzy concepts'. *Chicago Linguistic Society Papers* 8:183-228.

Lambrou, A.V. 1979. 'The form and function of definitions in undergraduate texts'. Unpublished MA thesis. University of Khartoum.

Lauren, C. and Norman, M. (eds.) 1989. *Special Languages: From Human Thinking to Thinking Machines,* selected papers from the 6th European Symposium on LSP, Vaasa, Finland, 3-7 August 1987, Clevedon: Multilingual Matters.

Li, Lan, 1989. 'Extending vocabulary in English for Specific Purposes - with special reference to Chinese university students'. Unpublished M. Phil Dissertation, University of Exeter.

Loots, M. 1986. 'Compiling a list of academically relevant words' in: Cornu, A. M., Vanparijs, J., Delahaye, M., and Baten, L. (eds.) 1986. **Beads or Bracelet: How Shall we Approach LSP?** Selected papers from the 5th European Symposium on LSP, Leuven, Belgium, 26-30 August 1985. Leuven: Oxford University Press. 264-273.

Lorge, I. and Chall, J. 1963. 'Estimating the size of vocabularies of children and adults: an analysis of methodological issues'. *The Journal of Experimental Education* 32 (2).

Lynn, R. 1973. 'Preparing word-lists: A suggested method'. *RELC Journal.* Singapore: Regional English Language Centre.

Mackay, R. 1978. 'Identifying the nature of the learners' needs' in: Mackay, R. and Mountford, A. 1978. (eds.) **English for Specific Purposes.** London: Longman. 21-37.

Mackay, R. and Mountford, A. 1978. (eds.) **English for Specific Purposes.** London: Longman.

Mackay, R. and Palmer, J. D. 1981. (eds.): **Languages for Specific Purposes: Program Design and Evaluation**. Rowley, Massachusetts: Newbury House.

Marshall, S. and Gilmore, M. 1993. 'Lexical knowledge and reading comprehension in Papua New Guinea'. *English for Specific Purposes* 12: 69-81.

Martin, A.V. 1976. Teaching academic vocabulary to foreign graduate students. *TESOL Quarterly* 10 (1): 91-99.

Martin, J.R. 1993. The discourse of geography: ordering and explaining the experiential world. In Halliday, M. A. K. and Martin, J. R. (eds.): **Writing Science. Literary and Discourse Power**. London: The Palmer Press.

McCarthy, M. 1990. **Vocabulary**. Oxford: Oxford University Press.

McCarthy, M. 1991. **Discourse Analysis for Language Teachers**. Cambridge: Cambridge Language Teaching Library.

McEnery, T. and Wilson, A. (eds.) 1996. **Corpus Linguistics**. Edinburgh: Edinburgh University Press.

Meara, P. 1984 'The study of lexis in interlanguage'. Daries, A. Criper, C. and Howatt, A.R.P.(eds.): **Interlanguage**. Edinburgh: Edinburgh University Press.

Meara, P. and Jones, G. 1987. 'Tests of vocabulary size in English as a Foreign Language'. *Polyglot* 8 (1).

Meara, P. and Buxton, B. 1987. An alternative to multiple choice tests. *Language Testing* 4 (2): 142-151.

Melka Teichroew, F. 1982. 'Receptive vs. productive vocabulary: A survey'. *Interlanguage Studies Bulletin Utrecht* 6 (2): 5-33.

Miller, C.R. and Selzer, J. 1985. 'Special topics of argument in engineering reports' in: Odell, L. and Goswami, D. (eds.) **Writing in Nonacademic Settings**. New York: Guilford Press: 309-341.

Moody, K. 1975. 'Levels of EST lexis'. *RELC Journal* 21 (1). Singapore: Regional English Language Centre, Singapore.

Morgan, C.L. and Bailey, W.L. 1943. 'The effect of context on learning vocabulary'. *Journal of Educational Psychology* 34: 561-565.

Morgan, C.L. and Foltz, M.C. 1944. 'Effect of context on learning a French vocabulary'. *Journal of Educational Research* 38: 213-216.

Murison-Bowie, S. 1993. **Microconcord Manual**. Oxford: Oxford University Press.

Nation, P. 1982. 'Beginning to learn foreign vocabulary: a review of the research.' *RELC Journal* 13 (1): 15-36. Singapore: Regional English Language Centre.

Nation, P. 1983. 'Testing and teaching vocabulary'. *Guidelines* 5 (1): 12-25.

Nation, P. 1990. **Teaching and Learning Vocabulary**. New York: Newbury House.

Nation, P. and J. Coady 1988. 'Vocabulary and Reading' in: Carter, R. and M. McCarthy (eds.) **Vocabulary and Language Teaching**. London: Longman.

Nunan, D. 1992. **Research Methods in Language Teaching.** Cambridge: Cambridge University Press.

Nuttall, C. 1982. **Teaching Reading Skills in a Foreign Language.** London: Heinemann.

Palmberg, R. 1990. Improving foreign-language learners' vocabulary skills. *RELC Journal* 21 (1). Singapore: Regional English Language Centre.

Praninskas, J. 1972. **American University Word List.** Longman: London.

Pueyo, I.G. and Val, S. 1996. 'The construction of technicality in the field of Plastics: a functional approach towards teaching technical terminology'. *English for Specific Purposes* 15 (4): 251-278.

Richards, J.C. 1974. 'Word lists: problems and prospects'. *RELC Journal* 5 (2): 69-84.

Robinson, P.J. 1988. 'A Hallidayan framework for vocabulary teaching: an approach to organizing the lexical content of an EFL syllabus'. **International Review of Applied Linguistics** 26 (3): 229-238.

Robinson, P. 1980. **English for Specific Purposes.** Oxford: Pergamon.

Robinson, P. 1989. 'An overview of English for Specific Purposes' in: Coleman, H. (ed.) Working with Language: A Multiplainary Consideration of Language Use in Work Contexts. Contributions to the Sociology of Language 52. Berlin: Mouton de Gruyter.

Robinson, P. 1991. **ESP Today: A practitioner's guide.** Hemel Hempstead: Prentice Hall.

Sager, J.C., Dungworth, D. and McDonald, P. F. 1980. **English Special Languages.** Wiesbaden: Oscar Brandstetter Verlag KG.

Salager, F. 1983. 'The lexis of fundamental medical English: classificatory framework and rhetorical function (a statistical approach)'. *Reading in a Foreign Language* 1: 54-64.

Salager-Meyer, F., Defives, G., Jensen, C. and de Filipis, M. 1989. 'Communicative function and grammatical variations in medical English scholarly papers: A genre analysis study' in: Lauren and Nordman: 151-160.

Salager-Meyer, F., Defives, G., Jensen, C. and de Filipis, M. 1989. 'Principal component analysis and medical English discourse: an investigation into genre analysis' in: *System* 17 (1): 21-34.

Saragi, T., Nation, P. and Meister, G.F. 1978. 'Vocabulary learning and reading'. *System* 6: 72-78.

Schiffrin, D. 1980. 'Metatalk: organisational and evaluative brackets in discourse'. *Sociological Inquiry* 50: 199-236.

Scott, M. and Johns, T. 1993. **Microconcord**. Oxford: Oxford University Press.

Seashore, R.H. and Eckerson, L.D. 1946. 'The measurement of individual differences in General English vocabularies'. *The Journal of Educational Psychology.*

Seibert, L.C. 1930. 'An experiment on the relative efficiency of studying French vocabulary in associated pairs versus studying French vocabulary in context'. *Journal of Educational Psychology* 21: 297-314.

Seligher, H.W. and Shohamy, E. 1989. **Second Language Research Methods**. Oxford: Oxford University Press.

Selinker, L. 1979. On the use of informants in discourse analysis and "Language for specialized purposes"'. *IRAL* 17 (3): 189-215.

Sheerin, S. 1981. 'Some difficulties, real and imagined, in conducting medical case conferences in the teaching of doctor/doctor language' in: Lexden Centre (Oxford) Ltd. 1981. **Lexden Paper 2**, essays on teaching English for specific purposes by the staff of the Colchester and Bedford English Study Centres. Colchester: Lexden Centre (Oxford) Ltd.. 34-44.

Sinclair, J. 1985. 'Lexicographic evidence' in: Ilson, R. (ed.) 1985. **Dictionaries, Lexicography and Language Learning**. Oxford: Pergamon. 81-94.

Sinclair, J. 1986. 'Basic computer processing of long texts'. Computers in ELT and Research, Selected papers from the 1984 Lancaster Symposium *Computers in English Language Education and Research.*

Sinclair, J. 1987. **Looking Up**. London: Collins.

Sinclair, J. 1991. **Corpus, Concordance, Collocation**. Oxford: Oxford University Press.

Strevens, P. 1988. 'The learner and teacher of ESP' in: Chamberlain, D. and Baumgardner, R. J. (eds.) **ESP in the Classroom: Practice and Evaluation**, ELT Document 128. Modern English Publications in association with the British Council. 39-44.

Svartvik, J. (ed.) 1992a. **Directions in Corpus Linguistics**. Proceedings of the Nobel Symposium 82, Stockholm 1991. Berlin: Mouton de Gruyter.

Svartvik, J. 1992b. 'Corpus linguistics come of age' in: Svartvik, J. (ed.) 1992a. **Directions in Corpus Linguistics**. Proceedings of the Nobel Symposium 82, Stockholm 1991. Berlin: Mouton de Gruyter. 7-13.

Swales, J. 1971. **Writing Scientific English**. London: Nelson.

Swales, J. 1976. 'Verb frequencies in scientific English'. *ESPMENA Bulletin* 4: 28-31.

Swales, J. 1981. 'Aspects of Article Introductions'. Birmingham: Language Studies Unit, University of Aston.

Swales, J. 1985. **Episodes in ESP: A source and reference book on the development of English for science and technology**. Oxford: Pergamon.

Swales, J. 1986. 'A genre-based approach to language across the curriculum' in: Tickoo , M. L. (ed.) 1986. *Language Across the Curriculum*, selected papers from the RELC seminar on 'Language Across the Curriculum', Singapore, 22-26 April 1985, Anthology Series 15, Singapore: SEAMEO Regional English Language Centre.

Swales, J. 1988. 'Discourse communities, genres and English as an international language'. *World Englishes* 7 (2): 211-220.

Thorndike, E.K. and Lorge, I. 1944. **The Teacher's Word Book of 30,000 Words**. New York: Teachers College, Columbia University.

Tickoo , M.L. (ed.) 1986. *Language Across the Curriculum*, selected papers from the RELC seminar on 'Language Across the Curriculum', Singapore, 22-26 April 1985, Anthology Series 15, Singapore: SEAMEO Regional English Language Centre.

Trimble, L. 1985. **English for Science and Technology: A Discourse Approach**. Cambridge: Cambridge University Press.

Trimble, L. and Trimble, M.T. (eds.) **English for Specific Purposes: Science and Technology**. Oregon: Oregon State University Press.

West, M. 1953. **A General Service List of English Words**. London: Longman.

Widdowson, H.G. 1983. **Learning Purpose and Language Use**. Oxford: Oxford University Press.

Williams, R. 1980. 'Lexical familiarization in content area textbooks'. in: Chapman, L. J. (ed.) **The Reader and the Text**. London: Heinmann.

Williams, R. 1985. 'Teaching vocabulary recognition strategies in ESP reading'. *ESP Journal* 4 (2): 121-131.

Winter, E.O. 1977. 'A clause-relational approach to English texts: a study of some predictive lexical items in written discourse'. *Instructional Science* 6 (1): 1-92.

Winter, E.O. 1982. **Towards a Contextual Grammar of English**. London: Allen & Unwin Ltd..

Working Group, 1989. **Reports of the Working Group Set up to Review Language Improvement Measures.** Hong Kong: Education Department.

Xue, Guo-yi and Nation, P. 1984. 'A University Word List'. *Language Learning and Communication* 3 (2): 93-342.

Yang, H. 1986. 'A new technique for identifying scientific/technical terms and describing science texts'. *Literary and Linguistic Computing* 1 (2): 92-103.

Yorio, C.A. 1971 'Some sources of reading problems for foreign language learners'. *Language Learning* 21: 107-115.

Zettersten, A. 1979. **Experiments in English Vocabulary Testing.** Malmo: Liber Hermods.

The Engineering Corpora

This list gives details of the Engineering textbooks for the development of both corpora for this research.

Baker, H. W. 1966. **Modern Workshop Technology**. London: Macmillan. 1-81.

Bhattacharyya, G. K. and Johnson, R. A. 1977. **Statistical Concepts and Methods**. Singapore: John Wiley and Sons. 1-59.

Carrie, A. 1988. **Simulation of Manufacturing Systems**. Chichester: John Wiley and Sons. 1-68.

Coiffet, P. and Chirouze, M. 1983. **An Introduction to Robot Technology**. London: Kogan Page Ltd. 1-120.

Craine, J. F. and Graham, R. M. 1985. **Micro-computers in Engineering and Science**, Finland: Addison-Wesley Publishing Company. 1-77.

Crawford, R. J. 1987. **Plastics Engineering** (2nd edition). Oxford: Pergamon Press. 1-102.

DeGarmo, E. P., J. T. Black and Kohser, R. A. 1988. **Materials and Processes in Manufacturing** (7th edition). New York: Macmillan. 1-81.

Eaton, J., J. Smithers and Curran, S. 1988. **This is IT: A Manager's Guide to Information Technology**. Hemel Hempstead: Philip Allan. 1-74.

Groover, M.P. and Zimmers Jr., E. W. 1984. **Computer Aided Design and Manufacture.** Englewood Cliffs: Prentice Hall. 1-51.

Hunt, R. and Shelley, J. 1988. **Computers and Common Sense** (3rd edition) Hemel Hempstead: Prentice Hall. 1-38.

Jensen, K. and Wirth, N. 1985. **Pascal: User Manual and Report**. New York: Springer-Verlag. 1-101.

Johnson, H. T. and Kaplkan, R. S. 1987. **Relevance Lost: The Rise and Fall of Management Accounting**. Harvard Business School Press. 1-58.

Lumby, S. 1994. **Investment Appraisal and Financial Decisions**. London: Chapman and Hall. 1-53.

McCarthy, E. J. and Perreault, W. D. 1993. **Basic Marketing: A Global Managerial Approach**. Boston: Irwin. 1-73.

Monks, J. G. 1987. **Operations Management: Theory and Problems**. Singapore: McGraw Hill Book Company. 1-102.

O'Conner, P. D. T. 1991. **Practical Reliability Engineering** Chichester: John Wiley and Sons. 1-63.

Rich, E. and Knight, K. 1991. **Artificial Intelligence**. McGraw Hill Inc.. U.S.A. 1-62.

Reiss, G. 1992. **Project Management Demystified: Today's Tools and Techniques**. London: E & FN Spon. 1-67.

Schonberger, R. J. 1986. **World Class Manufacturing**. New York: The Free Press. 1-76.

Tanenbaum, A. S. 1996. **Computer Networks**. New Jersey: Prentice Hall. 1-115.

Thelning, K. E. 1984. **Steel and its Heat Treatment** (2nd edition). Chatham: Butterworths. 1-88.

Vollmann, T. E., Berry, W. L. and Whybank, D. C. 1992. **Manufacturing Planning and Control Systems** (3rd edition). Boston: Irwin. 1-55.

Ward, J. and Griffiths, P. 1990. **Strategic Planning for Information Systems**. Chichester: John Wiley and Sons. 1-71.

Wild, R. 1990. **Production and Operations Management** (5th edition). London: Cassell Educational Ltd. 1-78.

Wilson, D. C and Rosenfeld, R. H. 1990. **Managing Organizations**. London: McGraw Hill Book Company. 1-62.

Winfield, A. 1983. **The Complete FORTH**. Cheshire: Sigma Technical Press. 1-74.

Appendix One

Interview with eight Engineering students

This appendix includes eight interview transcripts with eight students: Dicken, Eric, Frank, Leo, Michael, Raymond, Samuel and Tony. In the transcript, their names are abbreviated as D, E, F, L, R, M, S and T respectively; 'I' stands for the interviewer.

Before the interview, each student was asked to look through the list of modules for the Engineering programme and identify the modules that he was studying.

Dicken -

I: How do you find the modules? Are they easy to cope?

D: I started from Manufacturing Strategy.

I: Have you done some reading on it? You need to do some reading, I suppose. How do you find it?

D: The first one ('Manufacturing Strategy') I have done the assignment, and the second one ('Human Factors in Industry') for the second one, I do a lot of reading, first of all, I search a lot of reference book with the relevant information, and then read them all and then summarize and then type into the computer.

I: You are talking about 'Human Factors in Industry'.

D: Yeah, 'Human Factors'.

I: Is it your core module work?

D: Yeah, yeah, it's a core. I only have two elective, this one is the core.

I: Oh, I see.

D: The elective is 'Simulation of Production Systems'. Another one is 'Financial'.

I: Oh, right, so 'Simulations' and 'Human Factors'.

D: No, sorry, 'Simulations' and another elective.

I: While you are reading for the module work, do you find them OK, generally, in reading?

D: For some topic, it is quite difficult to learn some new vocabulary.

I: Which one?

D: A lot of. Maybe in some new topic, maybe 'Human Factors' or something is not quite engineering, for example, Finance, something else, other than, I'll even say, 'Finance' or 'Human Factors' quite a lot of new words.

I: There are a lot of new words in these modules?

D: Yes, they are outside engineering.

I: Right. Do you find it a problem to read?

D: Yeah (laugh). That's a problem. I need to check up the dictionary for a long, for a lot of time.

I: Do you check up dictionaries when you come across new words?

D: Yeah, sure, I buy two good dictionaries, one small, one bigger one (Oxford Advanced Learner's Dictionary of Current English with Chinese Translation).

I: Which one do you use more often?

D: It's a translation from English to Cantonese. Ah, I've found something quite useful, because our neighbour have a dictionary translate the Mandarin into English, but I have to find a book Cantonese into English. Yeah, they use the pronunciation of Mandarin translate into English, it's quite useful.

I: Really?

D: Yeah, but I haven't find it (laugh).

I: OK, do you happen to have a vocabulary notebook?

D: Yeah, maybe you can see.

I: Can I have a look?

D: ...

I: I mean when you have problems with a certain word, you look it up in the dictionary and then you write it down.

D: Yeah, ... just a little, quite lazy since I come here.

I: So all these are you ... OK, can I take it with me and make a photocopy? I'll return it to you tomorrow, OK?

D: No problem, no problem. Yeah, maybe if you don't come here maybe I'll forget (laugh).

I: No, I'll give it back to you. Is that all (notebook)?

D: Yeah, you can see all. Quite lazy, yeah, you've discovered about this (laugh).

I: All right. Suppose there are so many vocabularies for you in a certain paragraph, of course, you look up in the dictionary, do you look up all the new words or some new words?

D: That depends. If I just want to learn English, I will see the newspaper to check up all the vocabulary. If I'm doing work, I would just take the ideas, OK, because of the time limit, I need to ...

I: OK, if you can understand the idea just let them go.

D: Yeah (laugh). No matter understand or not just let them go because of the time. Sometimes the context is quite important, just the idea within ten or twelve book, you don't need to understand every word, just the idea.

I: Oh right. So even if you let them go, you can still understand the idea.

D: Yeah.

I: OK. Do you guess meanings sometimes?

D: Sometimes (laugh).

I: Sometimes. What do you think of your guessing? Does it make sense most of the time, or you don't know?

D: uh ... in these days within my work I think usually I understand, but actually the problem is I cannot write is when I'm doing the assignment, I cannot write in my own language, it's in English, represent my idea, it's quite difficult. Reading is OK. Is there any specific questions you want to ask me?

I: That's my questions (laugh).

D: Oh, sorry.

I: Or if you want to tell me anything about vocabulary and reading, go ahead.

D: You mean, about ... when .... reading habit, is it?

I: Yes, reading habit and vocabulary difficulties.

D: For the vocabulary, there got plenty of words I cannot remember, because it's all within the dictionary. I cannot remember everything within a year. And then sometimes I think although I know the word, but actually my pronunciation is not good, so the local students cannot understand.

I: I can understand every single word you are saying.

D: Because you are Chinese, you know the pattern. Actually for local speakers they only ..., they don't understand.

I: Actually you really speak good English.

D: Maybe. Maybe I come here, I do a lot of reading on newspapers. Actually my problem is the oral English and reading English. That's the problem.

I: Do you have problems in understanding your lecturers?

D: Yeah, quite a lot because some lecturers they cannot pronounce every word very clearly, speak slower, and use some popular vocabulary, some of them, just use their own style, they speak very loud and very fast, I never follow them, it's very difficult, so it's not good.

I: OK, that's it. Thank you for your time.

D: Thank you. Is there anything else?

I: Right now, no. I'll let you know OK?

Eric -

I: Can you tell me something about these modules first of all?

E: What kinds of modules, you mean?

I: Do you find them easy to cope or not, these programmes, these modules, these textbooks?

I: The textbook, I think generally it's OK but maybe I also have some vocabulary when I reading these references, especially for the business reference, the module about the business, I should find the reference in other business module, yeah.

I: So you are saying you find Business or Management module is more difficult.

E: Yeah, compared with this scientific module, engineering reference.

I: Is it because you can't understand the vocabulary, the word mainly?

E: Yeah.

I: So what do you do when you come across new words, do you do anything?

E: Maybe I'll check my electronic dictionary, or I'll check the Longman Dictionary, and then I write down or save it in my electronic dictionary.

I: Is that all you normally do? Do you ever guess meanings?

E: Yeah, I do.

I: How do you guess?

E: How do I guess? Similarly when you refer a paragraph you'll find that there is some vocabulary you'll guess from the some of the sentence or behind the sentence you'll guess the meaning of the words, and also maybe some of the words you can guess it's not very critical or it's only the adjective, not very meaningful.

I: You are not guessing the key word.

E: Yeah, yeah, yeah.

I: The important words you don't guess.

E: I don't guess.

I: Do you look up the dictionary?

E: Yeah, simply I can't guess (laugh).

I: You can't guess? Why is that?

E: Sometimes it's can't guess, sometimes I think it's critical then I'll make sure I'll guess it right or wrong. If I guess wrong, then maybe I'll not very understand maybe misunderstanding, not very well.

I: Right, I see.

E: Do you only consider reading and vocabulary, not conversation, daily life, other things?

I: Do you have anything to tell me? Go on.

E: I don't know is this useful it's necessary to include this in or not?

I: Yes.

E: Yeah. But I think this one that let me some disappointed. You know maybe you also can discover inside this corridor this flat, it's mainly Asian, most of them is Asian. I think it's very friendly also, but I think if I can live together with the foreigner, I am not mean, first I should mention I am not discrimination for Asia, but you know when you come here only for one year and uh ... except you get some knowledge in your courses, I think to know more the countries, cultures, to improve your English is also quite important things. But you know if you live all together with the Asian most of the time. The other Asians mainly use, they can speak English, but on their daily life especially when they saw you are a Chinese, they mainly will use Chinese or Mandarin to communicate with other Asians. So you have less chance for you to practise English. You know before I came here, I also write a letter to the Accommodation Office. Before you want to apply for Accommodation, also several line to give you to, if you have any request for the Accommodation.

I: So you wrote to them.

E: Yeah, I not also wrote it down, I also sent it a letter to the Accommodation Office, but you I think this is no use.

I: No, it seems not working.

E: Yeah, I think so.

I: What did you say in your letter? You want to live with overseas, no, English students?

E: First of all, we mention since before we came here we discussed with several other students who also come from Hong Kong. They, we make an agreement, the several students, we want to live together, except that we also want to live with other foreigners. And there are so many reasons, for example, we can improve English, know more culture, or something, so on and so forth. But our mention is several students not many, only three or four, and then we hope we can meet the other people or students who come from different countries but not all Asians (laugh). But I think right now to make an opinion is too late, maybe I think we can only stay here for three or four months. I hope the Accommodation Office can consider the overseas student's situation.

I: Yes, they should.

E: Yeah, I think to live, to mix with the other foreigners very useful, it's good. See me now I feel quite pity of this reason. If this can reflect to the International Office or Accommodation Office, I think this is quite serious problems. I think I already sent two letter, yeah.

I: How many students are living in this flat?

E: Twelve.

I: How many Hong Kong students?

E: Hong Kong, let me see, uh ... I think it's three.

I: Who are the other students?

E: All is Malaysian Chinese.

I: Do you speak Cantonese all the time?

E: ... They'll mainly use English, but only about, I mean two to three they also use half of their time using English, or Mandarin, Cantonese, except this three we use Cantonese. So you know also in this record you can see my English is not very good maybe due to this reason. Maybe is also my excuse (laugh).

I: Your English is all right.

E: No, not OK. I think for as far as I know, communication is OK. You know, maybe, since, this probably maybe. I found in the first term we go to Flat four to having our oral. In Flat four, there is only two Chinese people, two Hong Kong people, no more, always foreigners. And firstly I come here I think my English improved in the several months, and after that I think it's stable, no ...

I: Remain the same level, right?

E: Yeah, remain the same level. You know, after you talk about the general things, for example which course you are studying, the simple, general question, you will have more things to want to tell to the foreigners. But sometimes you don't know your vocabulary, it's not so good to express your ideas, so sometimes you want to talk to them, you don't know how (laugh).

I: Do you find you're having the same problems in listening to the lectures? Your vocabulary is not enough for understanding your lecturers?

E: Lecturers? I think it's OK, maybe you know in your lecture I think about half or more than half students maybe they saw them it's Asian, including Malaysian, Taiwan, Hong Kong. Maybe due to this situation, they use more simple sentences, except ... But if you ask me how many percentage I can understand, I can only answer you about sixty to seventy but not ...

I: Not everything?

E: No, maybe sometimes they'll speak to the foreigner, maybe not very formal, the English.

I: They use some colloquial words, slang words.

E: Yeah, yeah, so maybe sometimes quite daily vocabulary, you know what I mean daily? not very formal. You'll hear some people laughing, but you don't know why they laughing.

I: Yeah, I have the same problem sometimes.

E: Yes, I think it's due to my understanding I am not really understand what they are talking about.

I: So It's a certain part of vocabulary that you don't understand, right?

E: Yeah, in the lecture. Yeah, but not everything.

I: But in reading ...

E: Reading ... I think it's OK, but maybe it's I can use a dictionary.

I: You use a dictionary very often?

E: Yeah, very often, specifically electronic dictionary, it's quite convenient.

I: What kinds of words do you look up in the dictionary?

E: You mean ...?

I: Is it technical terms or what kind of words do you usually look up in the dictionary?

E: In the reference?

I: Yes, the reference.

E: What kinds of words?

I: Yes, let's say you don't understand a technical term or you don't understand a colloquial word, I mean which kind of vocabulary?

E: I think it's not technical words. For technical words, I think mostly I already saw it before. But for the new vocabulary, maybe I seldom contact before, only saw it several times, not inside in my mind. So I will find it in my dictionary, that kinds of words, not many technical terms.

I: No, OK.

E: OK, thank you (laugh).

Frank -

I: What do you think of these modules? Have you done some reading?

F: Yes, some are OK, but, uh .... May I speak according to my time schedule?

I: Yes, sure.

F: The first module I have taken here is 'Introduction to Manufacturing Systems'. In this module I have read a lot different because I am totally new to the Engineering field.

I: Are you? Really?

F: Yeah. Although I ... some technical subject in Hong Kong Polytechnic, but I'm not from the Engineering Department. So it's totally new for me, and in fact that's why they put me, they arrange me to take this module as compulsory.

I: Do you find it difficult?

F: Quite difficult, because for the module, I have to, while doing the post module work, I have to tell something about the Manufacturing, or the car industry. Although I have taken two weeks doing the modules, but in fact it's not enough. I have to read a lot of book, in this field, but obviously the university resources for the books is not so enough for them.

I: I'm sorry, which modules are you talking about mainly?

F: 'Introduction to Manufacturing Systems'.

I: Oh, I see.

F: It is compulsory for those who have not got an engineering background.

I: Right. What about modules like Business and Management and Finance. Do you find them equally difficult?

F: No, for those, for the other one, for example, EBM, for post module, I have done some mathematics in Hong Kong Polytechnic, in actually in A level, and in the certificate level. The first week, the first few days of the module is what I've learned before, and only the module of Thursday and Friday is new for me. And on the ITA, ITA is Information Technology Applications, it is ..., not really tough, I mean not only deep in the subject, only an introduction to the IT application. So it's OK, and I score, I think I score the first module, I have scored high marks there.

I: Which one?

F: ITA.

I: IT, really?

F: Yeah, it's quite easy to understand, quite interesting.

I: Really? Did you have some background knowledge about Information Technology or not?

F: uh ... a little background, because I've worked before in a company which used computer. But actually the subject itself is an introduction to the other ..., to this area. So it is very simple. And for the CAD module ...

I: Is it Computer-aided Design?

F: Yeah. It's very easy because for the practical side you just follow the instruction for the lectures, it's also OK, because it's very clear and simple.

I: What do you think of the vocabulary generally?

F: Vocabulary?

I: Yes.

F: No, they do not cover the difficult terms, for some of the lecturers, some of them may have to create their own notes, for example, I'm doing this one, instead of saying long term, in the person he is saying the long range, actually I've seen a number of books but there is one book that concerning the long range one. So I just have to assume that long range pen is equals to long term bank.

I: Long ...?

F: Long range pen. Most times the book not concerning with the medium range either, but using the long term and short term, so I have to think what is the medium range and what is the long range, is this equal to the long term one, something like that.

I: Oh I see, OK. Do you find the technical terms easy or difficult?

F: Technical terms? Actually they are quite .., the method for example, JIT means just-in-time, I haven't learned this term before, but when I come here it's very simple, they are using the just-in-time theory, ... I already have all the model in my mind.

I: Which term is it did you say?

F: JIT, just-in-time. ... It's very famous and useful in the book.

I: OK. What do you do when you come across a new word you do not know? Do you do anything at all or let it go?

F: That depends on the ... on how much I read the book. For example, if I have to ... if the module is going to talk about what the term is, actually you can sit there and listen and understand it easily, for example, if the lecturer didn't come across the term, but the post module work ask you to do something like that, you have to look around and see the books, there's so many knowledge we can get from the book.

I: Right, do you look up dictionaries sometimes?

F: Yeah, I do.

I: What dictionaries do you use?

F: I ... I used to have an electronic dictionary, but it is burned.

I: Really?

F: Yeah (laugh), so I'm now using this one.

I: Can I have a look? ... Burlington Universal Dictionary, I haven't heard of this one before.

F: This is very ..., it's not enough for me.

I: What kind of dictionary is it? Is it a technical dictionary?

F: No, no, no.

I: So it's a ....

F: I can't find many words inside. Actually I'm going to buy another one, this one, I'm using this one now, this one, I borrow it from my neighbour.

I: Really? Concise edition, that's an interesting dictionary actually, Burlington.

F: (laugh) it's very poor.

I: Where did you find this dictionary?

F: It's not mine, it is my senior student, some senior student gave it to me.

I: Printed in USSR.

F: Really (laugh).

I: Yeah (laugh).

F: So I'm not using this one, do you want it at all?

I: No, thank you. Is this a bilingual one?

F: Yeah, I borrowed it from my neighbour, who is in the kitchen now.

I: Oh I see. I know this one, the Advanced Learner's Dictionary.

F: It's much better. Tomorrow I'm going to the Coventry to ...

I: OK, do you guess meanings sometimes?

F: uh ... I do. For example, when I was doing 'Strategic Marketing' module, so I have to concentrate on the Fort and JM company, ... the Henry Ford, when he was in the 20th, some people asked him they want to have more models, more variation from the famous model T car, so he asked Henry Ford, better they can have that models. Henry Ford told them that they can have any model, any colour you wish as long as it is black. At first, I don't understand the meaning, I asked the other students, actually it is something related colour subtraction theory, but I'm not sure, so I asked the other. By guessing, it should be in this way.

Leo -

I: First of all, can you tell me something about the modules? I mean have you done lots of reading?

L: No, I ... due to all the modules have one this course, so we haven't time to see all the book, we only focus on the post module work question, only focus on the few books only.

I: Can you say something more about the post module work?

L: Post module work? After I read, attend the module, we have two question or three question to answer. This is called post module work.

I: I see. OK, while you are reading for the post module work, do you have any problems?

L: Sometimes.

I: Sometimes? What kind of problems are they?

L: Some people write that I can't understand.

I: You can't understand? You cant understand the writing, the textbook?

L: Yes, the textbook.

I: Is it because of ...? Why is that? Why you cant understand? Is it because you don't know the vocabulary or is it because of something else?

L: No, the ... the whole sentence.

I: The whole sentence.

L: I find them the dictionary, if they were, but sometimes also can't know the whole meaning.

I: I see. Do you mean even if you know every word in the sentence, you can't understand the sentence?

L: Yes (laugh).

I: Really?

L: Yes, I'm not sure. I don't know why, but the basic meaning I can get in. But some paragraph have one or two sentence I cant understand.

I: Right. What do you do when you can't understand the sentences or words?

L: Pass (laugh).

I: You pass? Does it have an effect on your whole understanding of the paragraph?

L: No, I don't think so.

I: You can still get the meaning.

L: Yes, I can get the meaning.

I: Without understanding the several words. I see. So generally you have difficulties in reading.

L: Yes, I also spend a lot of time on the dictionary to find the meaning of the Chinese and English.

I: So what dictionary do you use?

L: Electronic dictionary (laugh).

I: Is it bilingual, English and Chinese?

L: Yes.

I: Do you look up words very often?

L: Pardon?

I: Do you use your electronic dictionary very frequently?

L: Yes.

I: You do. Is it technical terms you are looking for or other words?

L: I think other words.

I: Other words.

L: Yes, not technical terms.

I: No. Do you have a vocabulary notebook by the way?

L: No.

I: You don't?

L: Yeah (laugh), so lazy.

I: So you find the meaning from the electronic dictionary, you just let it go, you don't write it down.

L: No.

I: Do you sometimes guess word meanings?

L: Yes, sometimes.

I: What kinds of words do you guess?

L: Not sure, if I can understand the whole sentence, I will pass the other sentence without check in the dictionary.

I: OK. As far as these modules are concerned, do you find they are equally easy, or they are equally difficult?

L: I think it's difficult in the Business subject.

I: Difficult in the Business subject.

L: Yes, due to I haven't studied before.

I: Oh, I see.

L: So sometimes I cannot get the meaning and first title('Production Planning and Control') here the term, so I'm not sure. And also the ..., if I haven't know the subject before (laugh), I haven't any basic ideas, so sometimes I cannot hear what he's talking about from the lecture. So I think this the listening problem.

I: Right. So you find difficult to do the Business subject.

L: Yes.

I: And you said it is because you don't have any background knowledge.

L: Yes.

I: Other students, did they study Business before?

L: No.

I: No? I see, so it's Business ...

L: We study engineer, uh ... mainly in the engineer before, so we haven't any basic idea in Business.

I: Ah, I see, so you have some basic background knowledge about Manufacturing or Engineering.

L: Yes.

I: But not Business.

L: No.

I: All right, I've got you. Thank you, that's it.

L: Thank you.

Michael -

I: Can you tell me something about these modules please?

M: Which one?

I: The ones you are doing. Have you done lots of reading?

M: It depends which modules and what's the question about. Yeah.

I: Which one have you read a lot so far? Are you concentrating on a specific module? Is there a particular one you like?

M: uh ... I think the module I read most is uh ...none of them (laugh). Maybe the first one.

I: Maybe the first one ... Materials ('Metallic Materials Selection')?

M: Yeah, 'Materials Selections'.

I: Do you have problems reading for it? How do you find it, easy, difficult, or enjoyable, interesting, boring?

M: It is the first module I attend. OK, I done it last year, last year, September, no, October, maybe November, OK?

I: A long time ago.

M: Yeah, a long time ago.

I: Well, for instance, you are reading something for a certain module, do you have any problems?

M: Reading?

I: Yes, reading your textbook for a certain module this term.

M: Sure.

I: You have problems.

M: I have problems.

I: Is it because you don't have enough vocabulary, or is it because of the subject itself? Why is it difficult?

M: Vocabulary ... be a problem.

I: Is NOT a problem?

M: IS a problem. If I didn't, if I don't know anything about the subject, it could be nil to me, sure it would be most difficult.

I: Yes. Is it technical terms that is a problem or other vocabulary?

M: Is other vocabulary.

I: It's not technical terms?

M: Actually the Management subject will be more difficult.

I: I see, all right. So it's Management and maybe Business?

M: Yes, Business, because most of the subject I had do more or less I study some, for a period of study, so technical subject will have less problem.

I: Right, yes, that's fine. For instance, you are reading a book on Management, and you have lots of vocabulary, what do you do when you come across new words. What do you do?

M: Check the dictionaries or guess the meanings.

I: Yes, first of all, what kind of dictionaries do you use, English or bilingual Chinese-English?

M: The first one I will use is electronic dictionary (laugh).

I: All of the Hong Kong students use electronic dictionary.

M: It's more convenient and fast. The next will be English to English.

I: Yes, You said you guess meanings sometimes. How do you find it? Are you OK with it?

M: No.

I: No? but still you do sometimes?

M: Yeah (laugh).

I: Do you guess all the words, or ... how do you guess?

M: It depends. If it is important, like, if I feel this chapter is very important, I'll try to understand it, and guess it and check it, like that.

I: All right, I see. Do you have a vocabulary notebook?

M: No (laugh).

I: No? When you look up a word in a dictionary, you don't write it down?

M: I'm sure I'll forget.

I: OK, so you don't have a vocabulary notebook.

M: No, even if I know it is I should have one, but I don't.

I: All right, thank you.

M: Sure. Thank you.

Raymond -

I: Are you doing well with these modules? Have you done some reading?

R: You mean my course, for the whole course, yeah, I've read a lot of books.

I: Do you have any difficulties in reading?

R: Not much difficulties.

I: Are you sure?

R: Yeah, sure.

I: So you are doing all the modules equally well?

R: Yes.

I: Do you find all the modules equally easy for you?

R: You mean in term of what?

I: In terms of everything, reading for instance.

R: For me, no, no difficulties, it's true. Yeah, I think my background is quite different from other students because I've done a similar course in Hong Kong.

I: Right. Looking at these modules again, what do you think of Management and Materials, I mean they are very different modules, aren't they?

R: Yeah, for me, No six ('Metallic Materials Selection') is a little bit different, because it's a very very technical side, yeah, it's very very technical and talking about a lot of Metallic theories, it's a little bit difficult, but not much.

I: Do you mean you have difficulties understanding the theories?

R: In terms of the theories behind, yeah, it's quite complicated and very confusing to understand.

I: So you are saying you don't have enough background knowledge about the subject, about the theories. It's not the language difficulties, I mean English.

R: No, in fact, the subject is quite difficult, the theories are also difficult.

I: There must be many terms, am I correct?

R: Yeah, many terms. You need to have a clear idea about each terms.

I: When you come across a term that you are not sure of, what do you do?

R: By reading, you mean?

I: Yes, what do you do to solve the problem? For example, do you ask help from your lecturer, or consult a dictionary in order to know the word meaning, or what?

R: Because it's not easy to reach, approach a lecturer, they are very busy, we all know they are very busy, so we only try ourselves to solve any problems.

I: How do you do that?

R: To read more and try to compare different books and then get more ideas. So that's the only way to do with these problems.

I: Do you use dictionaries at all?

R: Oh, dictionaries, sometimes if we don't know the meaning we must consult a dictionary, it's very common but the frequency is not very high, because these terms, you cannot find easy to find in the dictionary, it's technical term, you need to read the whole passage in the book.

I: Are you saying you have difficulties with terms basically?

R: Some technical terms in this module, yeah.

I: You are all right with the rest of the modules?

R: Yeah, I'm all right.

I: Do you keep a vocabulary notebook?

R: Vocabulary? No, I don't have because we study this course, basically I already attend a similar course, in my background in Hong Kong. So sometimes I only forget the meaning, sometimes not much new thing I can remember here, just elaborate more here.

I: Have other students attended a similar course in Hong Kong as you did?

R: Yeah, most students from Hong Kong they finished Manufacturing in Hong Kong. So it's quite similar in here.

I: So your background is similar more or less?

R: Yeah, the background is quite similar, studying Manufacturing and then doing Manufacturing, and again this is a Master course in Manufacturing. So we have our general idea. This module is a little bit difficult, so I need to work more harder in this module.

I: So generally you don't find modules difficult to cope, and you don't have any particular problems in reading?

R: Generally no.

I: OK, thank you very much.

R: OK (laugh).

Samuel -

I: Have you read lots of books for these modules?

S: Yes, I have read a lot of books about the modules. First of all we read a lot of books at the earlier ... For example, for the 'Production Planning and Manufacturing Strategies', but nowadays because we need to start our individual project, we spend, so I spend a little bit less time to come across the module work. So, yeah ...

I: Do you any problems in reading the textbooks?

S: Generally it's not too much problem. No, well, I can say, I can say no problems.

I: No problems at all?

S: No, no, I have problems, but not too big.

I: So you are saying you can understand everything, or most of what you are reading?

S: If I use what level you understand your literature survey, I should say I just understand 70%. So the rest of the 30% I don't understand.

I: So there is 30% of the textbook you don't understand very well.

S: Yeah, and one thing I don't, I miss your understanding, I miss your meaning. Did you say uh..., reading ...?

I: Any problems in your reading of the academic textbooks for the modules you are doing, any problems?

S: Do you mean English problems? The difficulties? All kinds of problems? OK, yeah, 70%, it's reasonable.

I: So what do you do when you come across the part you cant understand for the 30%?

S: Yes, the problem. uh ... one thing I can sum up to the several problems. One thing, because uh ... in my opinion, individually, only myself, my listening abilities in English is not very well, because I cannot catch all the information from the lecturer. So I spend more time to read the book, this is one thing. So maybe due to the time is not enough to read. So we just read one or two reference books. So I go on to start my project. That is one problem. Secondly is very common in Asian students, the English problem.

I: Do you mean problems with English? any particular areas? for instance, listening, speaking, writing, reading?

S: All of them. If you just say the English problems, everything is a little bit problems, reading, speaking, writing. I'm not sure, something, I'm not sure, just, maybe even if you don't understand, but you can, you know what the book is talking about, if you just know the theory.

I: Why is that? Can you be more specific?

S: Yeah, because I don't understand the program, for example, 'Financial Analysis', because Finance is not too difficult to understand if you understand the meaning, the program, you can catch the meaning, it's not too difficult, it's not too technical.

I: Are you saying that some modules are technical, some are not? Some are difficult, some are not?

S: I can sum up a little bit. Some subjects are more difficult than the others.

I: Can you tell me which modules are more difficult for you, please?

S: 'Production Planning and Control', 'Manufacturing Strategies', 'Financial Analysis and Control Systems'. Yes, the rest of the modules are not too difficult.

I: Are there some reason why these modules are more difficult?

S: First of all, because these modules are the first modules in the fist term, so maybe I'm not understanding the requirement of the subject. So I go on to study a lot of the books, spend a lot of time to do these subjects, but I cannot sum up all the information. So maybe I get the wrong directions, so nowadays, I ... sometimes can't understand these subjects (laugh).

I: Are you getting any better at all, do you think?

S: Yes, it's better because it's my subject ..., I need to go on to study more of these subjects.

I: Do you think you can catch up? What do you think you can do with these difficulties?

S: Not too difficulties, but it's a lot of materials, because these subjects are quite important.

I: You must have a lot to read.

S: Yeah, first I think because we have a lot of questions to answer for individual module work, you need to cover a lot of materials. And because uh ... no, if you can catch a lot of information from the lecture, you can, I think you can save the time to read the books, just understand and, after lecture read the books and catch more faster.

I: So what do you do when you can't understand well what you are reading? Do you do anything when you are stuck?

S: Yeah, so ... I have a dictionary. Can I tell you the dictionary?

I: Yes, sure.

S: It's the Oxford English Dictionary.

I: So what kinds of words do you look up in the dictionary? Do you look up all the words you don't know?

S: Yeah, when I don't understand at all, when I don't understand the whole sentence.

I: Are you happy when you've got the meaning from the dictionary? What do you do next?

S: Yes, I pick it down to my vocabulary notebook.

I: You have a vocabulary notebook? Can I have a look? Do you mind?

S: No. Where is it? (searching on the shelf)

I: (with the notebook) That's very good, excellent, Oh, what's this word?

S: "Imminent", "i-m-m-i-n-e-n-t", it's not very special for any kinds of words, nouns, or adjectives, I don't know, just anything I don't know.

I: Can I make it a copy of it?

S: No, of course not. You don't understand these words (laugh)?

I: I will pop it back to you pigeonhole when I finish with it, OK?

S: Yeah, that's fine, no problems.

I: Do you look up all the words you don't know or what happens when you use a dictionary?

S: I think uh ... the situation is that when I look up dictionaries, I am not reading the whole pages if there are a lot of new vocabularies. If there is only two or three words in the sentences, I can catch the meaning. Yeah, but sometimes some words is the key point, I need to look up the dictionary (laugh).

I: You said you don't look up all the words in the dictionary, what do you do with the rest of the words, do you ignore them, let them go or what?

S: I guessed at first in the context. Yes, if I can follow the whole meaning.

I: OK, thank you very much. That's it.

S: That's it? I want to do some more (laugh).

Tony -

I: About these modules, have you done some reading?

T: So far I've done eleven modules, for each module, we have different post module work, OK? We did quite a lot about book, for the post module work, normally about three to four books before I start doing the questions.

I: Can you tell me something about the post module work?

T: You see, here, for example, 'Automation and Robotics' is one module, OK? After the module, we have one post module work, which is like an assignment, and we'll get two or three questions about the subject, and we've asked for it in six weeks' time. And after the six weeks' time, we'll hand it in.

I: OK. When you are reading these books, do you have any difficulties?

T: uh ... yes, because my background is Management Science, and I'm doing IT, which is a bit on the Engineering, something I uh ... computer language, you now, the programming, so I find a little bit hard, problem. Others are OK. I need to read quite a lot of books before I can tackle the problems. It's a bit different to other people. I know some of the people, Like Frank, he did Textiles, and I know some of the people did Manufacturing, and they come to the course, the subject like ... for example, Automation and Robotics, they did it before in Hong Kong, so they don't find any problem. For me, I find it a little bit problem. But something like Financial Analysis, I did some Financial before, so when I come to this module, I find it is quite easy, I don't need to read a lot of books, I just do the work.

I: Yeah. So your background is more or less on the Management Science, and you don't find Management Science very difficult, right? But certainly you will find some other areas more difficult.

T: I do, yeah. Some of the technical like programming, computer programming as difficult.

I: Is it difficult because of the terms, or other vocabulary? Do you find the terms difficult or ...?

T: It's mainly the terms, I think, yeah, because if I don't understand the terms, all right, I cant do anything, unless I read a lot of books, understand it.

I: It's the concept.

T: Yeah, the concept, mainly concept, the definitions. Some of the definitions

I have to understand it from the very bottom before I can start do the work. That's the problem.

I: Right. So when you come across the words you don't know, what do you do?

T: uh ... normally I read quite a few books first, if I still get stuck on the question, I'll ask my friends, see how they feel about it, how they tackle the problem, then I just follow it.

I: Do you look up words in the dictionary?

T: uh ... sometimes I do, yes. Dictionary is ... because you see, for each module we have quite a lot of book list, references, it was given out by the lecturer, just follow the book list, get into the library, we just read.

I: Sometimes you cant understand, suppose your friend is not here, do you just pass, or do you refer to the dictionary, or what do you do? You cant always ask your friends for help.

T: The first module, when I first done, it's completely Engineering, this is the first module, right, at that time, I didn't know any friend at all, it's the first module and I got stuck, what I did, I just come to ask the module teacher, and he gave me the answers, yes, that's the way I did. Once I know a lot of friends around, I can.

I: Do you guess meanings sometimes?

T: No, I don't. I have to find out the definition.

I: All right, you are not sure about the guess.

T: No. I don't guess (laugh)

I: Why is that? You don't like it?

T: I don't like it, yeah. I want to know the answer.

I: Right, you want to be sure.

T: Be sure, yeah, certainly.

I: Do you have a vocabulary notebook?

T: Vocabulary notebook?

I: I mean ...

T: I've got a dictionary.

I: No, I don't mean dictionaries. I mean a list of words you are not sure of, you look up the dictionary, and then you write them down.

T: No, no, no.

I: OK, that's it. Thank you very much.

T: That's OK.

Appendix Two
Technical vocabulary Identified in the Large Corpus

This is a list of all technical vocabulary identified in the large corpus, arranged in rank order of frequency and range. The first figure following the word form is frequency, the second range, and the third is peak frequency.

marketing	552	11	466
bit	385	10	101
stress	375	14	147
objective	370	14	203
stack	341	9	336
FORTH	326	7	312
memory	320	17	92
robots	301	6	225
network	276	10	183
block	275	10	194
layer	274	7	258
loop	259	13	185
strain	252	9	150
reliability	224	10	195
accounting	221	10	150
storage	214	13	138
plastics	211	2	208
investment	206	11	147
constant	204	15	120
circuits	199	10	121
fracture	198	5	113
operating systems	198	8	113
channels	197	9	154
LISP	196	2	194
chip	192	11	142
sample	188	10	136
software	187	17	120
probability	186	9	120
materials requirement planning (MRP)	183	4	171
terminal	182	10	106
address	180	13	132
operations management	180	2	157
microcomputer	177	10	121
fracture	172	5	101
behaviour	171	14	76
ok	168	3	162
information technology	163	4	130
hardness	161	5	101
data processing (DP)	159	7	134
grain	150	5	140
metal	140	17	101
creep	135	4	115
carbon	131	5	103
integrated services digital network (ISDN)	128	1	128
electronic	127	15	101
just-in-time (JIT)	126	5	58
operations manager	124	3	113
interface	123	10	84
string	120	9	85
buffer	116	8	101
deformation	115	4	87
population	110	10	68
queue	110	6	81
user	109	14	90
dictionary	106	4	97
identifier	104	3	100
lead time	103	8	56
integer	102	7	77
supplier	102	13	28
argument	101	11	57
transport	101	9	74
array	99	7	68
central processing unit	98	6	57
project management	96	5	91
modulus	95	6	72
automation	94	9	69
metals	94	9	81
data bases	92	8	68
hardware	92	14	38
MPC system	92	1	92
PROLOG	92	2	87
fatigue	90	6	75
competitors	86	9	64
manufacturing system	82	8	59
managerial	79	13	50
statistics	77	9	55
addr	77	1	77
alloy	76	5	50
shear	76	4	61
data base	75	6	54
system structure	75	1	75
computing	74	12	55
information systems (IS)	74	6	65
computer-aided design	72	5	59

payback	71	2	63
programmer	71	7	46
degrees of freedom (DOF)	70	2	65
mathematical	70	10	50
planned order	70	1	70
machine tool	69	8	49
machining	68	9	48
register	68	9	46
gross requirement	67	1	67
ring	67	8	43
bandwidth	66	1	66
transformation	66	8	43
colon definition	65	1	65
elastic	65	7	27
customer service	64	4	56
management accounting	64	1	64
raw materials	64	9	41
austenite	61	1	61
compute	61	13	25
flag	61	4	58
notation	61	11	26
syntax	61	4	53
computer-aided manufacture (CAM)	60	6	42
protocol	60	1	60
toughness	60	3	28
workpiece	60	4	48
bus	59	8	26
purchasing	58	8	30
rotation	58	10	34
entity	57	7	49
disc	56	8	47
martensite	56	1	56
scheduled receipt	55	1	55
boolean	54	3	32
cutting force	53	1	53
personnel	53	15	30
project manager	53	1	53
transactions	53	10	38
crystal	52	10	40
median	52	2	43
packet	52	1	52
rake	52	1	52
keyboard	50	6	24
variance	50	6	24
vocabulary	50	5	44
automated	49	9	30
built-up edge	49	1	49
standard deviation (SD)	49	4	44

frame	48	6	40
numerical control (NC)	48	5	32
rings	48	8	37
shorten	48	4	44
syntax diagram	48	1	48
voltage	48	9	36
gate	47	3	40
hardening	47	2	34
promotion	47	5	25
translation	47	8	20
nylon	46	1	46
spindle	46	4	42
utility	46	8	23
equilibrium	45	8	26
database	44	8	29
session	44	3	42
total revenues (TR)	44	4	38
utilization	44	10	31
carrier	43	6	35
grain size	43	2	34
transmit	43	6	29
articulation	42	1	42
manufactured	41	12	10
microprocessor	42	6	33
barchart	41	1	41
fetch	41	5	30
elongation	41	4	26
manufacturing process	41	7	30
DO loop	40	3	38
financial decision	40	1	40
host	40	8	27
milling	40	4	23
simulation	40	5	33
stiffness	40	5	27
cooling	39	7	29
declaration	39	3	37
lattice	39	2	23
moulding	39	1	39
working memory (WM)	39	3	36
demand curve	38	2	37
ferrite	38	2	37
formulation	38	8	27
pearlite	38	1	38
peg	38	2	37
constraint	37	9	27
controller	37	7	23
cutting edge	37	1	37
marketing concept	37	1	37
plastic deformation	37	3	23

strategic planning	37	4	30
textile	37	6	27
clause	36	4	28
magnetic tape	36	4	21
management services (MIS)	36	2	35
normal distribution	36	3	27
profitability	36	6	15
systems network architecture (SNA)	36	2	34
random access			
end effector	35	1	35
memory (RAM)	35	7	25
assembler	34	4	23
cassette	34	5	23
diffusion	34	4	31
hazard	34	2	22
marketing mix	34	2	31
numerical control (NC)	34	3	25
machine shop	34	3	24
macro-marketing	34	1	34
polypropylene	34	1	34
return on investment(ROI)	34	5	12
actuator	33	1	33
atomic	33	4	22
coefficient	33	10	15
frequency distribution	33	3	30
hardened	33	3	25
heating	33	7	21
label	33	6	26
involvement	33	6	16
portfolio	33	3	27
carbide	32	3	17
endurance limit	32	2	25
FORTRAN	32	4	24
integrated circuit	32	6	24
machine code	32	3	20
main memory	32	4	27
planner	32	3	21
primary memory	32	2	31
relative frequency	32	1	32
depreciation	31	3	19
mechanical properties	31	4	15
molecules	31	2	24
present value	31	4	27
pricing	31	6	17
query	31	5	26
subnet	31	1	31
torque	31	5	21
world-class manufacturing (WCM)	31	1	31
arithmetic and logic unit (ALU)	30	2	20
class interval	30	1	30
creep curve	30	2	28
deformed	30	4	19
fabrication	30	5	19
planned order release	30	1	30
radius	30	9	15
tensile strength	30	3	20
trait	30	3	26
behavioural science	29	1	29
control unit	29	5	16
finished goods	29	7	17
notch	29	4	20
non-repairable	29	1	29
ordinal type	29	1	29
production system	29	3	25
receiver	29	4	18
slip	29	7	15
tensile test	29	3	15
destination	28	4	24
ductility	28	3	18
IF statement	28	3	25
mainframe	28	9	15
offerings	28	2	25
percentile	28	2	25
production system	28	2	20
quartile	28	1	28
replenishment	28	4	18
racket	28	1	28
slot	28	4	23
strategy planning	28	1	28
broadcast	27	2	20
declared	27	3	20
dup	27	1	27
fracture toughness	27	2	23
grain boundary	27	2	21
hazard rate	27	1	27
job shop	27	4	20
marketing strategy	27	3	25
warehouse	27	8	16
precedence	27	5	19
rake angle	27	1	27
read only memory (ROM)	27	4	20
return on capital employed (ROCE)	27	2	26
subrange	27	2	22

working capital	27	5	21
assembly line	26	6	19
geometry	26	9	15
practitioner	26	1	26
compression	26	4	15
decimal number	26	6	12
fraction	26	10	15
modelling	26	9	15
ownership	26	9	15
practitioner	26	1	26
robotics	26	4	15
scientific management	26	4	12
tool face	26	1	26
binary number	25	3	13
cementite	25	1	25
cubic	25	4	18
data link	25	1	25
decnet	25	1	25
expected value	25	3	17
failure rate	25	1	25
feasibility	25	4	22
heated	25	4	16
IS/IT planning	25	1	25
polyester	25	1	25
shop order	25	1	25
tension	25	4	17
aluminium	24	4	16
business planning	24	2	21
capacity planning	24	5	18
computed	24	9	14
elasticity	24	7	16
memory location	24	5	16
on-line	24	6	17
penetration	24	4	15
potentiometer	24	1	24
reliability engineering	24	1	24
vendor	24	3	19
bearings	23	2	18
COBOL	23	3	21
component part	23	1	23
cond	23	1	23
cutting tool	23	3	19
distributor	23	2	21
indentation	23	3	19
physical layer	23	1	23
polyvinyl chloride (PVC)	23	1	23
strategic information system	23	1	23
thermoplastic	23	1	23
thermoset	23	1	23
variant	23	4	20
viscoelastic	23	1	23
conveyor	22	2	21
economics	22	10	11
emit	22	2	21
networking	22	7	14
packaging	22	5	13
session layer	22	1	22
artificial intelligence	21	7	12
bainite	21	1	21
carbon content	21	2	20
coaxial cable	21	1	21
digital computer	21	6	15
editing	21	5	14
FORTH word	21	1	21
hexadecimal	21	4	11
investment appraisal	21	1	21
local loop	21	1	21
master production schedule (MPS)	21	2	18
null hypothesis	21	1	21
operating system structure	21	1	21
organizational behaviour	21	2	20
packet switching	21	1	21
platter	21	3	15
polyethylene	21	1	21
polystyrene	21	1	21
pressed	21	7	15
psychological	21	4	18
recrystallization	21	2	19
broadcasting	20	3	14
centered	20	3	15
certainty	20	6	14
coupling	20	7	14
heading	20	6	13
mass production	20	9	10
outlay	20	3	18
polycarbonate	20	1	20
resource utilization	20	2	18
taper	20	3	18
target market	20	1	20
tempering	20	1	20
tool wear	20	1	20
top handle assembly	20	1	20
thermoplastics	20	1	20
algorithm	19	9	10
ALLOT	19	1	19
arm	19	7	11

computation	19	9	9
costing	19	6	14
IF structure	19	1	19
macro-marketing system	19	1	19
microscope	19	5	13
microwave	19	3	17
mill account	19	1	19
multiplexing	19	1	19
net requirement	19	1	19
output stock	19	1	19
polymers	19	1	19
predicate	19	2	17
servocontrol	19	1	19
SWAP	19	2	18
time value of money	19	2	17
valuation	19	3	11
yield point	19	3	13
acetal	18	1	18
acrylic	18	1	18
circuit switching	18	1	18
decision tree	18	1	18
dot diagram	18	1	18
impact test	18	2	8
inflation	18	6	10
lubricant	18	4	16
pegging	18	1	18
sensor	18	2	17
statistical methods	18	2	14
subscript	18	3	16
tachometric generator	18	1	18
tool life	18	1	18
transfer function	18	1	18
unloading	18	5	11
vehicle	18	5	10
activity cycle diagrams (ACD)	17	1	17
arpanet	17	1	17
capacity management	17	3	15
circuitry	17	7	9
connector	17	3	15
cylindrical	17	7	9
dislocation	17	2	13
ductile	17	3	14
FOR statement	17	3	15
grinding	17	4	11
incoming	17	1	17
mathematics	17	7	8
molecular	17	3	13
monetary	17	3	15
ord	17	1	17
optics	17	4	13
polyetheretherκetone (PEEK)	17	3	15
pressure group	17	1	17
project planning	17	1	17
rectangular	17	4	11
relaxation	17	3	15
rotor	17	1	17
amorphous	16	2	12
angular	16	2	13
broadband	16	1	16
data type	16	2	11
declaration part	16	1	16
end user	16	2	14
enumerated	16	3	13
eutectoid	16	1	16
feeding	16	7	8
geometric	16	3	10
initiation	16	3	12
inventory control	16	5	8
mean time between failures (MTBF)	16	2	14
minicomputer	16	7	10
projected available balance	16	1	16
psychology	16	2	15
return stack	16	1	16
secondary memory	16	1	16
shareholder	16	2	14
shortage	16	2	14
target customer	16	1	16
transport layer	16	1	16
velocity	16	4	12
waveform	16	1	16
absenteeism	15	2	14
armature	15	1	15
carbon steel	15	3	11
cos	15	5	9
cosine	15	5	8
cost accounting	15	4	11
cumulative frequency	15	2	14
economical	15	4	10
economists	15	5	10
error signal	15	2	14
exponent	15	3	13
header	15	2	14
heat treatment	15	2	14
inertia	15	2	14
noun	15	1	15

numeric	15	5	10
octal	15	3	13
output line	15	1	15
permissible cutting	15	1	15
polling	15	1	15
polyphenylene oxide (PPO)	15	1	15
retail	15	6	9
silicon	15	5	9
time-phrased record	15	1	15
tooling	15	1	15
ultimate tensile strength	15	3	11
assembly line	14	2	13
assembly language	14	2	12
assignment statement	14	2	11
binding	14	3	12
binomial distribution	14	1	14
break-even point (BEP)	14	3	11
clearance	14	1	14
cm	14	1	14
compound statement	14	2	11
computer network	14	2	13
discrete simulation	14	1	14
economic system	14	2	12
expert system	14	2	8
flexible manufacturing systems (FMS)	14	1	14
flow shop	14	3	10
floppy disk	14	3	11
fracture mechanics	14	2	12
material flow	14	2	13
front end	14	3	11
goto statement	14	2	11
hardness number	14	2	13
incompatible	14	4	11
intuitive	14	3	12
iteration	14	3	12
locus	14	4	12
management accounting system	14	1	14
micro-marketing	14	1	14
multiply	14	6	9
numerically	14	6	8
polymer	14	1	14
record type	14	1	14
recursion	14	1	14
reliability programme	14	1	14
seller	14	3	12
statistical process control (SPC)	14	1	14

subrange type	14	2	13
subroutine	14	2	13
system capacity	14	3	10
thermal conductivity	14	2	13
time scale	14	4	12
type definition	14	1	14
type identifier	14	1	14
variance	14	3	10
vector	14	2	13
absorption	13	4	11
acid	13	5	12
annealing	13	3	11
backtracking	13	1	13
binary number system	13	4	11
competitiveness	13	3	11
computational	13	6	8
duplex	13	1	14
evaluator	13	1	13
extreme value distribution	13	1	13
factory ledger	13	1	13
financing	13	4	10
flaw	13	2	12
FORTH operation	13	1	13
geometrical model	13	1	13
hex	13	2	12
input line	13	3	10
internal rate of return (IRR)	13	1	13
lubrication	13	4	10
machined surface	13	2	11
maintainability	13	4	11
modem	13	1	13
multiprogramming	13	3	12
network diagram	13	1	13
planning system	13	4	10
prime	13	2	11
pulse code modulation (PCM)	13	1	13
quotient	13	3	12
rationale	13	4	10
rotational	13	3	9
safety lead time	13	1	13
setq	13	1	13
step-down gear	13	1	13
structured type	13	1	13
subscript	13	1	13
subsystem	13	4	10
UNTIL loop	13	2	12
variable declaration	13	1	13

weaving	13	2	12
yardstick	13	3	10
yield stress	13	2	9
acceleration	12	3	9
addressing	12	4	9
adjective	12	3	10
alloying	12	3	8
arithmetical	12	5	8
banking	12	7	6
caller	12	1	12
carry-on-ban	12	1	12
chip formation	12	1	12
cube	12	2	10
discounts	12	4	8
electronically	12	5	8
facsimile	12	3	8
grammar	12	3	10
histogram	12	1	12
indexing	12	5	9
induction	12	3	9
information exchange	12	3	8
leverage	12	2	10
line printer	12	3	9
literal	12	5	10
marketing function	12	2	11
maximization	12	2	11
missile	12	2	11
molding	12	4	8
network layer	12	1	12
nextprime	12	1	12
nucleus	12	2	11
operands	12	3	9
polyethylene terephthalate (PET)	12	1	12
polytetrafluoroethylene (PTFE)	12	1	12
phonenumber	12	1	12
predefined	12	4	10
pure competition	12	1	12
quenched	12	1	12
radial	12	4	8
reorder level (ROL)	12	1	12
repairable system	12	1	12
separator	12	1	12
set type	12	1	12
signaling (CCIS)	12	1	12
sine	12	4	8
skewed	12	3	9
strategic management	12	2	11
subtraction	12	3	10
tailstock	12	1	12
thermal conductivity	12	2	7
tool material	12	1	12
transporting	12	6	6
true stress	12	1	12
unitary	12	3	10
valence electron	12	1	12
viscous	12	3	8
voltages	12	4	7
warranty	12	4	8
algorithms	11	6	6
alternating	11	4	8
amplitude	11	3	8
assortment	11	2	10
base type	11	1	11
card reader	11	3	9
carrier	11	3	8
comp	11	1	11
critical path analysis (CPA)	11	2	10
decision objective	11	1	11
design process	11	4	8
desirability factor	11	1	11
dice	11	2	10
digital transmission	11	1	11
dissolution	11	2	10
enterpreneur	11	4	8
eqn	11	1	11
equivalent annual cost	11	1	11
fatigue test	11	2	10
filler	11	3	9
fixed period system (FPS)	11	1	11
gear box	11	1	11
gripper	11	2	10
instruction set	11	1	11
job shop	11	2	10
impact strength	11	2	10
interpreter	11	2	10
leased	11	2	10
linkage	11	3	9
magnification	11	2	9
marketing strategy	11	3	8
maximum stock level (MSL)	11	1	11
necking	11	1	11
non-zero	11	1	11
planning	11	1	11
orthogonal	11	2	10

portfolio model	11	1	11
predictability	11	2	9
probability distribution	11	3	7
prog	11	1	11
programmable read only memory (PROM)	11	1	11
proprietary	11	3	9
public network	11	1	11
quarterly	11	6	9
record variable	11	1	11
red light	11	1	11
servo-motor	11	1	11
shear plane	11	1	11
shearing	11	3	9
simp	11	1	11
singularity	11	2	9
single chip	11	3	9
simulator	11	2	10
stack manipulation	11	1	11
stress strain curve	11	2	10
stressful	11	2	10
textfile	11	1	11
transfer line	11	2	10
transmitter	11	3	9
true strain	11	1	11
turret	11	2	10
undeformed chip thickness	11	1	11
verb	11	2	10
video display unit (vdu)	11	4	8
with statement	11	1	11
accounting profit	10	1	10
activity on arrow	10	1	10
alphanumeric	10	4	7
analog transmission	10	1	10
apics	10	2	9
APT	10	1	11
bracket assembly	10	1	10
capstan	10	1	10
case statement	10	2	9
cathode ray tube (CRT)	10	2	9
carburizing	10	2	9
census	10	2	9
continuous distribution	10	2	9
corrosion	10	3	6
crack propagation	10	2	9
cutting direction	10	1	10
dashpot	10	2	9
data link layer	10	1	10
definition part	10	1	10
direct labour	10	2	9
equilibrium price	10	2	9
ethics	10	2	9
field identifier	10	2	9
flapper	10	1	10
foamed	10	1	10
formability	10	1	10
gage length	10	1	10
high level language	10	2	9
human-resource	10	1	10
bill of materials (BOM)	10	1	10
insight	10	3	8
insulation	10	3	8
keyword	10	2	9
localized	10	3	8
mapping	10	3	8
matrices	10	4	7
mobile	10	3	8
non-linear	10	3	8
parse	10	2	9
perpendicular	10	3	8
polybutylene terephthalate (PBT)	10	1	10
predeclared function	10	1	10
presentation layer	10	1	10
production operation	10	2	8
reinforcement	10	3	8
semiconductor	10	4	6
sender	10	2	9
set constructor	10	1	10
sintered	10	2	8
special symbol	10	4	7
stator	10	1	10
stocked	10	2	9
tangential	10	3	8
transient	10	3	8
relative frequency			
teletypewriter	10	1	10
thermosetting	10	1	10
topology	10	1	10
trailing	10	2	8
transition temperature	10	3	8
undefined	10	2	9
unix	10	1	10
shear strength	10	1	10
uniform random number	10	1	10
value base	10	2	9
antenna	9	1	9
array type	9	1	9

articulated mechanical system (AMS)	9	1	9
articulated variable	9	1	9
backing storage	9	1	9
boolean algebra	9	1	9
break-even analysis	9	1	9
character string	9	2	8
computer graphics	9	1	9
computerized	9	3	6
congestion	9	2	8
copolymer	9	1	9
delimiter	9	1	9
deliverable	9	1	9
detector	9	2	8
dividend flow	9	1	9
fatigue fracture	9	2	8
flow shop	9	2	8
formaldehyde	9	1	9
FORTH arithmetic	9	1	9
function declaration	9	1	9
grained	9	2	7
headstock	9	1	9
heterogeneous	9	3	7
indentor	9	1	9
infix	9	2	8
mass storage device	9	2	8
mechanically	9	3	7
migration	9	2	8
multiplexer	9	1	9
network analysis	9	2	8
network architecture	9	2	8
nuclei	9	2	8
operational decision	9	1	9
packed array	9	1	9
phenol	9	1	9
planning horizon	9	4	6
polyamide	9	1	9
polyurethane	9	1	9
prefix	9	3	7
print statement	9	2	8
project plan	9	1	9
promotional	9	1	9
prospective	9	2	8
recruitment	9	2	7
regeneration	9	1	9
relational operator	9	1	9
resilience	9	2	8
resource productivity	9	1	9
retained austenite	9	1	9

shop floor control	9	3	8
shorthand	9	2	7
simple expression	9	1	9
stipulated	9	2	8
surface roughness	9	1	9
tempered	9	1	9
tensile stress	9	3	6
time bucket	9	1	9
timestable	9	1	9
torsion	9	2	8
thermal properties	9	2	8
tradeoffs	9	3	6
truncated	9	3	6
twisted pair	9	1	9
WHILE loop	9	1	9
while statement	9	1	9
yellow light	9	1	9
address bus	8	2	7
algol	8	2	6
append	8	2	7
application layer	8	1	8
application software	8	2	7
ascii value	8	1	8
austenitic	8	1	8
baseband	8	1	8
carbon atom	8	2	7
cardinal	8	1	8
cartridge	8	2	7
cellulose	8	1	8
chr	8	1	8
close-packed plane	8	1	8
const	8	2	7
conveyor	8	2	7
crystalline plastics	8	1	8
cutter	8	3	6
decision-making	8	1	8
declarative	8	2	7
defectives	8	1	8
degradation	8	1	8
denominator	8	4	5
discount	8	4	5
draughting	8	1	8
drill press	8	2	6
drilling	8	1	8
echo suppressor	8	1	8
electromagnetic	8	4	5
fatigue crack	8	2	7
flowchart	8	2	7
flow of materials	8	3	6

gauge	8	1	8
geographical	8	2	6
hexagonal	8	2	6
hydraulic system	8	2	7
incentive pay	8	3	6
inductor	8	1	8
instantiated	8	1	8
interarrival value	8	1	8
interstitial	8	2	7
kinematic model	8	1	8
lathe	8	1	8
logarithmic	8	2	6
machine language	8	2	6
magnet	8	1	8
manpower	8	2	6
materials handling	8	3	6
mechanized	8	2	6
memory unit	8	2	7
mousetrap	8	1	8
multiplexed	8	1	8
multiprocessor	8	2	6
negate	8	3	6
nonprofit	8	1	8
normal stack	8	1	8
nucleation	8	2	7
ozone	8	3	6
peer process	8	1	8
pegs	8	1	8
phasing	8	2	7
piston shaft	8	1	8
planing	8	1	8
plastically	8	3	6
point process	8	1	8
polysulphone	8	1	8
positional servo-system	8	1	8
precipitation	8	1	8
pred	8	1	8
punctuation	8	3	6
residual stress	8	3	6
removal	8	1	8
rubbing	8	2	7
scientific theory	8	2	7
selector	8	2	7
servovalve	8	1	8
simulated	8	2	7
spinning	8	2	7
spool	8	1	8
statistic	8	3	6
statistical model	8	2	7
stockturn	8	1	8
stoppage	8	2	6
stress concentration	8	1	8
subscripted variable	8	1	8
tabulator	8	1	8
target marketing	8	1	8
total preventive maintenance (TPM)	8	1	8
tool point	8	1	8
tonnes	8	1	8
unit cell	8	2	7
vectors	8	2	6
viscoelastic material	8	1	8
vocabulary	8	2	7
wedge	8	2	7
wordlength	8	2	7
adhesion	7	2	6
alignment	7	1	7
analytical engine	7	1	7
aptitude	7	2	6
brittle fracture	7	2	6
bucketless	7	1	7
calibrated	7	2	6
cellular	7	1	7
constant definition	7	1	7
conversion process	7	3	5
clamping	7	2	6
commercially	7	2	6
compile	7	2	6
cracking	7	3	5
data link control	7	1	7
decomposition	7	3	5
dentist	7	1	7
descriptive measure	7	1	7
design capacity	7	1	7
dielectric	7	1	7
digitized	7	1	7
direct labor	7	1	7
dp application	7	2	6
educator	7	1	7
elastic limit	7	2	6
electromechanical	7	2	6
engineering plastics	7	1	7
epoxy	7	1	7
evolutionary	7	2	5
exception code	7	1	7
exponential distribution	7	2	6
fabricated	7	3	5
fabricating	7	3	5

fast Fourier transform (FFT)	7	1	7
frequency division multiplexing (FDM)	7	1	7
file type	7	1	7
functor	7	1	7
flow industry	7	2	6
form ι tility	7	2	6
fourth generation language	7	1	7
fractured	7	2	6
gamma distribution	7	1	7
glass fibre	7	1	7
housekeeping	7	3	5
identified variable	7	1	7
impact value	7	1	7
index type	7	1	7
industrialized	7	3	5
interatomic	7	1	7
intercept	7	2	6
inversion	7	2	5
invoicing	7	3	5
isometric	7	1	7
kilometer	7	1	7
laser printer	7	3	5
lead time offset	7	1	7
line diagram	7	1	7
logic gate	7	2	6
lot sizing	7	1	7
low density polyethylene (LDPE)	7	1	7
lubricating	7	3	5
machinability	7	3	5
manufacturing cost	7	2	6
manufacturing lead time	7	1	7
marketer	7	3	5
mass selling	7	1	7
mass marketing	7	1	7
meter	7	3	5
modeling	7	2	6
nibble	7	1	7
non-repairable item	7	1	7
oligopoly	7	1	7
operational system	7	1	7
overload	7	3	5
parameter value	7	2	5
physical resource	7	2	6
plate martensite	7	1	7
pointer value	7	1	7
poisson distribution	7	1	7
polythene	7	1	7
possession utility	7	2	6
pre-automation	7	1	7
process technology	7	2	6
PROLOG system	7	1	7
queuing	7	3	5
retailing	7	3	5
remuneration	7	2	6
repeat statement	7	1	7
resolver	7	1	7
salvage value	7	1	7
sampling distribution	7	1	7
secant modulus	7	1	7
servosystem	7	1	7
set point	7	2	6
shipments	7	3	5
solvent	7	1	7
superposition	7	1	7
synchro	7	1	7
tangent	7	2	5
temporary entities	7	1	7
theme	7	1	7
time dimension	7	2	5
torsional	7	3	5
total float	7	1	7
translational	7	1	7
transmission medium	7	1	7
transport and service systems	7	1	7
transport connection	7	1	7
unsigned integer	7	1	7
uniaxial	7	2	6
variable gain amplifier	7	1	7
welded	7	3	5
wholesalers	7	3	5
additives	6	1	6
algebraic	6	2	5
aligned	6	2	5
anvil	6	1	6
apostrophe	6	1	6
assoc	6	1	6
bargaining	6	1	6
bathtub	6	1	6
bit rate	6	1	6
blueprint	6	1	6
bookkeeping	6	2	5
brittleness	6	2	5
categorisation	6	1	6
cemented	6	1	6

chip friction	6	1	6
chuck	6	2	5
clamped	6	2	5
cleavage fracture	6	1	6
coating	6	2	5
cost management	6	1	6
creep behaviour	6	1	6
creep test	6	2	5
crosspoint switch	6	1	6
cutting operation	6	1	6
defect	6	1	6
defun	6	1	6
demand schedule	6	1	6
determiner	6	2	5
diesel	6	1	6
directives	6	2	5
discounted payback	6	1	6
distributed system	6	1	6
effective marketing	6	2	5
elastomer	6	1	6
engineering stress	6	2	5
esc	6	1	6
external	6	1	6
financial intermediary	6	1	6
fixed quantity system (FQS)	6	1	6
flammability	6	1	6
flexural stress	6	1	6
flow industry	6	1	6
FORTH variable	6	2	5
free float	6	1	6
harmonics	6	1	6
headend	6	1	6
homogenizing	6	2	5
hoop	6	1	6
housing	6	1	6
incremental encoder	6	1	6
indentations	6	2	5
indenter	6	1	6
industrial robot	6	1	6
intermittent loading	6	1	6
ionic	6	1	6
ISDN interface	6	1	6
item master file	6	1	6
iterative	6	2	5
kilogram	6	1	6
jet spray	6	1	6
level of automation	6	2	5
linear piston	6	1	6

LISP expression	6	1	6
LLDPE	6	1	6
looping	6	1	6
lubricated	6	2	5
maclisp	6	1	6
management development	6	1	6
manganese	6	2	5
marketing management			
metallic bond	6	1	6
metallurgical	6	1	6
microelectronics	6	2	5
molten	6	2	5
moulded	6	1	6
nonmetallic	6	1	6
online	6	2	5
order launching	6	1	6
payoffs	6	2	5
pert	6	2	5
physics	6	2	5
pneumatic system	6	1	6
pre-defined	6	1	6
process	6	1	6
monopolistic competition	6	1	6
MTYPE	6	1	6
multidrop	6	1	6
non-profit organization	6	1	6
object code	6	1	6
phase shift	6	1	6
platelets	6	1	6
procedure and function declaration	6	1	6
process planning	6	1	6
questionnaire	6	2	5
rectangle	6	1	6
redesigned	6	2	5
reorder level	6	1	6
repaired	6	1	6
resin	6	1	6
rotary	6	1	6
salaried	6	2	5
scrap value	6	1	6
serial printer	6	1	6
service production systems (SPS)	6	1	6
shear angle	6	1	6
slip system	6	1	6
statistical population	6	1	6
statistically	6	2	5
stepping motor	6	1	6

supply system	6	2	5
swarf	6	1	6
textiles	6	2	5
time division multiplex	6	1	6
time division switch	6	1	6
total quality control (TQC)	6	1	6
transponder	6	1	6
ungrouped	6	1	6
vacancy	6	2	5
vee	6	1	6
wearout	6	1	6
Weibull distribution	6	1	6
WIP tracking	6	1	6
workflow	6	1	6
abort	5	1	5
abrasion	5	1	5
adverb	5	1	5
address space	5	1	5
analog-digital			
audit	5	1	5
converter (ADC)	5	1	5
anatomy	5	1	5
application program	5	1	5
appliance	5	1	5
appliance plug	5	1	5
apprenticeship	5	1	5
articulated mechanical			
chain	5	1	5
auction	5	1	5
austenite grain	5	1	5
automatic lathe	5	1	5
barprint	5	1	5
bottom-up replanning	5	1	5
bureaucracy	5	1	5
capital rationing	5	1	5
caravan	5	1	5
centralised network	5	1	5
chemistry	5	1	5
climatic	5	1	5
clientpointer	5	1	5
clip	5	1	5
codec	5	1	5
compensated	5	1	5
compensation (PD)	5	1	5
competency	5	1	5
concentration differential	5	1	5
continuous simulation	5	1	5
continuous function	5	1	5
coolant	5	1	5

coordinate axe	5	1	5
covalent bond	5	1	5
creep strength	5	1	5
crystal lattice	5	1	5
data flow control	5	1	5
data statement	5	1	5
data transfer rate	5	1	5
decision support			
system (DSS)	5	1	5
dialect	5	1	5
digital bit pipe	5	1	5
direct access storage	5	1	5
discontinuous	5	1	5
discrete change	5	1	5
discrete Fourier			
transform (DFT)	5	1	5
disk pack	5	1	5
dissemination	5	1	5
distributed network	5	1	5
elastic demand	5	1	5
empty statement	5	1	5
encode	5	1	5
expectancy	5	1	5
experiential learning gap	5	1	5
factory management	5	1	5
fractional recovery	5	1	5
frequency response	5	1	5
fused	5	1	5
getreplies	5	1	5
glass transition			
hazard function	5	1	5
inventory decision	5	1	5
temperature	5	1	5
historian	5	1	5
homopolymer	5	1	5
hub polling	5	1	5
human-resource planning	5	1	5
hypoeutectoid	5	1	5
identically	5	1	5
in-band signaling	5	1	5
incorporation	5	1	5
inelastic demand	5	1	5
informative	5	1	5
injection moulding	5	1	5
integer value	5	1	5
integrated manufacturing			
internal accounting system	5	1	5
intrinsic function	5	1	5
isochronous graph	5	1	5

isothermal	5	1	5
lamellae	5	1	5
latch	5	1	5
lath	5	1	5
life cycle costs (LCC)	5	1	5
liquidity requirement	5	1	5
location address	5	1	5
lognormal distribution	5	1	5
magnet	5	1	5
magnetic drum	5	1	5
market-directed economy	5	1	5
message switching	5	1	5
mnemonics	5	1	5
mold	5	1	5
motel	5	1	5
movement sensor	5	1	5
NAND gate	5	1	5
negative deviation	5	1	5
network addressable units (NAU)	5	1	5
newprime	5	1	5
non nil determinant	5	1	5
nonmanufacturing	5	1	5
optic	5	1	5
perceptual	5	1	5
perspective	5	1	5
pertchart	5	1	5
phenolic	5	1	5
plasticity	5	1	5
polycrystalline	5	1	5
polyethersulphone	5	1	5
polymeric material	5	1	5
positive deviation	5	1	5
positive ion	5	1	5
precedence convention	5	1	5
precept	5	1	5
pretty printer	5	1	5
primary bond	5	1	5
procedure statement	5	1	5
production orientation	5	1	5
production system (IMPS)	5	1	5
program counter	5	1	5
prom	5	1	5
protons	5	1	5
propagation delay	5	1	5
putprop	5	1	5
quantify	5	1	5
rally	5	1	5
random access storage	5	1	5

rate of occurrence of failures (ROCOF)	5	1	5
readjustment	5	1	5
regression line	5	1	5
relative frequency line diagram	5	1	5
resatisfy	5	1	5
reschedule	5	1	5
retransmitted	5	1	5
revolving	5	1	5
ring network	5	1	5
rivets	5	1	5
sample interquartile range	5	1	5
s-normal distribution	5	1	5
self-excited vibration	5	1	5
semantics	5	1	5
semiannual	5	1	5
sequential access storage	5	1	5
servocontrolled	5	1	5
setpart	5	1	5
shear fracture	5	1	5
signal-to-noise ratio	5	1	5
simple configuration	5	1	5
simple data type	5	1	5
simplifier	5	1	5
source code	5	1	5
spindle speed	5	1	5
squared deviation	5	1	5
stabiliser	5	1	5
standard error of the difference	5	1	5
standard linear solid	5	1	5
stereotype	5	1	5
strategic business units (SBU)	5	1	5
strategic framework	5	1	5
strategy decision	5	1	5
stress-strain relation	5	1	5
structural foam	5	1	5
styrene	5	1	5
subchannel	5	1	5
sublist	5	1	5
subscriber	5	1	5
swivel	5	1	5
SWOT analysis	5	1	5
synthetic	5	1	5
telecommnunication	5	1	5
telemarketing	5	1	5
terrestrial link	5	1	5

three-era	5	1	5
three-era model	5	1	5
three-phase rule	5	1	5
tool profile	5	1	5
transmission control	5	1	5
trimmed mean	5	1	5
trunc	5	1	5
tuner	5	1	5
tungsten carbide	5	1	5
uncontrollable	5	1	5
unimodal	5	1	5
unpack	5	1	5
unplasticised	5	1	5
variable identifier	5	1	5
variate	5	1	5
verbatim	5	1	5
viscoelastic behaviour of plastics	5	1	5
wear rate	5	1	5
welding	5	1	5
word symbol	5	1	5

Appendix Three
Subtechnical Vocabulary Identified in the Large Corpus

This is a list of all subtechnical vocabulary identified in the large corpus, arranged in rank order of frequency and range. The first figure following the word form is the frequency and the second range.

basic	426	26
variable	398	15
fig.	356	9
equation	247	20
user	247	18
complex	234	26
techniques	215	24
overall	176	23
environment	172	21
characteristics	163	23
relatively	157	25
role	153	21
normally	149	24
sequence	148	18
competitive	133	13
potential	132	21
procedures	129	18
procedure	121	22
logical	110	11
associated	109	23
external	107	15
typically	107	17
focus	106	12
ensure	103	24
differentiate	101	20
initial	100	22
specified	100	21
corresponding	99	19
implementation	99	16
technique	99	17
alternative	97	20
related	95	20
fundamentally	93	18
multiple	91	18
accuracy	84	18
evaluate	83	18
symbol	83	16
context	81	17
emphasis	81	14
minimum	77	22
monitor	77	10
symbols	77	17
dependent	74	17
oriented	72	13
preceding	72	15
traditional	72	18
similarly	71	19
primarily	70	17
summary	70	17
calculations	69	20
alternatives	68	9
plus	68	18
complexity	67	18
potentially	66	17
sophisticated	66	19
capabilities	65	14
setting	65	18
conventional	64	18
matrix	63	13
relevant	62	18
representation	61	15
subjected	60	14
complications	59	14
effectively	59	17
mechanism	57	12
respectively	57	18
drawing	56	17
uncertainty	56	15
scope	54	20
batch	53	13
calculation	53	18
essentially	53	17
evaluation	53	14
specification	53	16
effectiveness	52	15
generate	51	18
chart	51	17
subsequent	51	17
valid	51	15
complicated	50	18
measurement	50	19
appendix	49	10
acceptable	48	15
evaluated	48	15
corporate	47	9
duration	47	12
modified	47	15
reliable	47	16
variability	47	10
advanced	46	16
availability	46	17
increasingly	46	12
communicate	45	11
cycles	44	13
decrease	44	14
stability	44	10
version	44	14
capability	43	13
feasible	43	12
simplified	43	13
static	43	12
virtually	43	17
correctly	42	14
dynamic	42	15
extended	42	16
generated	42	16
implications	42	13
integration	42	14
monitors	41	10
criterion	40	9
dimensions	40	15
core	39	13
eventually	39	18
feedback	39	10
flexibility	39	12
intermediate	39	13
offering	39	9
television	39	10
approximately	38	15
environmental	38	11
functional	38	12
technological	38	13
implemented	37	12
spectrum	37	13
accurately	36	16
environments	36	14
exclusive	36	10
planners	36	9
precise	36	17

precision	36	13
categories	35	15
classification	35	14
consistent	35	17
designers	35	10
evolution	35	11
simultaneously	35	16
specify	35	15
evaluating	34	15
manual	34	14
readily	34	17
trend	34	8
validity	34	11
denoted	33	8
inspection	33	11
percentage	33	17
quantitative	33	9
senior	33	7
successive	33	12
versions	33	11
perspective	32	12
segments	32	8
specialists	32	9
allocation	31	8
analogous	31	11
experimental	31	6
inference	31	6
magnitude	31	11
rigid	31	15
status	31	13
theoretical	31	12
underlying	31	12
vertical	31	11
electronics	30	10
initially	30	13
mechanisms	30	12
media	30	9
specifications	30	12
formula	29	11
intensive	29	10
internally	29	10
limitations	29	14
motivation	29	11
options	29	9
repetitive	29	11
specifies	29	10
standardized	29	11
ultimately	29	15
versa	29	16
versus	29	11
allocated	28	8
arbitrary	28	11
assess	28	12
assumptions	28	10
continuously	28	12
decades	28	9
enclosed	28	10
format	28	10
identical	28	11
implement	28	12
integral	28	7
interaction	28	10
interactive	28	7
interpretation	28	11
straightforward	28	16
transaction	28	10
widespread	28	11
adjacent	27	12
amplitude	27	5
analogy	27	10
applicable	27	11
basically	27	12
category	27	13
criteria	27	10
finite	27	10
forecast	27	11
forecasting	27	9
layout	27	10
producer	27	10
schematic	27	10
technologies	27	9
transmitted	27	6
approximation	26	6
breakdown	26	12
disadvantages	26	10
evolved	26	12
focused	26	8
generating	26	9
rely	26	9
researchers	26	8
roles	26	9
separately	26	14
setup	26	15
subsequently	26	16
compact	25	9
compiled	25	10
complement	25	6
expansion	25	11
horizontal	25	9
necessitate	25	6
optical	25	7
perception	25	7
significantly	25	8
specifically	25	13
terminology	25	11
usage	25	10
attached	24	12
classified	24	13
diverse	24	12
indirect	24	7
outcomes	24	7
systematic	24	8
transistor	24	7
unlikely	24	12
alternatively	23	14
client	23	5
dimensional	23	9
efficiently	23	10
essence	23	12
implementing	23	12
innovation	23	10
micro	23	10
modify	23	11
payroll	23	8
phenomena	23	11
physically	23	10
priority	23	7
specifying	23	9
academic	22	8
accessed	22	7
crucial	22	13
mutually	22	8
sequences	22	11
skilled	22	9
uncertainties	22	8
acceptance	21	10
assessing	21	10
assumption	21	8
classical	21	7
depicted	21	7
designated	21	6
facilitate	21	11
inadequate	21	11
indication	21	8
interact	21	10
macro	21	5
maximize	21	7

monitoring	21	8
priorities	21	9
simplify	21	13
specialized	21	10
approximate	20	11
commitment	20	8
conjunction	20	10
devoted	20	13
earnings	20	4
empirical	20	7
hierarchy	20	7
incorporated	20	11
independently	20	10
inherent	20	12
minus	20	11
participation	20	4
replacement	20	10
sliding	20	7
survival	20	4
analysed	19	10
asset	19	5
coordination	19	9
disks	19	6
distortion	19	9
enhance	19	12
explicit	19	8
hierarchical	19	9
limitation	19	10
manually	19	8
restricted	19	9
advent	18	10
analysing	18	11
compete	18	11
cope	18	9
denote	18	6
differently	18	10
disadvantage	18	12
displacement	18	8
duplicate	18	6
global	18	8
inevitably	18	5
integrate	18	7
outcome	18	9
overlap	18	11
routing	18	6
sector	18	8
totally	18	12
trends	18	9
tv	18	10
update	18	6
utilize	18	8
analyzing	17	11
awareness	17	7
calculator	17	7
characterized	17	10
domain	17	8
experienced	17	11
explicitly	17	8
identification	17	10
incorporate	17	9
installation	17	8
interactions	17	9
methodology	17	5
partial	17	8
predictable	17	9
react	17	9
resultant	17	4
retailer	17	5
specialist	17	7
subjective	17	6
summarized	17	8
worldwide	17	7
accommodate	16	10
adjustments	16	8
airline	16	10
analyses	16	7
auto	16	5
axial	16	5
competing	16	10
comprises	16	6
conceptual	16	8
correlation	16	7
creative	16	7
exert	16	5
fluctuations	16	7
increments	16	7
lesser	16	14
maturity	16	5
modules	16	6
multiplication	16	7
option	16	8
periodic	16	7
reserved	16	9
reset	16	4
summarize	16	10
supervisor	16	5
terminated	16	5
transmitting	16	5
twentieth	16	6
valued	16	8
virtual	16	3
adjustment	15	8
assets	15	5
autonomous	15	5
competitor	15	6
complementary	15	5
consistency	15	8
dealt	15	9
dedicated	15	5
depicts	15	5
ensures	15	7
historically	15	5
homogeneous	15	6
imperfections	15	5
informal	15	6
mid	15	5
minimize	15	9
modification	15	9
procedural	15	4
regardless	15	9
rotating	15	5
symbolic	15	5
throughput	15	7
traditionally	15	9
utilized	15	8
visual	15	9
airlines	14	5
assessment	14	7
clients	14	5
coherent	14	7
comparable	14	11
completion	14	10
conversely	14	8
cyclic	14	5
dependence	14	7
drawings	14	8
enhanced	14	7
establishment	14	7
estimation	14	5
fitting	14	7
focusing	14	5
formulae	14	6
generates	14	9
initiated	14	5
meaningful	14	10
module	14	5
optimum	14	7

passive	14	5
predetermined	14	11
preparation	14	8
randomly	14	6
repetition	14	9
respective	14	7
segment	14	9
turnover	14	5
undesirable	14	10
updated	14	5
variants	14	6
viewpoint	14	5
accountant	13	9
adequately	13	11
allocate	13	8
aluminum	13	5
ambiguity	13	8
analyse	13	8
characterised	13	3
communicating	13	8
comprehensive	13	7
denotes	13	4
differentiation	13	3
dramatically	13	7
experimentally	13	4
externally	13	9
forecasts	13	7
guidelines	13	7
harmonic	13	5
implementations	13	6
incurred	13	6
inexpensive	13	6
instantaneously	13	5
interrelated	13	8
manipulate	13	9
markedly	13	6
massive	13	7
minimizing	13	6
optimal	13	8
originated	13	8
permanently	13	9
progressive	13	5
realistic	13	12
reasonably	13	11
rewrite	13	4
rigidity	13	6
schematically	13	7
simulate	13	8
subtract	13	10
summarised	13	6
summarizes	13	6
tabulated	13	5
tedious	13	9
updating	13	8
accessible	12	8
acquisition	12	8
analyze	12	7
assessed	12	7
bias	12	7
comprise	12	7
continual	12	7
coordinated	12	3
coordinates	12	5
correction	12	8
decentralized	12	7
dispersed	12	5
dominant	12	5
evolving	12	5
excessive	12	8
expertise	12	9
graphical	12	6
graphically	12	7
handy	12	4
imaginary	12	5
indirectly	12	8
integrating	12	8
interconnected	12	7
jargon	12	6
likelihood	12	6
monitored	12	6
peripheral	12	6
predecessor	12	4
prerequisite	12	6
qualitative	12	7
rating	12	6
readable	12	5
relevance	12	6
sensitivity	12	8
simplistic	12	4
simultaneous	12	7
standardization	12	4
substantially	12	8
successively	12	7
summation	12	5
tangible	12	6
tolerance	12	6
adoption	11	4
aggregate	11	6
aggressive	11	5
ambiguous	11	6
auxiliary	1	6
batches	11	5
clarity	11	6
conceptually	11	8
decade	11	7
dial	11	3
ensuring	11	6
evaluates	11	3
exerted	11	4
exposure	11	5
extensively	11	6
ideally	11	8
incorrect	11	9
inspect	11	7
instantaneous	11	7
interchangeable	11	5
interference	11	6
marketplace	11	8
mesh	11	4
misleading	11	8
mobility	11	3
nearby	11	5
nickel	11	5
overtime	11	5
popularity	11	5
predictions	11	5
rational	11	6
reproduced	11	7
signifies	11	5
simplification	11	6
suitability	11	6
supervision	11	8
terminating	11	4
unchanged	11	6
vertically	11	7
warning	11	8
abbreviated	10	5
accessing	10	5
administrative	10	6
bin	10	5
cantilever	10	4
cognitive	10	3
compensate	10	7
compilation	10	7
compressive	10	4
cumbersome	10	7
deflection	10	4

delete	10	4
distorted	10	5
exclusively	10	7
fractions	10	4
generic	10	5
gradual	10	7
hybrid	10	7
inappropriate	10	7
incomplete	10	6
inefficient	10	8
initiate	10	7
insufficient	10	9
integrity	10	4
interchange	10	5
interdependence	10	4
investors	10	5
isolated	10	5
logically	10	7
overview	10	8
partially	10	7
periodically	10	5
prediction	10	6
prone	10	7
redundant	10	5
reinforced	10	7
repeatedly	10	6
robust	10	6
satellite	10	6
specialization	10	5
stressed	10	6
successor	10	3
temporarily	10	6
theoretically	10	6
trivial	10	8
typewriter	10	5
unacceptable	10	7
uniformly	10	4
unrelated	10	4
utilisation	10	4
verbal	10	5
wiring	10	5
workpieces	10	4
workplace	10	4
abstract	9	7
additionally	9	5
amplification	9	3
appreciably	9	3
appropriateness	9	3
audio	9	3
automate	9	3
backward	9	5
basics	9	6
bike	9	2
broadly	9	8
coil	9	2
collective	9	2
commas	9	3
communicated	9	4
compatible	9	5
constituents	9	4
cultural	9	3
defence	9	3
embedded	9	6
equivalents	9	4
facilitates	9	6
failing	9	6
focuses	9	5
graphic	9	4
implicit	9	5
inclusion	9	7
indicators	9	5
inflows	9	4
innovative	9	5
insight	9	8
irrespective	9	7
labelled	9	5
landing	9	3
manipulated	9	6
milliseconds	9	4
minimized	9	4
modifications	9	7
necessitates	9	3
obsolete	9	6
optional	9	4
orderly	9	6
orientated	9	5
platters	9	2
poorly	9	7
premium	9	6
protective	9	5
pyramid	9	5
reasoning	9	5
rectangles	9	3
retrieved	9	5
risky	9	4
rot	9	2
sectors	9	7
sequentially	9	6
span	9	3
stockroom	9	3
subordinate	9	4
subset	9	4
symptoms	9	6
technicians	9	2
terminate	9	4
thorough	9	6
treasurer	9	3
unforeseen	9	6
unpredictable	9	6
usable	9	7
vehicles	9	4
versatile	9	6
accordance	8	7
allocating	8	6
analysts	8	4
arbitrarily	8	6
calculators	8	5
categorized	8	3
cited	8	6
clerical	8	6
coordination	8	3
compressed	8	6
comprising	8	7
cone	8	4
consequent	8	4
consistently	8	5
constituent	8	5
contractor	8	3
controllable	8	3
conveniently	8	5
defensive	8	5
deliberate	8	7
departmental	8	6
deterioration	8	4
distinctive	8	5
diversity	8	6
documentation	8	5
drawback	8	5
dynamics	8	5
economically	8	5
emergence	8	4
excluded	8	3
exploited	8	5
fixtures	8	5
formulated	8	6
guidance	8	7
hamburger	8	2

hourly	8	3
importantly	8	7
increment	8	3
indicator	8	5
ingredient	8	7
instrumentation	8	6
insulating	8	3
interdisciplinary	8	4
internationally	8	5
manipulating	8	7
methodologies	8	4
motivate	8	4
multiples	8	5
outlays	8	3
outset	8	7
peripherals	8	5
pollution	8	4
propagate	8	4
qualified	8	4
radiation	8	4
recommendations	8	4
relied	8	6
relies	8	6
repetitions	8	4
retention	8	4
retrieve	8	6
rotate	8	5
shortages	8	4
similarity	8	8
simplifies	8	5
socially	8	2
specialised	8	4
specialize	8	3
subordinates	8	4
susceptible	8	4
tapers	8	3
theorem	8	4
tidy	8	4
underline	8	5
unreliable	8	5
usefulness	8	5
valve	8	4
accelerated	7	3
accumulator	7	2
adaptable	7	6
aggregated	7	3
albeit	7	4
alternate	7	5
assertion	7	7
cellular	7	4
clarify	7	6
classifications	7	4
classify	7	6
colleagues	7	3
comma	7	5
compatibility	7	3
complexities	7	5
composite	7	5
conforms	7	4
consecutive	7	4
consensus	7	5
consultants	7	5
corrective	7	4
deterministic	7	4
diagnosis	7	5
diagnostic	7	5
dialogue	7	3
distort	7	4
diversification	7	2
dominated	7	4
dummy	7	3
elimination	7	6
exceptional	7	6
exhaustive	7	5
exploit	7	3
exploitation	7	4
exploration	7	4
filter	7	5
formulate	7	4
formulating	7	5
gearbox	7	3
headings	7	4
incidence	7	3
inclined	7	3
incremented	7	5
indefinitely	7	4
individually	7	4
ingredients	7	4
innovations	7	3
intersection	7	5
intervention	7	5
linearly	7	4
manuals	7	4
mbytes	7	2
meaningless	7	6
methodological	7	3
microscopic	7	3
midpoint	7	3
miscellaneous	7	4
modulated	7	2
monopoly	7	4
motivated	7	5
obsolescence	7	5
onhand	7	3
outstanding	7	4
overriding	7	5
participative	7	3
pendulum	7	2
perceptions	7	3
pertaining	7	4
petroleum	7	4
pose	7	5
posed	7	6
postal	7	3
procurement	7	4
progressively	7	4
prototype	7	4
retirement	7	5
rework	7	2
roller	7	3
seasonal	7	4
segmentation	7	3
settings	7	4
shelf	7	4
similarities	7	4
specially	7	5
specialty	7	5
strict	7	4
subassemblies	7	3
subtracting	7	5
surprisingly	7	7
technically	7	4
transmits	7	3
transparent	7	5
usefully	7	5
valves	7	4
verify	7	4
wasteful	7	6
aerospace	6	5
allocates	6	3
allocations	6	2
alphabet	6	3
ambulance	6	3
amplified	6	3
amplify	6	4
attainable	6	2
attentive	6	2

backup	6	2
beverages	6	4
bibliography	6	3
breakthrough	6	4
calendar	6	2
chaos	6	2
cheque	6	3
classifying	6	4
coiled	6	3
collectively	6	4
commonplace	6	2
composites	6	2
conduction	6	3
containers	6	4
contraction	6	4
contractors	6	2
decibels	6	3
deleted	6	3
derivatives	6	4
dilemma	6	4
discretion	6	5
disposable	6	3
disposal	6	6
disruption	6	3
dual	6	4
duplication	6	2
elapsed	6	4
elementary	6	5
enhancements	6	3
equality	6	4
equity	6	5
erroneous	6	4
exciting	6	4
exclude	6	5
exit	6	4
exponentially	6	2
faulty	6	4
filtered	6	2
filters	6	2
flaws	6	3
foremen	6	2
formulas	6	3
frustration	6	3
groupings	6	5
highlight	6	5
hypotheses	6	3
inability	6	5
incorporates	6	5
inherently	6	5
inspected	6	5
intensity	6	5
interconnection	6	2
interdependent	6	5
investigators	6	2
justification	6	5
laborious	6	3
minded	6	4
modifying	6	4
multiplications	6	2
negotiate	6	5
newclient	6	2
nitrogen	6	3
nomenclature	6	4
notches	6	3
nuclear	6	6
observable	6	2
oscillations	6	2
outermost	6	4
outsiders	6	4
parity	6	3
pertinent	6	6
pipeline	6	3
portable	6	3
postulated	6	3
printout	6	3
proliferation	6	4
quantifiable	6	4
raisers	6	2
readability	6	4
recursive	6	2
redundancy	6	3
refined	6	5
regional	6	3
relying	6	4
researcher	6	3
respondents	6	2
restrict	6	6
rivalry	6	2
rotates	6	3
satisfactorily	6	4
scan	6	5
scanning	6	4
scoop	6	2
semiconductors	6	3
setup	6	2
setups	6	3
shortcomings	6	4
shortened	6	6
slack	6	4
solidify	6	2
spanning	6	4
sporting	6	3
subsidiary	6	4
substitutional	6	2
suitably	6	5
superficial	6	2
supplement	6	5
symmetrical	6	2
synchronization	6	2
synonymous	6	4
terminates	6	4
tester	6	2
textual	6	3
thousandth	6	2
transparency	6	4
uncommon	6	4
understandable	6	5
unidirectional	6	4
universally	6	6
unload	6	3
unstable	6	3
unstructured	6	4
urgent	6	4
vastly	6	3
welders	6	2
wholesale	6	4
withstand	6	5
zinc	6	3
academics	5	5
accelerating	5	4
acknowledgement	5	2
activated	5	3
adjustable	5	4
advantageous	5	5
advisable	5	4
aeroplane	5	3
agenda	5	2
alphabetical	5	3
alphabetically	5	2
analyzed	5	3
appended	5	2
appraise	5	3
appreciable	5	4
architectural	5	4
assertions	5	2
assesses	5	3
athletic	5	2

attenuation	5	2
automobiles	5	4
automotive	5	3
beneficial	5	3
billion	5	5
buffered	5	4
calibration	5	2
cams	5	2
castings	5	3
centralization	5	2
centred	5	2
chaotic	5	3
chlorine	5	2
chrome	5	4
chromium	5	3
commonality	5	4
configured	5	3
consciously	5	5
console	5	4
container	5	3
contexts	5	3
continuum	5	3
controversy	5	4
converse	5	5
cutoff	5	3
cutters	5	3
decrement	5	4
deduced	5	4
definitive	5	2
deflections	5	2
dependable	5	3
derivation	5	3
designate	5	2
destructive	5	5
detection	5	4
detrimental	5	4
diagnose	5	2
diagonal	5	4
diligence	5	3
directional	5	3
directory	5	4
discard	5	5
displaced	5	4
economist	5	3
edit	5	4
eds	5	3
encoded	5	2
encompass	5	4
equates	5	3
ethical	5	2
ethnic	5	2
excellence	5	3
excluding	5	5
exemplified	5	4
expedite	5	2
extrapolation	5	3
facilitated	5	5
familiarity	5	5
fashionable	5	2
fittings	5	2
fixture	5	3
formally	5	4
formats	5	2
foundry	5	3
generalized	5	3
genetic	5	3
geographically	5	3
germanium	5	2
glossary	5	3
hopefully	5	4
hypothetical	5	4
inaccuracies	5	5
inclination	5	3
inclusive	5	2
inconvenient	5	4
influential	5	4
infrared	5	2
infrastructure	5	4
initiative	5	3
innermost	5	3
inputting	5	4
inspections	5	3
installations	5	3
instrumental	5	4
insulator	5	3
intact	5	4
intangible	5	3
intentionally	5	4
invariably	5	4
inverse	5	3
lateral	5	4
liable	5	4
linearity	5	2
lowercase	5	5
magnetised	5	2
manipulator	5	2
marital	5	5
markings	5	3
maze	5	2
measurable	5	3
mechanization	5	3
membership	5	3
merchants	5	3
millionths	5	2
misconception	5	5
mismatch	5	2
mould	5	2
negligible	5	4
nicely	5	4
nineteenth	5	5
normative	5	2
optimised	5	5
orbit	5	2
organic	5	3
overlapping	5	5
pancake	5	2
paperwork	5	3
participants	5	3
percentages	5	3
pinion	5	5
portability	5	4
positively	5	3
practicable	5	4
predecessors	5	4
predominant	5	4
preferable	5	5
preferably	5	5
prevention	5	3
principally	5	3
profitably	5	4
progression	5	3
propagates	5	3
psychologists	5	2
queries	5	2
ratings	5	3
reconstructed	5	2
redefine	5	4
redefining	5	4
repetitively	5	3
resistant	5	2
resolvable	5	2
resonance	5	3
responsiveness	5	2
restrictive	5	5
rewritten	5	2
rigidly	5	4
rivet	5	5

rocket	5	3
rollers	5	3
roulette	5	2
sausage	5	2
secondly	5	4
seemingly	5	4
selfconcept	5	2
semicolon	5	5
severely	5	5
shrinkage	5	3
skeletal	5	5
spherical	5	3
sponsored	5	5
stainless	5	4
steelcase	5	2
subsystems	5	3
subtle	5	4
superseded	5	3
synchronous	5	2
tag	5	2
targeted	5	3
tolerances	5	3
tolerated	5	5
transit	5	3
tricky	5	2
turnaround	5	2
tyres	5	2
unaffected	5	3
unbiased	5	4
unlimited	5	5
unnecessarily	5	4
unrealistic	5	4
unreliability	5	5
unsatisfactory	5	4
unskilled	5	2
unsuccessful	5	4
upwards	5	3
utilised	5	3
vacuum	5	3
verification	5	5
verified	5	3
versatility	5	3
viability	5	3
visualize	5	3
volts	5	2
whiskers	5	2
workable	5	4
workstations	5	3
worthwhile	5	5
xerox	5	5
zero	5	5

Appendix Four

Invented words for Study One and Two

Section One

CORE

1. 'Applied Statistical Methods'
   illumbent, oplimation, bractile, apperstrictive, disclusivity

2. 'Financial Analysis and Control Systems'
   diagmatic, versive, uncouncible, megocentred, peg drive management

3. 'Computer-aided Design and Computer-aided Manufacturing'
   marlobetics, distentive, stross control, rale programme, computer underfly

4. 'Quality Reliability Maintenance Systems'
   exertibility, deliction, ecostition rate, cartol failure, system obligance

5. 'Basic Computing'
   simblet, parmertype, bullage, wanderfix unit, dialitical digit

IT

6. 'Automation and Robotics'
   bactivator, crappor, casp, articulated subvincible system, peliographic robot

7. 'Computer Engineering'
   shellnet, trinode, arno-loop, genucology, interplexing

8. 'Computer Language'
   transcent, reduct, rebation, semantic defractics, quantiple separator

9. 'Artificial Intelligence'
   hartflow, numantics, spandex, obtex, rallifier

10. 'Programming Language'
    emorate, gradics, pipe sigmentals, venerous definition, FORTH claw

EBM

11. 'Manufacturing Strategy'
    sub-vincibility, plexi-control, resparated, demobility switching,
    quality degagement

12. 'Human Factors in Industry'
    observement, sensiblication, classitude, martricular depense,
    development potentiation

13. 'Information Systems Strategy'
    tribation, enventilement, over-angle systems, addation model, blore pointer

14. 'Strategic Marketing'
    endo-marketing, byesale, emorcument, enhalement, market displement

MSE

15. 'Metallic Materials Selection'
    aurelite, antite, lucanite, vexite, thermocrinal

16. 'Logistics and Supply Chain Management'
    intellation, decupulate, reploy, maintream underthrow, inventory deprobe

17. 'Polymer Materials, Processes and Products'
    bi-potic, pharicise, parathane, polythane, polyphathane

18. 'Industrial Engineering'
    fallity, hesticate, festivation, cumulity target, tubertile system

19. 'Machine Tool'
    torce, entrave, sedulant, incudency, circumhinge

IT/EBM

20. 'Information Technology Fundamentals'
    infendor, hastillation, ludomechanical, electropod circuit, pressor integration

21. 'Production Planning and Control Systems'
    exation, randomation, wardenment, reclearment control, disponent service

IT/MSE

22. 'Programmable Systems in Measurement and Control'
    hyperbanner, galvant, rebect, single cranning, pake space

SELECTIVE

23. 'Simulation of Production Systems'
    rectication, simulactment, linehanding, notion platern, lentalization control

24. 'Project Planning Management and Control'
    holdmarker, wedgescape, hobstick, systile management, relancement technology

25. 'Financial Decision Making'
discountment, everisation, flow masterance, loss juration, abtile productivity

26. 'Introduction to Manufacturing Systems'
masket, hadding, nome, sorband, redget

Section Two

subvential, integratory, helochronic, obolic, dispect, magnostify, stillic, mortence, extracute, plausive, insculation, rowball, nellistic, tonid, actonic, bioreferant, ultivistic, relovation, bifector, tresist, legidity, marmented, rostical, invexicate, shabidity, pedget, condulation, despodation, windstore, routance, intertaken, flossid, anavistic, intortment, barhood, disclusivity, decrevity, vectalise, flectative, scopical

Appendix Five

Letter to Subject Lecturer

Dear ______

Thank you very much for your valuable comments on the results of an experimental study at the beginning of this year.

I am writing to you again for your help to identify the technical terms in the module(s) you have been teaching on the Msc. Engineering programme.

As mentioned in the first letter, I am interested in researching into the vocabulary needs of students from Southeast Asia studying the Msc. Engineering programme. The research is based on a corpus of the recommended textbooks, which cover the twenty six modules.

Technical terms were identified for each module mainly on the basis of frequency and range distribution. Distribution here means range of occurrences on the texts selected. If a word occurs with high frequency and low distribution, it is identified as a technical term in this study. By high frequency, I decided to count as a technical term if this word has five or more occurrences in one text. By distribution, the lowest distribution is one, which means that a word occurs with high frequency only in one text. There are variations in this case.

Apart from what is called single word terms (for example, *hardware, software*), there are multi-word terms (for example, *polyvinyal chloride*) too.

The attached is a list of technical terms from the module 'Logistics and Supply Chain Management'. Please tick ten technical terms from the list that you think are typical of, and best represent the module. Thank you for your help.

Technical terms from 'Logistics and Supply Chain Management':

		Frequency	Range
bottom-up replanning		5	1
bracket assembly		10	1
capacity planning	√	24	5
component part	√	59	6
data base		75	6
exception code		7	1
fabrication		30	5
finished goods	√	29	7
flow of materials		8	3
gross requirement		67	1
bill of materials (BOM)		10	1
inventory control		15	4
item master file		6	1
lead time offset		7	1
lot sizing		7	1
manufacturing planning and control (MPC)		159	1
manufacturing process	√	41	7
master production schedule (MPC)	√	19	2
material flow		14	2
material requirements planning (MRP)	√	183	4
MPC system	√	92	1
net requirement		19	1
order launching		6	1
pegging		18	1
planning horizon		9	4
planned order		70	1
planned order release		30	1
project team		15	4
projected available balance	√	16	1
raw material		64	9
regeneration		9	1
safety lead time	√	13	1
scheduled receipt		55	1
shop floor control	√	9	2
shop order		25	1
time bucket		9	1
top handle assembly		20	1

Appendix Six

Technical Words Not Recognized in Study One and Two
(words not recognized in both studies asterisked)

CORE

1. 'Applied Statistical Methods'
   class interval, cumulative frequency, descriptive measures, experimentation generality, inference*, informative, quartile, relative frequency histogram sample, sample interquartile range, statistical population statistical method*, uncertainty

2. 'Financial Analysis and Control Systems'
   accounting, corporation, cost accounting, cost management, factory ledger, information technology, internal accounting system, management accounting*, management accounting system* process control*, return on investment, stockturn, transaction*

3. 'Computer-aided Design and Computer-aided Manufacturing'
   automatically programmed tool, cathode ray tube, centralised network, computer graphics*, database*, design process*, draughting*, geometric*, intuitive, mass storage device, process planning*

4. 'Quality Reliability Maintenance Systems'
   binomial distribution, expected value, failure rate*, hazard rate*, life cycle cost, maintainability*, mean time between failures*, non-repairable, normal distribution, probability, rate of occurrence of failures, reliability*, repairable system, standard deviation

5. 'Basic Computing'
   analytical engine, anatomy, central processing unit, encode*, FORTH, location address, logic gates

IT

6. 'Automation and Robotics'
   actuator*, analog-digital converter, articulated mechanical system, degrees of freedom*, end effector*, gripper, incremental encoder, piston shaft, potentiometer, resolver, rotation, servocontrol, vector*

7. 'Computer Engineering'
   analog transmission, arpanet, baseband, broadband, coaxial cable, decnet, header, integrated services digital network, modem, multiplexing, packet switching, subnet, topology*

8. 'Computer Language'
algol, assignment statement*, compound statement, constant definition, field identifier, function declaration, identified variable, integer value, iteration, pointer value, record variable, separator, subrange type*, type definition, unsigned integer, variable declaration

9. 'Artificial Intelligence'
backtracking, cardinal, duplex, functor, iteration, LISP, notation*, parse, PROLOG system, recursion, semantics, simplifier*, working memory

10. 'Programming Language'
allot, colon definition*, delimiter*, FORTH arithmetic, FORTH variable, FORTH word, infix, normal stack, quotient*, stack manipulation

EBM

11. 'Manufacturing Strategy'
direct labour*, flow industries, flow shop*, job shop*, lead time*, manufacturing lead time, return on investment, statistical process control*, total quality control, total preventative maintenance

12. 'Human Factors in Industry'
apprenticeship, aptitude*, assembly line, autonomous, behavioural science, management development, organizational behaviour*, psychology*, recruitment*, scientific management*, trait

13. 'Information Systems Strategy'
data processing, information system, IS/IT planning, leverage, management information system, management services, operational system*, portfolio, portfolio model*, retailer, strategic framework, strategic information system*, strategic planning*, telemarketing, three-era model

14. 'Strategic Marketing'
economic system, effective marketing, elasticity, equilibrium price, form utility, macro-marketing, market-directed economy, marketing mix, mass marketing, micro-marketing, mass production, middleman, possession utility, strategy planning, target customers, target marketing, total revenues

MSE

15. 'Metallic Materials Selection'
bainite, carbide, cementite, ferrite*, isothermal, lamellae, martensite*, nuclei, pearlite, precipitation, ultimate tensile strength

16. 'Logistics and Supply Chain Management'
capacity planning*, component part, inventory level, master production schedule*, MPC system, projected available balance, replenishment, safety lead time, shop floor control*, turnover, vendor

17. 'Polymer Materials, Processes and Products'
fatigue, moulding, polyamide*, polycarbonate*, polyester, polyethylene, polypropylene, polystyrene, polyurethane

18. 'Industrial Engineering'
constraint, necessitate, output stock, physical resource, remuneration, resource utilization, resource productivity, service system, supply system, transport and service system, utilization*

19. 'Machine Tool'
built-up edge, orthogonal, plastic deformation*, rake*, rake angle, spindle, tangential, tool face, tool wear, velocity, wedge

IT/EBM

20. 'Information Technology Fundamentals'
artificial intelligence, broadcast, data processing, dp application, electromechanical*, end-user, information exchange, instruction set, mainframe, multiprogramming, object code, source code, tabulator, telecommunication*

21. 'Production Planning and Control Systems'
break-even analysis, internal rate of return, nonmanufacturing, operations management, production control*, production operation*, production system*, quality control*, sampling distribution, service system, subsystem

IT/MSE

22. 'Programmable Systems in Measurement and Control'
address bus*, address space, boolean algebra, hexadecimal*, latch, NAND gate, octal, silicon*, single chip

SELECTIVE

23. 'Simulation of Production Systems'
coating, equation, expertise, machine shop, queue, uniform random number

24. 'Project Planning Management and Control'
barchart*, critical path analysis, free float, monitor, precedence, precedence convention, rally, time scale*, total float, yardstick

25. 'Financial Decision Making'
accounting profit, capital rationing, decision objective, dividend, dividend flow*, financial decisions*, investment, investment appraisal, liquidity requirements, maximization, return on capital employed, time dimension, time value of money, valuation, value base

26. 'Introduction to Manufacturing Systems'
cutting tool, dislocation, drilling, endurance limit, flow shop, job shop, lathe, level of automation, manufacturing cost, materials handling, mold, process technology, tooling

Appendix Seven

Technical Words Recognized in Study One and Two
(words recognized in both studies asterisked)

CORE

1. 'Applied Statistical Methods'
   frequency distribution, population, statistics

2. 'Financial Analysis and Control Systems'
   direct labour, information technology, raw material

3. 'Computer-aided Design and Computer-aided Manufacturing'
   numerical control*, robotic

4. 'Quality Reliability Maintenance Systems'
   non-repairable item

5. 'Basic Computing'
   arithmetic and logic unit, arithmetical, binary, binary digit, control unit, input device, interface*, output device, random access memory, read only memory, storage

IT

6. 'Automation and Robotics'
   human operator, industrial robot, sensor

7. 'Computer Engineering'
   computer network, data base, interface, loop, node, protocol

8. 'Computer Language'
   Boolean, identifier

9. 'Artificial Intelligence'
   artificial intelligence, congestion, expert system, PROLOG

10. 'Programming Language'
    array, integer, operand*, stack, variable

EBM

11. 'Manufacturing Strategy'
    customer service, just-in-time*, mass production, world-class manufacturing

12. 'Human Factors in Industry'
    behaviour, human-resource*, human-resource planning

13. 'Information Systems Strategy'
business planning

14. 'Strategic Marketing'
marketing, marketing strategy

MSE

15. 'Metallic Materials Selection'
austenite*, ductile, equilibrium, grain size, plastic deformation, yield stress

16. 'Logistics and Supply Chain Management'
finished goods, manufacturing process, material flow, materials requirement planning*, warehouse

17. 'Polymer Materials, Processes and Products'
creep, ductility, modulus, molecule, polymer*, stiffness, thermoplastics toughness

18. 'Industrial Engineering'
capacity management*, feasibility, operations management, operations management decision-making, system structure

19. 'Machine Tool'
clearance, fracture, lubricant, shear*, tungsten, tungsten carbide

IT/EBM

20. 'Information Technology Fundamentals'
assembler, central processing unit, compiler, magnetic tape

21. 'Production Planning and Control Systems'
apics, break-even point, inventory control*, mass production

IT/MSE

22. 'Programmable Systems in Measurement and Control'
assembly language, byte, data type, high level language, integrated circuit, microcomputer, microprocessor, storage

SELECTIVE

23. 'Simulation of Production Systems'
activity cycle diagram, discrete change, discrete simulation, entity, flexible manufacturing system, job shop, manufacturing system, maximum stock level, parameter value, probability distribution, simulation, simulator, temporary entity, three-phase rule

24. 'Project Planning Management and Control'
breakdown, information technology, network analysis,
project management*, project manager, project planning*

25. 'Financial Decision Making'
payback, working capital

26. 'Introduction to Manufacturing Systems'
numerical control, elastic limit, fracture toughness, grain boundary,
hardness, strain, tensile strength