

# Vocabulary and the Teaching of English for Specific Purposes

Faculty of Social Information Science, Kure University

Simon A. Fraser

---

**Abstract** This paper is a brief survey of the pedagogic issues and literature in the field of vocabulary in English for Specific Purposes (ESP). There is agreement among ESP practitioners that the teaching of vocabulary needs to be improved if second language learners are to cope adequately with academic writing in a specialist discipline such as medicine. The problem still remains, however, as to the nature of the vocabulary that should be taught. The paper reviews some of the work which has been done by researchers who have attempted to answer this question by categorizing and quantifying the vocabulary of scientific texts.

**Key Words:** English for specific purposes (ESP), Technical vocabulary, Subtechnical vocabulary

## 1 INTRODUCTION

Vocabulary has been described by Meara<sup>1)</sup> as “a neglected aspect of language learning”. In the past, there has been a tendency to ignore vocabulary in favour of structure, and many share the view that vocabulary teaching remains unsatisfactory. Part of the blame for the current situation has been placed on more modern communicative approach methodologies introduced in the 1970s which shift the focus away from vocabulary (see, for instance, Harvey<sup>2)</sup> and McCarthy<sup>3)</sup>). In the field of English for Specific Purposes, the neglect of vocabulary has been no less apparent. Recently, however, there have been signs that researchers are realizing that the EFL (English as a foreign language) learner needs to have an excellent command of vocabulary. Johnson<sup>4)</sup> believes the 80s to have been the period in which the importance of vocabulary instruction was re-discovered. Researchers, materials designers

and ESP practitioners have become more interested than ever before in questions related to lexis. Ulijn<sup>5)</sup>, for example, realizes that the most serious difficulties faced by foreign learners in the ESP field are lexical, particularly at intermediate and advanced levels of reading. Ulijn believes that syntactic analysis is superficial except when insufficient cues are provided by the lexicon and the conceptual system to comprehend the text. Laufer and Sim<sup>6)</sup> go further and state that foreign readers sometimes distort the syntax of the text so that it can fit their interpretation of words. According to Laufer<sup>7)</sup>, massive vocabulary expansion should be one of the major goals in any course in Language for Academic Purposes. Laufer has evidence to suggest that 5,000 words seems to be the lexical threshold beneath which other facilitating factors in reading comprehension may not be very effective, and he believes that many students beginning their academic studies have poorer vocabularies than this.

## 2 THE VOCABULARY OF SCIENTIFIC TEXTS

There is agreement, then, that vocabulary building has not received the attention it requires if second language learners are to be efficient readers. The question remains, though, as to exactly what kind of vocabulary should be taught.

### 2.1 Technical Vocabulary

It is commonly thought in ESP that a knowledge of scientific terminology is sufficient for non-native readers. Martin<sup>8)</sup> defines this technical vocabulary as “the specific vocabulary related to a particular discipline”, and she gives examples of some specialized EFL books which have focused on this. A more precise definition of the specialized vocabulary in medicine is given by Salager<sup>9)</sup>, who states that it “comprises those high-frequency, context-bound, or topic-dependent, terms particular to a given medical specialty”. Walsh<sup>10)</sup> points out that specialist vocabulary is an obvious focal point in any examination of the vocabulary of scientific texts. Often only the teachers, practitioners or students of a subject will be familiar with this store of words used regularly in a particular subject area, which are most often found in science textbooks or research papers. Medical science is particularly well-known for its use of words which are incomprehensible to the layman.

Because these words are used almost exclusively in scientific texts there is a temptation to believe that a mastery of technical vocabulary is all that is required for success in ESP reading. However, the consensus of opinion among ESP practitioners appears to be that it is not the job of the English teacher to teach technical vocabulary. Cowan<sup>11)</sup>, in fact, believes that learning the meanings of these words is an automatic consequence of studying the discipline which uses them. Investigations into the reading problems of non-natives demonstrate that teachers should be concentrating on

what is often termed “subtechnical” vocabulary.

### 2.2 Subtechnical Vocabulary

Trimble<sup>12)</sup> claims that subtechnical vocabulary is important in the teaching of ESP because items are common across fields, and because this area, although often overlooked by teachers, is a cause of problems for students. The term “subtechnical” for categorizing vocabulary was introduced in 1974 by Cowan<sup>11)</sup>, who defines it as “context-independent words which occur with high frequency across disciplines”. Martin<sup>8)</sup>, while agreeing with this definition, prefers the term “academic vocabulary”. Trimble<sup>12)</sup> extends Cowan’s definition to include “those common words which occur with special meanings in specific scientific and technical fields”. He gives the example *fast (arsenic-fast virus)* in the sense of “resistant to”. Howard<sup>13)</sup> prefers the term “crypto-technical” for this group of polysemous words because their technical sense is hidden and can cause confusion to the learner.

Trimble believes that students should come to recognize that there are familiar words that have technical senses. This indeed will help to reassure teachers that scientific texts are not simply unintelligible, but that learners’ underlying prior knowledge can be activated to enable them to relate new vocabulary to existing knowledge structures.

#### 2.2.1 Frequency Counts and Statistical Analyses

A number of workers have attempted to quantify the vocabulary of scientific texts using computer-aided frequency counts and statistical analyses.

Inman<sup>14)</sup>, who carried out a computerized study of a 114,000 word corpus from professional journals, divided into 10 scientific and technological fields, follows Cowan’s definition. Inman classified words with low distribution across the fields as “technical” (21% of the corpus), and after identifying “function” words (9%), considered the remainder (70%) to be subtechnical words.

A list of the most frequent of these subtechnical words included *result*, *process*, *system* and *function*. The problem with this categorization is that the 70% of “subtechnical” words comprises everything in the texts that is not a technical or functional word. Frequency counts on corpuses have shown that whatever the size of the corpus, very few items occur with high frequency. The majority of items are *hapax legomena* –very low frequency words– and up to 50% of the corpus consists of items which occur only once.

Salager<sup>15)</sup> attempted to determine the core lexis of medicine, or what she calls “fundamental medical English”, by carrying out a statistical study of a corpus of 100,000 words. This corpus comprised 50 sample texts (drawn from 50 different textbooks, journals and reviews) of approximately 2,000 words each. The analysis took the “root” of a word as being the basic unit. Salager describes a statistical procedure which enables her to split the medical English (ME) lexicon into three superimposed strata: (1) Basic English (BE) which is common to all kinds of writing; (2) Fundamental Medical English (FME) which is made up of those roots whose frequency of occurrence is homogeneously distributed on the whole ME corpus but in a significantly different way from that in BE; and (3) Specialized Medical English (SME) made up of specialized terms particular to a given discipline and known only to a small group of workers in the field. The results of Salager’s investigation into the nature of FME vocabulary showed the following:

- (a) The verbal “roots” are composed of items which describe the general methodology of scientific inquiry (e.g. description, analysis, comparison) and those which indicate the evolution of diseases and the patient’s clinical state.
- (b) The noun “roots” describe the medical procedure, and express measurement.
- (c) The adjectival roots describe illness or injury and the quality and/or timing of treatment.
- (d) The “function” words mainly express causality, opposition and purpose.

Yang<sup>16)</sup> describes a computerized method of identifying scientific/ technical terms using a corpus of 300,000 words. He found that words with high distribution and high frequency tended to be function words, and if they showed a peak of frequency in one field, they would tend to be a technical term. The remaining words (lower frequency with high distribution) were identified by Yang as subtechnical – words such as *factor*, *problem*, *conclusion* and *disease*. Yang also found that there could be overlap between subtechnical words and technical terms. He found *solution*, for example, to be a subtechnical word because of its high distribution. However, *solution* is also found with peak frequencies of occurrence in the fields of chemistry and mathematics.

Perhaps the most thorough investigation into the notion of subtechnical vocabulary has been carried out by Baker<sup>17)</sup>. She discusses the various definitions of this category of vocabulary and brings them all together in the following list:

- (1) Items which are used to express notions general to specialized disciplines, e.g. *factor*, *method* and *function*.
- (2) Items which, in addition to their meaning in general language, have a specialized meaning in one or more disciplines, e.g. *bug* in computer science. These correspond to Howard’s “cryptotechnical” words.
- (3) Items which are not used in general language but which have different meanings in specialized disciplines. *Morphological*, for example, has different meanings in linguistics and botany.
- (4) General language items which have restricted meanings in certain specialized disciplines. For example in botany, *effective* simply means *take effect*.
- (5) General language vocabulary items which are used frequently to describe technical

processes and functions.

- (6) Items which are used in specialized texts to perform specific rhetorical functions. These signal the writer's intention or his evaluation of the material presented.

In Baker's view, the last type of "sub-technical" item is the most important and also the most difficult to acquire. She believes that a knowledge of rhetorical/ organizational lexis is important if learners are to develop a sense of style relevant to the various text types used in their specific disciplines and to be able to interpret and evaluate an argument or piece of specialist information. However, this category of words would appear to be totally incompatible with the set of words defined by Cowan.

In her study of 18 medical articles, Baker attempted to identify rhetorical/ organizational lexis. In order to narrow down the number of possible candidates for the category of items which may be used for rhetorical reasons, Baker decided first to eliminate specialized and general lexis. General lexis was identified on the basis of its even distribution across medical and general English. Specialized lexis was identified on the basis of (a) its uneven distribution among medical texts (text specificity) and (b) a significant difference between its frequency of occurrence in medical and general English. If an item was not identified as either Specialized or General, it was considered to be a rhetorical/ organizational item – an item considered to have a potentially significant role in structuring the writer's argument. Baker arrived at a list of 65 items, of which five (*findings, report, reported, diagnosis and evidence*) were selected for a close analysis using KWIC (Key Word in Context) concordances. The computer was able to show whether a particular pattern was typical or atypical of the genre, and to give an indication of whether an item was used in the same way in all sections, or used for specific rhetorical purposes in specific rhetorical sections. The Discussion Section, unsurprisingly, was found to contain the largest

number of rhetorical items.

### 2.2.2 Discourse-Organizing Words

Baker's category of subtechnical/ rhetorical lexis appears to have much in common with what McCarthy<sup>18)</sup> calls "discourse-organizing words". According to McCarthy it is the job of words such as *problem, issue* and *assessment* to organize and structure the argument of a text, rather than answer for its content or field. Some of these words in fact indicate the larger text-patterns that the author has chosen, and allow the reader to predict the shape of the whole discourse. Such words, then, are clearly very important, as the language learner who is not familiar with them may find it difficult to decode the whole text efficiently. As McCarthy points out, the inability to understand discourse-organizing words or misinterpretation of them could cause problems.

Some linguists have attempted to define more precisely sub-sets of words in this lexical area. Winter<sup>19)</sup>, for instance, identifies "Vocabulary 3" words, which act as "lexical items" of connection. This list includes words like the following:

achieve, addition, basis, case, change, differ, explanation, fact, function, hypothetical, identify, kind, manner, point, result, technique, way.

These items closely paraphrase basic clause-linking relationships, such as reason in a context like "He left for the reason that he did not like it there" where *for the reason that* could be replaced by *because*.

Subtechnical/ rhetorical lexis is also linked to the category of "anaphoric nouns" put forward by Francis<sup>20)</sup>. Francis gives extensive examples of nouns that frequently occur to refer back to chunks of text. These nouns operate as pro-forms and are presented as the given element in a clause containing new information. Words such as *hypothesis, concept, position, diagnosis, interpretation* and *viewpoint* fulfil these conditions. For example, consider the following sentence:

“Here I want to spend some time examining this issue”.

In the example, “this issue” is acting anaphorically —referring back to the preceding text. It also helps us to understand the writer’s attitude towards the content — we know to look for something problematic.

### 3 CONCLUSION

The neglect suffered by vocabulary in ESP in the past is all the more surprising since the most serious difficulty identified by the language learner is likely to be “I don’t know enough words!”. However, we have seen that researchers now acknowledge that learners in the ESP field need to have an excellent command of vocabulary.

The category of vocabulary known as “subtechnical” seems to be of particular importance, with this term being used to cover a wide range of items which are neither highly technical and specific to a particular field of knowledge nor obviously general. Although this category of words is a source of confusion for many teachers, it has been useful in drawing attention to the fact that it is not only technical terms which pose problems for learners. Students need a command of both technical and subtechnical vocabulary.

Swales<sup>21)</sup> has argued that ESP cannot “come of age” until more efficient ways have been found of coping with the vocabulary problem. Progress has most definitely been made, but a great deal of research is still required on the part of ESP practitioners.

### References

- 1) P. M. Meara. 1980. Vocabulary acquisition: a neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts* 13: 221-46.
- 2) P. D. Harvey. 1983. Vocabulary learning: the use of grids. *English Language Teaching Journal* 37 (3): 243-250.
- 3) M. J. McCarthy. 1984. A new look at vocabulary in EFL. *Applied Linguistics* 5 (1): 12-22.
- 4) D. D. Johnson. 1986. Journal of reading: a themed issue on vocabulary instruction. *Journal of Reading* 29 (7): 5-6.
- 5) J. M. Ulijn. 1984. Reading for professional purposes: Psycholinguistic evidence in a cross-linguistic perspective. In J. M. Ulijn and A. K. Pugh (eds). *Reading for Professional Purposes: Studies and Practices in Native and Foreign Languages*. Heinemann.
- 6) B. Laufer and D. D. Sim. 1985. Taking the easy way out: non-use and misuse of clues in EFL reading. *English Teaching Forum* 23 (2).
- 7) B. Laufer. 1989. What percentage of text-lexis is necessary for comprehension? In Lauren and Nordman (eds). *Special Language: From Humans Thinking to Thinking Machines*. Multilingual Matters.
- 8) A. V. Martin. 1976. Teaching academic vocabulary to foreign graduate students. *TESOL Quarterly* 10 (1): 91-97.
- 9) F. Salager. 1985. Specialist medical English lexis: classificatory framework and rhetorical function — a statistical approach. *EMP Newsletter* 2 (2): 5-17.
- 10) V. Walsh. 1982. Reading scientific texts in English. *English Teaching Forum* (3): 24-28.
- 11) J. R. Cowan. 1974. Lexical and syntactic research for the design of reading material. *TESOL Quarterly* 8 (4): 389-399.
- 12) L. Trimble. 1985. *English for Science and Technology*. Cambridge University Press.
- 13) R. Howard. 1991. Teaching medical English vocabulary systematically. *EMP Newsletter* 8: 15-21.
- 14) M. Inman. 1978. Lexical analysis of scientific and technical prose. In Trimble et al. (eds). *English*

- for Specific Purposes: Science and Technology*. English Language Institute, Oregon State University.
- 15) F. Salager. 1983. The lexis of fundamental English: classificatory framework and rhetorical function (a statistical approach). *Reading in a Foreign Language* 1 (1): 54-64.
  - 16) H. Yang. 1986. A new technique for identifying scientific/technical terms and describing science text (an interim report). *Literary and Linguistic Computing* 1 (2): 93-103.
  - 17) M. Baker. 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): 91-105.
  - 18) M. J. McCarthy. 1991. *Discourse Analysis for Teachers*. Cambridge University Press.
  - 19) E. O. Winter. 1978. A look at the role of certain words in information structure. In K. P. Jones and V. Horsnell (eds). *Informatics* 3: 1: 85-87. ASLIB.
  - 20) G. Francis. 1986. *Anaphoric Nouns*. ELR, University of Birmingham.
  - 21) J. Swales. 1985. *Episodes in ESP*. Prentice Hall International.