A Corpus Analysis of English Vocabulary Input in Course Materials Used for Engineering Students

Siwalaluk Chanchanglek [1]
Nongnuch Sriussadaporn [2]

ABSTRACT

The purpose of the study was to analyze vocabulary input in English for Engineering courses in three areas: 1) proportion of general service vocabularies based on the general service list (GSL) 2) proportion of academic words based on the academic word list (AWL) and 3) most frequently used academic words in a corpus. The textbooks were collected from universities running English for Engineering courses: 1) Thammasart University (TU) 2) Rangsit University (RSU) 3) Rajamangala University of Technology (RMUT). The lexical analysis software applied in the study was WordSmith. The finding revealed that the majority (around 80%) of every text comprised of words from the GSL. Regarding academic words, they accounted for 5-6 % in each text approximately. Finally, the study found that 363 academic words were used in English for Engineering textbooks.

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อวิเคราะห์คำศัพท์ที่อยู่ในหนังสือเรียนภาษาอังกฤษจากมหาวิทยาลัย ที่มีการเรียนการสอนวิชาภาษาอังกฤษเพื่อนักศึกษาคณะวิศวกรรมศาสตร์ ซึ่งกลุ่มตัวอย่างประกอบด้วย มหาวิทยาลัยธรรมศาสตร์ มหาวิทยาลัยรังสิต และมหาวิทยาลัยเทคโนโลยีราชมงคล โดยวิเคราะห์ใน 3 ด้านต่อไปนี้ 1) อัตราส่วนของคำศัพท์พื้นฐานที่ใช้ทั่วไป หรือ general service list 2) อัตราส่วนของ คำศัพท์ทางวิชาการ หรือ academic word lists และ 3) นำเสนอคำศัพท์ ทางวิชาการที่พบในการศึกษา ครั้งนี้เพื่อที่จะสร้างประมวลคำศัพท์ทางวิชาการที่ใช้ในคณะวิศวกรรมศาสตร์ และเครื่องมือที่ใช้ในการ วิเคราะห์คำศัพท์คือโปรแกรม WordSmith จากการศึกษาพบว่า หนังสือแต่ละเล่มมีอัตราส่วนของคำศัพท์พื้น ฐานทั่วไปโดยเฉลี่ยร้อยละ 80 และมีอัตราส่วนของคำศัพท์ทางวิชาการร้อยละ 5-6 นอกจากนี้จากการศึกษา พบว่ามีคำศัพท์ทางวิชาการที่ใช้ในหนังสือเรียนภาษาอังกฤษเพื่อวิศวกรรมศาสตร์ทั้งหมดจำนวน 363 คำ

[1] *Graduate Student, English for Specific Purposes Program, Graduate School, Kasetsart University*

[2] *Assistant Professor, Department of Foreign Languages, Faculty of Humanities, Kasetsart University*

Background of the Study

It has been a staple of Thai education policy for at least 30 years that undergraduates in science and Engineering should study their specialist subject matter by reading textbooks in English (Prapphal, 2001). Engineering university students have to use English as a basis for exploring new knowledge and information in their own fields of study. However, over the past 30 years or so it has become apparent that incoming university Engineering undergraduates in many developing nations fall far short of the basis linguistic and in particular lexical knowledge necessary to read materials in English (Nurweni and Read, 1999). Even though reading skills are emphasized and Thais have all studied for at least six years at secondary school and indeed probably before that, university students have still confronted the difficulties of lexis and reading. This caused many interesting questions related to teaching and learning English in Thailand. For example, "how many words of vocabulary should student know and learn?" or "what strategies do learners use to acquire vocabulary?"

Another consideration focuses on vocabulary input. That is, how do teachers or course designers decide which and how many vocabularies should be incorporated? The most common paradigm for this lexical problem comes from the work of Paul Nation and his associates (1995). The chief approaches are that learners should learn first the 2000 or General Service Word List (GSL) by West (1953) as well as a set of academic words common to all academic disciplines. In practice this is taken to mean the 2000-word family and the 570-word family Academic Word List or AWL (Coxhead, 2000). The course designers and English teacher, thus, have to consider both commonly words used and technical terms when designing teaching material for their students.

This leads to the interest in the study of corpus-based analysis that has been developed as another resource for developing English proficiency. According to Hunston (2002), corpus is a representative of how texts can be collected, contained and accessed by specific software program so that concordance lines are presented and frequencies are calculated.

Besides, a number of advantages in corpus-based approach for developing teaching materials were found. Thurstun and Candlin (1998) mention that corpus-based study has a potential role to play in expanding the target vocabulary of the learners. That is, the outputs gained from corpus based study resulting in practical input of materials used in classroom course designers and teachers can get information from authentic sources. Flowerdew (1997) explains that the value of the corpus- based word list lies in the fact that course designers can be certain that the selected words are the most relevant and useful for learners. It also provides criteria for teachers to choose the materials or textbooks, syllabus selection and grading of lexical items; therefore, it is a very practical and effective tool.

Consequently, investigating the vocabulary input of the course material in an English course, regarding frequency in particular, is worthwhile to see what vocabulary is actually taught to students and whether the materials are suitable for a specific group of learners. Such research can contribute to pedagogy for non-native English speakers. The samples of the study, moreover, can be employed as the guideline for designing a course.

Objectives:

1) To compare vocabulary appearing in intensive English for Engineering courses with the General Service List (GSL) (Bauman and Culligan, 1995) so as to find out proportion of vocabulary input based on GSL.

2) To compare vocabulary appearing in intensive English for Engineering courses with the Academic Word List (AWL) (Coxhead, 1998) so as to find out proportion of vocabulary input based on AWL.

3) To find out the most frequently used academic vocabulary items in a corpus of Engineering texts.

Previous Research: Flowerdew (2003) conducted a study of the similarities and differences between professional and nonprofessional writing in the pattern of the problem-solution. The research employed the WordSmith Tool to explore key words that automatically provide linguistic evidence for the problem-solution pattern.

Kaewphanngam, Broughton, and Soranasataporn (2002) studied the specialized vocabulary used in psychology texts to meet the need of non-native graduated students majoring in Psychology at Silpakorn University. The corpus included three sub-fields of psychology which are Counseling Psychology, Education Psychology and Social Psychology respectively. The comparison between the content words in a general corpus and the specialized corpus were conducted in the study. These findings suggested the need to produce field-specific word lists which incorporate all frequent specific lexical items necessary for the expression of the rhetoric of the specific subject areas.

Materials and Methods

Materials: The materials used in this study were printed course materials or textbooks used in the course of "English for Engineering" The course materials were taken from the universities which had the Engineering program in undergraduate level and run "English for Engineering" course. They used that course materials for their Engineering students in 2009-2010 academic year. So, textbooks from Thammasart University (TU), Rangsit University (RSU) and Rajamangkala University of Technology (RMUT) were selected. Moreover, in-house materials were used in TU while commercial course books were used in RSU and RMUT.

Methods: Data Collection and Procedure: 1) Searching universities operating Engineering faculty 2) Finding out lists of universities offering English for Engineering courses. 3) Collecting textbooks and materials used in the course of English for Engineering.

Data Analysis: Firstly, all textbooks were scanned and divided into sub corpus which were TU (textbook used in the first academic year of Thammasart University), TU1 (textbook used in the second academic year of Thammasart University), RSU (textbook used in the first academic year of Rangsit University), RSU1 (textbook used in the second academic year of Rangsit University) and RMUT (textbook of Rajamangala University of Technology). Concordancer was employed to analyze word frequency and categorize word group (GSL, AWL and Other words) in the form of tables for separated corpus. At the last stage, in order to build AWL words list of the study, concordancer was used again with former process and the results was counted and presented.

Reliability and Validity: The analysis of the data was validated by using qualified and suitable program in order to eliminate the error of the data which may be caused by unsuitable program. In addition, the materials from various writers were used to create generalization in order to avoid bias which may be caused by interests, beliefs or personality of the writers. The validity of research will be also verified by peer debriefing technique.

Findings

The data can be grouped into three important areas: 1) general service words: percentage coverage of the texts and the number of absent words from the corpus 2) academic words: percentage coverage of the texts, overlapping academic words of the texts, and the number of absent words and 3) the corpus of academic words from the studied texts.

The findings of the study were presented in percentage through tables and examples of the results were also shown.

Table 1: Word families, Frequency and Coverage of GSL words and AWL words

Name of Corpus	General Service Words			Academic Words		
	Word Families	Word Occurrences	Coverage	Words Families	Words Occurrences	Coverage
TU	1,266	23,571	83.29%	322	1,961	6.93%
TU1	1,031	14,570	78.16%	283	1,207	6.47%
RSU	1,194	33,119	84.28%	267	2,189	5.57%
RSU1	1,260	32,289	84.13%	323	2,307	6.01%
RMUT	685	6,324	83.15%	113	427	5.61%

Note: TU stands for textbook used in the first academic year of Thammasart University.

TU1 stands for textbook used in the second academic year of Thammasart University.

RSU stands for textbook used in the first academic year of Rangsit University.

RSU1 stands for textbook used in the second academic year of Rangsit University.

RMUT stands for textbook of Rajamangala University of Technology.

As shown in Table1, among five corpora of English for Engineering courses, the GSL words had the highest coverage with around 80%. In conclusion, most part of any text was comprised of words from the GSL. Accordingly, these words should receive attention first because without these it is not possible to understand English without knowing these words. High frequency words or general service words are worth spending time on since they are found over and over in reading. Teachers may deal with these words by pre-teaching them, putting the unknown words in an exercise after the text, or spending time looking at the range of their meanings. Clearly, this group of words is very useful and important for learners of English (Warning 2001).

However, 693 out of a total of 2,284 general service words were not found in the study and these are some examples from absent GSL words.

ability	accept	brave	childhood	debt
abroad	boast	bravery	chimney	deceit
absence	boundary	charm	choice	deceive
absent	bow	chicken	death	decision

Regarding academic words, they accounted for 6% in each approximately. From the study, the course materials for first and second year students from Thammasart University gave the highest coverage by AWL words at 6.93% and 6.47% respectively. However, the results did not meet Coxhead's 10% coverage of academic texts. Moreover, the RU corpus and RMUT corpus were found to have the smallest coverage of AWL. However, as for the AWL comprising 570 word families, 693 of them never occurred in the study.

The examples of absent academic words are listed below

add	immigrate	analogy	ideology
administrate	implicate	anticipate	implicit
advocate	incentive	append	inhibit
aggregate	incline	attribute	insight
albeit	inevitable	attribute	intrinsic

The examples of the corpus of AWL words with their frequency

abstract (8)	decade (2)	inspect (47)	accurate (32)	deny (2)
academy (4)	decline (1)	instance (2)	react (10)	recover (8)
access (13)	define (47)	institute (3)	register (1)	regulate (17)
accommodate (9)	definite (6)	instruct (93)	intelligence (4)	refine (17)
accompany (2)	demonstrate (4)	integral (3)	reinforce (13)	integrate (7)

The AWL can be very helpful in establishing goals for English for Academic Purposes courses. Today, there are many tests based on the AWL that can be used to identify if learners know these lexical items and if effort to teach and learn these words have been successful. Course books specifically designed to teach academic words are also being improved based on the AWL (Coxhead, 2000).

Discussions and Conclusions

Discussions: The results of the study clearly revealed that words from the General Service List played a key role in teaching and learning materials used in English for Engineering courses at Rangsit University, Thammasart University and Rajamangala University of Technology. The majority (around 80%) of corpora were comprised of vocabulary items from the GSL. (83.29% of TU, 78.16% of TU1, 84.28% of RSU, 84.13% 0f RSU1, and 83.15% of RMUT). These results are in agreement with Nation (2001) that high frequency words in the GSL provide over 80% text coverage of most types of reading texts. A number of studies such as Laufer and Nation (1995), Kennedy (1998), and Worthington and Nation (1996), and so on also made the same findings regarding the coverage of GSL words in different texts.

Regarding academic words, the course materials for first and second year students from Thammasart University gave the highest coverage by AWL words at 6.93% and 6.47% respectively. However, the results did not meet Coxhead's 10% coverage of academic texts. Moreover, the RU corpus and RMUT corpus were found to have the smallest coverage of AWL (5.57% and 5.61% respectively). It is obvious that all of these studied texts focus more on general vocabularies than AWL vocabularies.

The results gained from the analysis could be directly useful for teachers or course designers who wish to know what vocabulary is actually taught to Engineering students and if it is worth teaching and suitable for the students. The findings could also aid teachers to decide whether to retain or redesign the teaching materials

Conclusions: An analysis of vocabulary input in English for Engineering courses at Thammasart Universtiy, Rangsit University and Rajamangala University of Technology was carried out to investigate if English for Engineering course materials used in universities provide the right vocabularies for their students.

The results of the study showed that English for Engineering courses used for first year students at Thammasart, Rangsit, and Rajamagala University had a primary emphasis of vocabulary input on high frequency vocabulary or general service words. Similarly, the GSL still played the most key role in the printed course materials for Engineering students used in the second year. Moreover, the majority of general service words occurring in the second year material is a repetition of general service words in first year textbooks. Moreover, the first 1,000 words frequently occurring in the GSL were mostly used in Engineering course materials for both first and second year. In addition, general service words in the study cover nearly three fourths (1,591) of the words in the GSL. It indicated that if Engineering students are able to be exposed to words appearing in their English for Engineering course materials, they will be able to understand over 80% of any kinds of texts. From the findings, it was found that academic words were introduced to students in the first year in all universities (See Table 1). In addition, some of academic words provided for second year Engineering students have already been confronted in the first year course materials. However, a number of unseen academic words were inserted in course materials used in the second year. It can be said that Engineering students are going to encounter more new academic words that are essential for their academic study.

This relationship between textbook used in the first and the second year implied that vocabulary input in English for the first year Engineering students in every university was taught as a foundation in order to provide students with basic words (general service words) and some frequent occurring academic words that could assist students in reading academic texts.

References

Bauman, J. and B., Culligan. 1995. **General Service List**. [Online]. 20 Feb. 1995. Available: http://plaza3.mbn.or.jp/%20bauman/gsl.html. [Accessed 27 Dec. 2009].

Coxhead, A.1998. **An Academic Word List, English Language Institute Occational Publication Number 18**. Wellington: Victoria University of Wellington.

_____. 2000. "A new academic word list." **TESOL Quarterly**. 34, 2: 213-238.

Flowerdew, L. 1997. "Interpersonal strategies: investigating interlanguage corpora." **RELC Journal**. 28, 1: 72-108.

_____. 2003. " A combined corpus and systemic-functional analysis of the problem solution pattern in a students and professional corpus of technical writing." **TESOL Quarterl**. 37, 3: 489-511.

Hunston. S. 2002. **Corpora in Applied Linguistics**. Cambridge: Cambridge University Press.

Kaewphanngam, C., Broughton M.M., and S. Soranasataporn. 2002. "Corpus-based analysis: guidelines for getting practical language input in materials development." **SLLT Journal of Mahidol University**. 2 (December): 16 – 32.

Kennedy, G. 1998. **An Introduction to Corpus Linguistics**. [n.p.]: Addison Wesley Longman.

Laufer, B. and P. Nation. 1995. " Vocabulary size and use: lexical richness in L2 written production." **Applied Linguistic.** 16, 3: 307-322.

Nation, P. and Hwang. 1995. "Where would general service vocabulary stop and special purposes vocabulary begin?." **System**. 23, 1: 35-41.

Nurweni, A. and Read, J. 1999. "The English language knowledge of Indonesia university students." **English for Specific Purposes**. 18, 2: 161-175.

Prapphal, K. 2001. **English Proficiency of Thai Learners and Directions of English Teaching and Learning in Thailand**. Bangkok: Language Institute, Chulalongkorn University.

Thurstun, J. and Candlin, C. 1998. "Concordacing and the teaching of the vocabulary of academic English." **English for Specific Purposes**. 17, 3: 267-280.

Warning, R. 2001. **How Should Teachers Incorporate Vocabulary Teaching into Their Classroom?**. [Online]. Available: http://www1.harenet.ne.jp/%20warning/papers/eltnews.html. [Accessed 3 Apr. 2002]

West, M. 1953. **General Service list of English words**. London: Longman, Green & Co.

Worthington, D. and P. Nation. 1996. "Using texts to sequence the introduction of new vocabulary in an EAP course." **RELC Journal**. 27, 2: 1-11.