

Vocabulary Analysis On Reading Texts Used By EFL Students

C. Sutarsyah

Universitas Lampung, Bandar Lampung

Abstract: The vocabulary in the texts is the aspect that needs to identify. It is claimed that the condition of the words in a text has a great influence to readers' comprehension. It is also commonly believed that comprehension depends on the extent that the words in a text are familiar to the readers. This case study was carried out in the English Education Department of University of Malang. The aim of the study is to identify and describe the vocabulary in the text and to seek if the text is useful for reading skill development. The reading materials under investigation were a collection of reading passages based on the syllabus (Reading Comprehension I) and limited to the passages that were used in class during the second semester, 1999. Based on the nature of the investigation, a descriptive qualitative design was applied to obtain the data. For this purpose, some available computer programs were used. They were used to find the description of vocabulary in the texts. The vocabulary analyses in the texts reveal some constrains. It was found that the texts, containing 7,945 words of 20 different texts, are dominated by low frequency words which account for 16.97% of the words in the texts. In terms of high frequency words occurring in the texts, function words dominate the texts. Of the 50 most frequent words, only two content words (*people* and *say*) were found. In the case of word level, it was found that the texts being used have very limited number of words from GSL (General Service List of English Words) (West, 1953). The proportion of the first 1,000 words of GSL only accounts for 44.6%. The data also show that the texts contain too large proportion of words which are not in the three levels (the first 2,000 and UWL). These words account for 26.44% of the running words in the texts. Based on the findings, some conclusions were drawn, it is believed that the constraints are due to the selection of the texts which are made of a series of short-unrelated texts (20 different topics). This kind of text is subject to the accumulation of low frequency words especially those of content words and limited of words from GSL.

This vocabulary condition could defeat the development of students' reading skills and vocabulary enrichment.

Key Words: Vocabulary, low frequency words, high frequency words, word list, word level.

Reading is a practical, purposeful accomplishment that is vital to learning in school and all through life. Most information passes through the written words, and any one who finds reading difficult is seriously handicapped in the civilized struggle for a place in the world (Rubin, 1982). The ability to read English texts for students who want to increase their knowledge seems inevitable. This is particularly true of English as so much professional, technical and scientific literature is published in English today (Alderson, 1984).

English textbooks dominate most literature, they constitute about ninety per cent of all required textbooks in many colleges (Retmono, 1980). In many countries, like Indonesia, English is learned by a large numbers of students who will never have the opportunity of conversing with native speakers, but who will have access to the literature and periodicals, or scientific and technical journals written in the language they are learning (Rivers, 1981). Thus, reading skill is virtually more important in this context than other skills. In fact, it is not enough for a student to have reading skills in their first language.

On the other hand, despite this specific need for a foreign language, very frequently, students reading in a foreign language seem to read with less understanding than one might expect them to have (Alderson, 1984). It is observable that many students are still not able to comprehend texts when they enter the tertiary level. Day and Bamford (2000) also claim that in general, students learning to read English as a foreign language find it a difficult process, and as a result, they do not enjoy it.

On the other hand, reading skill has become the most important component of English teaching in Indonesia. It has been decided that since the 1967 Ministerial Decree and the 1984 syllabus, the general objective of ELT in the secondary school is to develop reading skill in English (see Huda, 1992). This decision is enhanced by the result of the 1987-1990 national survey of the teaching of English in the secondary school (Huda, 1990) which becomes the basis of the emergence of the 1994 English syllabus for the secondary school. The objectives of ELT in

the secondary school in these two syllabi give a strong emphasis on the development of reading, in addition to listening, speaking, and writing skills. These documents suggest that reading in English is an important skill for secondary school students to equip them in tertiary education and their future life as well.

Of the skills stated, the ability to read today is the most realistic in terms of language use. Indonesian language is used in all schools and colleges. Students are expected to be able to read academic texts in English and have to master a total of 4,000 vocabulary items as set by the 1984 curriculum. For some reasons, in the 1994 curriculum, the number of vocabulary items was reduced to about 2,500 items for Social and Science Program and 3,000 for Language Program (Depdikbud, 1993:1). This vocabulary size is actually sufficient for a student in order to read academic texts. Nevertheless, these objectives are not well achieved.

However, some studies show that the students have low achievement in reading skills and vocabulary gain. It is predicted that the SMU graduates have only mastered about 1,000 words. Results of some researches also support the view that reading in a language which is not the learners' first language is a source of considerable difficulty (Alderson, 1984).

Many efforts have been made to improve the learners' reading ability. Some of them are shown by the existence of books on theories of reading and some books on reading exercises or learning to read. However, as stated by some experts (Baradja and Machmoed, 1998; Sadtono, 1995; Retmono, 1980; Adams, 1980; Eskey, 1988), the fact shows that many learners still encounter problems or difficulties when trying to understand an English text.

If we want to identify what makes for success and what causes failure in learning to read, we will find a very long list of variables that can be drawn up. However, some experts agree that there are three factors which are very dominant, that is, the learners, the teacher, and the materials. In spite of the interdependency of these three factors, each can be considered independently of the others.

Dakin (1973) has an interesting idea about these three variables. She explains that the materials have their own structure because they are written by different people. The material writers, again, have their own perception towards the learners. Thus, the materials have their own plan of development, and their own weaknesses, which both teacher and learners must overcome.

The interdependency can be illustrated in Figure 1 as taken from Dakin (1973: 100). The figure shows that each variable has its own identity and its own rate of development. The two-way arrows show that they have relationship and

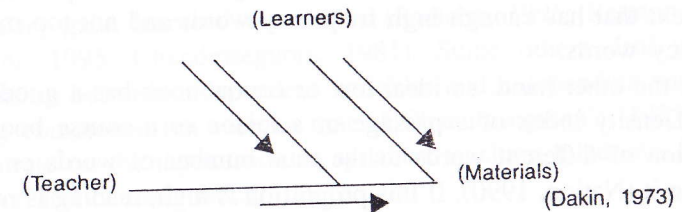


Figure 1 The interdependency of the three factors in learning-teaching

interdependency that are set up in the classroom. The teacher has his own method of teaching and using the materials and his own method towards the learners. The learners develop their own methods of learning to respond to both the teacher and to the materials. Finally, the materials are written by authors who have a particular teacher and particular learners in mind.

Success or failure in learning to read depends on what is learned, not on what the teacher and materials set out to teach. The job of the teacher and the materials is to promote in the learners successful technique of learning (Dakin, 1973). Thus, even though our attention is paid to the factors, we still have too many variables to look at. Based on Dakin's illustration in Figure 1.1, this research, therefore, is only focused on the reading materials used by EFL learners.

According to Nuttall (1985), other common problems which are often faced by the students are familiarizing the features of a text that are known to give rise to difficulties for the readers. These include vocabulary and sentence structure, cohesive devices and discourse markers.

The vocabulary in the texts is the aspect that needs to identify. It is claimed that the condition of the words in a text has a great influence to readers' comprehension. It is also commonly believed that comprehension depends on the extent that the words in a text are familiar to the readers. On the other hand, the familiarity of words implies that the readers have already met the words for many times and have been stored in their long-term memory. These words usually belong to high frequency words and

occur in wide ranges of texts and most of them belong to the most frequent words of General Service List of English Words (West, 1953). In contrast, the words that are infrequent are usually considered difficult for most readers. Some of these words are of technical words and are not in the three *Base Word List* of Nation's (1997). Thus, an ideal text is the text that has enough high frequency words and not too many low frequency words.

On the other hand, an ideal text or course book has a good density index. Density index of a passage or a lesson or a course book is the proportion of different words to the total number of words or running words in it (Nation, 1990). If this proportion is high, reading is relatively easy. In other words, in order to get high-density index, many of the different words must be frequently repeated. This also implies that the text containing too many low frequency words is relatively difficult. According to Nation (1990), the density index of an English course book is usually of the ratio of 1:2.4. This means that, on an average, each word is repeated between two or three times. An efficient course book for beginning students should have density index of around 1:20 (Nation, 1990). In later years the density index should be decreased to about 1:10 or 1:12. A course book containing a large number of words that occur less than five or six times can not be efficient. This course book usually has density index of 1:2.4 or 1:4.

This case study was carried out in the English Education Department of UM. The students are coming from high school at the age of 18. This teacher education program lasts about four years. In this department, English is used as a medium of instruction for most subjects taught in the department. It prepares English teachers to teach English as a foreign language in high schools where reading receives the most emphasis. The main goals of English Department are, first, to improve the students' command of English for the purpose of teaching the language, and, second, to enhance their knowledge in the field of language teaching. As the latter goal, content courses are offered and make up 40 percent of the total courses.

Based on the discussion above, the texts being used for a reading class at the English Department of UM are believed to contain some constraint that make them difficult for the students to read. The study is guided by one main research question: to what extent can the texts being used for a reading class support students' reading skills? Specifically, the

study is focused on the research question: to what extent is the vocabulary in the texts useful for reading skill development?

To the writer's knowledge, there is no thorough investigation on reading materials especially on vocabulary analysis. Most studies deal with the output of learning-teaching, such as, reading achievement and vocabulary gain, learners' test scores (Sadtono, 1995; Retmono, 1980; Nurweni, 1995; Chandrasegaran, 1981). Some others deal with the relationship between independent variables and a dependent variable in reading comprehension such as that of Chandrasegaran's (1981). Some concentrate on the input, such as syllabus, teaching materials, learners' behavior and characteristics.

METHOD

The texts being investigated were the ones used by the students in the class and the ones used for the tests. A computer-based calculation was used by using available program in the computer. To do this, all the texts were retyped to be computer readable data. During the semester, 20 texts were used in both the classroom and the tests. All information indicates the complexity of a text and is used to judge the condition of a text. As it has been discussed, the text condition can influence students' reading performance. The investigation on the texts was made on vocabulary condition. This vocabulary examination was divided into two sections. The first one is dealing with the examination of high and low frequency words occurring in the texts. The second one deals with the classification of words in the texts into three levels.

Computer programs called *WORD* and *Vocaprofile* were used to analyze words in these two investigations. These computer programs were prepared and designed in Computing Service Center, Victoria University of Wellington, New Zealand (Nation, 1997). Word program simply uses an input file of reading text and analyzes it to find the frequency of occurrence of each word. The output of this program is a series of word lists arranged alphabetically with their frequency of occurrences and rank order of word frequency list with their proportion. From this rank order list we can see the most frequent words up to the least frequent words. Both word lists were produced in the form of word types or word forms.

Many experts have claimed that word frequency count is important for a text evaluation and to find the most important words to learn,

because word selection is needed since it is impossible to learn all the words at once. At this context this investigation is useful to see if the texts are appropriate to help develop students' reading skills.

Other investigation on vocabulary is based on word levels. By using the computer program, *Vocaprofile*, the words in the texts were grouped into three levels (three Base Word Lists). The *Vocaprofile* provides a series of selected three Base Word Lists that are used to run the program. The program processes an input file (reading text) and automatically classifies words in the text into three levels based on these three Base Word Lists. The output is a series of word lists arranged alphabetically with their frequency of occurrences.

The main purpose of this analysis is to see what and how many words in the texts are and are not in the lists. By having this, we can see how many words in the texts are among the high frequency words of English (Level I and II or Base Word List I and II) and how many in the University Word List or Level III (Xue and Nation, 1984). This investigation is also useful to see to what extent the texts with the vocabulary they contain can help the learners develop their reading skills.

Base Word List one and Base Word List two consist of the first 2,000 most frequent words taken from *A General Service List of English Words* (West, 1953), called GSL for short. According to West (1953, vii), the words represent a list of 2,000 General Service Words and considered suitable as the basis of vocabulary for learning English as a second language. Thus, Base Word List one consists of the first 1,000 most frequent of GSL and Base Word List two consists of the next 1,000 most frequent words of GSL. Base Word List Three consist of words not in any of these lists, the first 2,000 most frequent words, but which are frequent in university texts from wide range of disciplines. These words were originally from the University Word List prepared by Xue and Nation (1984).

The selection of words for Base Word List one and two is based on word frequency. Words with frequency of 332 or more occurrences per five million running words in the GSL were put in Base Word List One (1st 1,000 most frequent words). Words with frequency of 331 or less per five million running words in the GSL were put in the Base Word List two, 2nd 1,000 most frequent words.

It should be taken into account that the classification of words in the text is mainly based on word families. It means that words with the same

family are considered one word family. Take, for example, the root *relate* with its derivational words, was divided into several word families as follows:

- * relate – relates – related – relatedness – relating - unrelated
- * relation – relations
- * relationship
- * relative – relatively - relativity

Thus, if all these words occur in a text, they are considered as 4 word families, *relate*, *relation*, *relationship*, and *relative* whose frequency of occurrence is 6, 2, 1, and 3 respectively. This word family classification is based on derivational affixation of level 1 up to 4 of Bauer and Nation (1993). The idea is that when a student knows the word *relate*, it is assumed that he knows the word *relates*, *related*, *relatedness*, *relating*, and *unrelated*. Therefore, these words are considered one word family. It is also assumed that he cannot recognize the words *relative* and *relationship* based on the base word *relate*. Therefore, they are considered different word families. In addition, the number of running words and different words or word types in the texts are presented. This data can be used to find the density index of the text being used

RESULTS AND DISCUSSION

As the aim of the study, the investigation on the reading materials is limited to the condition of vocabulary in the texts. This aspect is usually used as the basis to judge the acceptability of a reading text. The discussion on vocabulary is divided into two subsections, that is, the discussion on high and low frequency words in the texts and the classifications of word levels.

High and Low Frequency Words

This investigation identifies the proportion of high and low frequency words that occur in the texts. As has been discussed, the computer program reveals that the texts consist of 7,945 running words. Table 1 lists the distribution of word occurrence beginning from most frequent to least frequent. It was found that the most frequent word occurred 515 times and this accounts for 6.49% of the words in the text. There is only one item (the) found with this frequency. The next most frequent word occurs 238 times and it covers 3.01% of the words. The cumulative percentage of this word together with the previous one is 9.50% and so on.

Looking at the low frequency words, the condition is somewhat different. The table shows that the least frequent words, the words occurring once, dominate the texts. There are 1,346 words or 16.97% of the words of this type found in the texts. This means that in average of ten words about one to two words occur once. Thus, having met these words for the first time, the reader would not meet again. Obviously, no repetition for these words is found by the reader working through the texts.

Furthermore, the number of words whose frequency is 4 and less is 2,847 words or 35.90% of the running words in the texts. While the occurrences of the four most frequent words are, 515, 238, 229, 211 which account for 6.49%, 3.01%, 2.88%, and 2.66% respectively. These, altogether, account for only 15.04% of the words. This indicates that the proportion of low frequency words in the texts is greater than that of the high frequency words.

In the case of vocabulary learning through reading, too much proportion of low frequency words is of a little value. This suggests that students working their way through all the texts would be continuing meeting words that they would not meet again or would meet only a few times in the texts. Ideally, the texts do not consist of too many low frequency words but they have more high frequency words, especially

Table 1: The distribution of high and low frequency words in the texts

No	Frequency/range	Number of words	Cumulative percentage (%)	Percentage of words (%)
1	515	1	6.49	6.49
2	238	1	9.50	3.01
3	229	1	12.38	2.88
4	211	1	15.04	2.66
5	201	1	17.58	2.54
6	187	1	19.94	2.36
7	117	1	21.41	1.47
8	102	1	22.70	1.29
9	66-87	4	26.39	3.69
10	46-65	7	31.29	4.90
11	40-45	5	34.02	2.73
12	28-39	8	37.38	3.36
13	25-27	6	39.36	1.98
14	22-24	6	41.11	1.75
15	19-21	7	42.89	1.78
16	17-18	9	44.85	1.96

No	Frequency/range	Number of words	Cumulative percentage (%)	Percentage of words (%)
17	14-15	13	47.24	2.39
18	12-13	13	49.27	2.03
19	10-11	27	52.86	3.59
20	9	9	53.88	1.02
21	8	12	55.09	1.21
22	7	26	57.39	2.30
23	6	42	60.57	3.18
24	5	56	64.10	3.53
25	4	93	68.79	4.69
26	3	153	74.58	5.79
27	2	335	83.03	8.45
29	1	1346	100.00	16.97

those of content words. At this point, long continued texts of the same topic might be more valuable than a series of unrelated texts. This kind of text can provide a lot of repetition that is important for optimal learning to occur. Whilst, a series of unrelated texts must use a bigger proportion of low frequency words.

Furthermore, it was also found that most of the high frequency words in the texts consist of function words. Of the first 50 most frequent words, there are only two content words, i.e., *people* and *said*; the rest are, of course, function words. Whilst, in a study done by Sutarsyah *et al.* (1994), there are 18 content words in the 50 most frequent words from economic corpus.

From the list, we can also identify some function words from these first 50 most frequent words successively, i.e., *the, to, of, a, and, in, he, was, it, and had*. In the second 50 most frequent words, more content words are found, such as *read, book, like, now, come, long, make, eyes, look, money, manager, see, top*, etc. This indicates that many function words dominate high frequency words in the texts. The texts being used are not rich enough as to develop students reading skills because too many function words and limited content words occur in the texts. Many content words are of low frequency in the texts. This means that the students have a little opportunity to learn and enrich their vocabulary as to increase their reading skills because they are not provided by enough repetitions of these words. This text condition is of a little value for the students who learn to read.

To conclude, we can have some points that can be taken into account. The proportion of low frequency words found in the texts is

much greater than that of the high frequency words. Thus there is a tendency for a reader to forget the low frequency words when meeting for the first time in the series of unrelated texts.

The function words dominate the texts with considerably high frequency. It is true that function words are used more frequently in any text and they are relatively fewer than content words (Bolinger, 1975). These words belong to classes that are relatively closed. Whilst, content words, which are of a great value for reading learners, are very limited in the texts, and yet their occurrences are considered low. This implies that learning content words are more important than function words because function words do not develop. Learning function words does not give much contribution to increase students' reading performance.

Word Levels

As has been mentioned, this study also classifies words in the texts into three levels based on the three base word lists provided in the computer program. Table 2 shows that most words belong to the first level. There are 6515 tokens or running words (82%) which account for 1147 word types or 54%. This makes up 706 word families with the coverage of 44.65%. In other words, less than a half of the word families in the texts are included in the most frequent 1,000 words of GSL.

At level two, the numbers of word families is almost a half as many as the number of word families at level one (the first 1,000 words), that is, 335 word families or 21.19%. These two groups of words belong to the first 2,000 words GSL,

Table 2 The number of words and percentage of coverage in the texts

Word Level	Tokens	Types	Families
The first 1,000	6515 / 82.0%	1147 / 54.40%	706 / 44.65%
The second 1,000	608 / 7.7%	402 / 19.1%	335 / 21.19%
UWL	185 / 2.3%	140 / 6.6%	122 / 7.72%
Others	637 / 8.0%	418 / 19.8%	418 / 26.44 %
Total	7945/ 100.00%	2107 / 100.00%	1581 / 100.00%

that is 65.84% of the word families in the texts. In fact, with this proportion of high frequency words in the texts, it seems rather difficult for the students to read the texts. Ideally, to read optimally, a reader is

expected to know 90% coverage of words in the texts. These words are mostly available in the first 1,000 words and some in the second 1,000 words. In other words, the first 2,000-word level is the basic vocabulary needed for every learner of English as a second language to read English text and to have other skills as well. Numerous studies, for example, Hwang, 1989; Hwang and Nation, 1995; Sutarsyah *et al.*, 1994, have confirmed that this high frequency vocabulary accounts for around 80% of the total running words (tokens) of the texts

Further more, the number of the words at level three or University Word List in the texts is very limited. It consists of 122 word families or 7.72% of word families. In fact this group of words is important for the learners to develop their reading skills in academic study.

Studies on word analysis with two corpora, economics and academic texts by Sutarsyah *et al.* (1994) and Sutarsyah (1993) show a different result from the present study (see Table 3). The table shows that the proportion of word families at level one in both corpora 77.72% and 74.11% and at level two, 4.78% and 4.32%. Thus, the texts in the present study are dominated by words from the second 1,000 words level. In fact, the texts need more words from the first 1,000-word level so that they can read optimally. The table also shows that the other two word analyses are comparable in that they have almost the same proportion. However, the proportions of words in the texts in the present study are quite different from the other two.

Table 3 The comparison of word coverage of the present study and the study of Sutarsyah *et al.* (1994)

Word Level	Present study	Sutarsyah <i>et al.</i> (1994) and Sutarsyah (1993)	
		Economic corpus	Academic corpus
The first 1,000	44.65%	77.72%	74.11%
The second 1,000	21.19%	4.78%	4.32%
UWL	7.72%	8.74%	8.40%
Others	26.44%	8.77%	13.16%

Another constraint is that a group of words not in any of these levels is too large, that is, 26.44% of the total word families in the texts. Most of the words in this group are not very useful for student's reading development. These words are uncommon in most texts. Some of them

are the name of people, places or countries such as *Africa, America, British, Burma, China*, etc. The words on the name of people were even found to have high frequency of occurrence such as *Della* (7), *Cheseborough* (11), *Henry* (15), *Jim* (14), etc. More over, it was also found that a number of words belong to technical or subtechnical words which are usually difficult for learners. Take, for example the words *apothecaries, audible, catastrophes, casmo, damning, drought, em-bankments, extinguish*, etc. This seems to be particular to and useful for a specific area of knowledge. On the other hand, in this group we can find many complicated words such as *apologetically, drifting, complacency, disastrous, engenders*, etc. Therefore, readers whose experience is limited in such specific area will find the texts difficult.

In terms of density index, the text or the course book being used is considered rather difficult for students. It is found that, based on the data in Table 2, the density index of the text is 1: 3.77 (2107:7945), which means that on an average each word is repeated between three to four times. This density index is still considered low. In terms of vocabulary learning and reading development for the first year students the text with this index is inefficient. As has been mentioned, an efficient textbook for this group of students has density index of about 1:20 with a low number of "one timers". In fact, most modern English textbooks have density index of 1: 2.4 containing 40% of words of one occurrence (Nation, 1990).

It seems that the unfavorable aspect of the texts is mainly due to the choice of texts which are made of a series of unrelated texts. This kind of texts can increase a vocabulary load of the course enormously because the texts consist of many low frequency content words and technical words. Thus, it seems that the texts cannot provide enough repetition for content words which are valuable for skill development.

This finding is in line with the study done by Sutarsyah *et al.* (1994) which compared one coherence, continuous text by one writer on one topic and a series of unrelated texts made up of a variety of academic texts by many writers. The study reveals that a coherence text by a single writer on a single broad topic uses a very much smaller vocabulary than a series of unrelated texts. This means that vocabulary load increases when learners are provided with reading materials containing many different texts of unrelated topics.

CONCLUSIONS AND SUGGESTIONS

The investigation on vocabulary yields some points. This part has examined the vocabulary used in the texts in terms of frequency of occurrences and word levels. It reveals that function words dominate high frequency words in the texts. The texts are also dominated by low frequency content words. This condition is not favorable for skill development because the students have a little opportunity to learn these important words. In other words, the texts cannot provide enough repetition of content words for students to learn and thus, the text is difficult to read.

In terms of word levels, the texts consist of too many words outside the three levels. This implies that the students face with vocabulary learning load because the texts contain large number of vocabulary of this type. In addition, the texts provide limited words from the first 1.000 words of GSL so that the students have difficulty to read them. At the same time, the proportion of the words in the first 2,000-word level is also not enough to provide the students to read optimally.

One reason for this constraint is that the texts were made of a series of unrelated texts with different topics or themes. Because of different topics, the texts or the course book tend to have many different words and they are usually of low frequency of occurrences.

Thus, the teachers and course designers are suggested to avoid the use of reading materials consisting of a series of unrelated texts, even though it is quite impossible to have the whole reading materials made of related texts for one semester. It is worth considering making the course consist of a few themes so that the texts within a theme bear more relationship to each other and thus make use of smaller vocabulary.

REFERENCES

- Adams, M.J. 1980. Failure to Comprehend and Levels of Processing in Reading. In R.J. Spiro, B.C. Bruce, and W.F. Brewer. (Eds.), *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education*. Hillsdale, New Jersey: Lawrence Earlbaum Associates: 11-32.
- Alderson, J.C. 1984. Reading in a Foreign Language: A Reading Problem or a Language Problem? In J.C. Alderson and A.H. Urquhart (Eds.), *Reading in a Foreign Language*. London: Longman.

- Baradja, M.F. and Machmoed, Z. 1998. *The Teaching of English in Indonesia: Then and Now*. A paper presented at the National Conference on ELT. Malang.
- Bauer, L. and Nation, I.S.P. 1993. Word Families. *International Journal of Lexicography*, 64, 253-279.
- Bolinger, D. 1975. *Aspect of Language*. New York: Harcourt Brace Jovanovich.
- Chandrasegaran, A. 1981. *Problems of Learning English as a Second Language*. Singapore: SEAMEO Regional Language Center.
- Dakin, J. 1973. The teaching of Reading. In H. Fraser and W.R. O'Donnell (Eds.), *Applied Linguistics and the Teaching of English*. London: Longman.
- Day, R.R. and Bamford, J. 2000. Reaching Reluctant Readers. *English Teaching Forum*, 38(3): 12-17.
- Depdikbud, 1993. *Garis-garis Besar Program Pengajaran (GBPP) Mata Pelajaran Bahasa Inggris Sekolah Menengah Umum (SMU) tahun 1994*. Jakarta: Depdikbud.
- Eskey, D.E. 1992. Holding in the Bottom: An interactive approach to the language problems of second language readers. In P.L. Carrell, J. Devine, and D.E. Eskey (Eds.), *Interactive Approach to Second Language Reading*. (pp. 93-100) Cambridge: Cambridge University Press.
- Huda, N. 1990. A Survey of the Teaching of English in Secondary Schools in Eight Provinces. *TEFLIN Journal*, 3 (1).
- Huda, N. 1992. The 1994 English Syllabus for Secondary Schools: Issue and Problems. *Bahasa dan Seni*, 20 (51): 2-14.
- Hwang, K. 1989. *Reading Newspaper for the Improvement of Vocabulary and Reading Skills*. Unpublished MA Thesis, English Language Institute, Victoria University of Wellington.
- Hwang, K and Nation, I.S.P. 1995. Where would General Service Vocabulary Stop and Special Purposes Vocabulary Begin? *System*, 23(1): 35-41.
- Nation, I.S.P. 1990. *Teaching and Learning Vocabulary*. Boston, Massachusetts: Heinle & Heinle.
- Nation, I.S.P. 1997. The Language Learning Benefits on Extensive Reading. *The Language Teacher*, 21(5): 11-18.
- Nurweni, A. 1995. *The English Vocabulary Knowledge of the First Year Students in an Indonesian University*. Unpublished M.A. Thesis. Wellington: Victoria University of Wellington, New Zealand.
- Nuttall, C. 1985. *Teaching Reading Skills in a Foreign Language*. London: Heinemann Educational Books.
- Retmono. 1980. Developing Culturally Appropriate Reading Materials for the Teaching of English as a Foreign Language in Indonesia. In E.M. Anthony and J.C. Richards (Eds.), *Reading: Insights and Approaches*. Singapore: SEAMEO Regional Language Center.
- Rivers, W.M. 1981. *Teaching Foreign Language Skills*. Chicago: The University of Chicago Press.
- Rubin, D. 1982. *A Practical Approach to Teaching Reading*. New York: Holt, Rinehart and Winston.
- Sadtono, E. 1995. *Perspektif Pengajaran Bahasa Inggris di Indonesia*. Malang: FPBS IKIP Malang.
- Sutarsyah, C. 1993. *The Vocabulary of Economics and Academic English*. Unpublished MA Thesis. English Language Institute, Victoria University of Wellington, New Zealand.
- Sutarsyah, C., Nation, P., and Kennedy, G. 1994. How Useful is EAP Vocabulary for ESP? A corpus based case study. *RELC Journal*, 25 (2): 34-50.
- Wadsworth, B. J. 1978. *Piaget for the Classroom Teacher*. London: Longman.
- West, M. 1953. *A General Service List of English Words*. London: Longman.
- Xue, G. and Nation, I.S.P. 1984. A University Word List. *Language Learning and Communication*, 3(2): 215-229.