

**VOCABULARY PROFILE OF GRADED READERS USED IN
THE FACULTY OF LANGUAGE AND LITERATURE, SATYA
WACANA CHRISTIAN UNIVERSITY**

THESIS

Submitted in Partial Fulfillment
of the Requirements for the Degree of
Sarjana Pendidikan



Pawinda Tiffany Putri

112012126

**ENGLISH LANGUAGE EDUCATION PROGRAM
FACULTY OF LANGUAGE & LITERATURE
SATYA WACANA CHRISTIAN UNIVERSITY
SALATIGA**

2016





PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : PAWINDA TIFFANY PUTRI
NIM : 112012126 Email : putri.tiffany@gmail.com
Fakultas : BAHASA DAN SAstra Program Studi : PEMBIDIKAH BAHASA INGGRIS
Judul tugas akhir : VOCABULARY PROFILE OF GRADED READERS USED IN THE FACULTY OF LANGUAGE AND LITERATURE, SATYA WACANA CHRISTIAN UNIVERSITY
Pembimbing : 1. Prof. Dr. Gusti Aprika, M.A
2. Anne Indrayanti Timotius, M.Ed

Dengan ini menyatakan bahwa:

1. Hasil karya yang saya serahkan ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar kesarjanaan baik di Universitas Kristen Satya Wacana maupun di institusi pendidikan lainnya.
2. Hasil karya saya ini bukan saduran/terjemahan melainkan merupakan gagasan, rumusan, dan hasil pelaksanaan penelitian/implementasi saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing akademik dan narasumber penelitian.
3. Hasil karya saya ini merupakan hasil revisi terakhir setelah diujikan yang telah diketahui dan disetujui oleh pembimbing.
4. Dalam karya saya ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali yang digunakan sebagai acuan dalam naskah dengan menyebutkan nama pengarang dan dicantumkan dalam daftar pustaka.

Pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari terbukti ada penyimpangan dan ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya saya ini, serta sanksi lain yang sesuai dengan ketentuan yang berlaku di Universitas Kristen Satya Wacana.

Salatiga, 16 Juni 2016



Aprianti

PAWINDA TIFFANY PUTRI

Linda sirip & sama sirip melautkan



PERNYATAAN PERSETUJUAN AKSES

Saya yang bertanda tangan di bawah ini:

Nama : PAWINDA TIFFANY PUTRI
NIM : 112012120 Email : putrytiffany@gmail.com
Fakultas : BAHASA DAN SAstra Program Studi : PENDIDIKAN BAHASA INGERIS
Judul tugas akhir : VOCABULARY PROFILE OF GRADED READERS USED IN THE FACULTY OF LANGUAGE AND LITERATURE, SATYA WACANA CHRISTIAN UNIVERSITY

Dengan ini saya menyerahkan hak *non-eksklusif*** kepada Perpustakaan Universitas – Universitas Kristen Satya Wacana untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tugas akhir elektronik sebagai berikut (beri tanda pada kotak yang sesuai):

- a. Saya mengizinkan karya tersebut diunggah ke dalam aplikasi Repositori Perpustakaan Universitas, dan/atau portal GARUDA
- b. Saya tidak mengizinkan karya tersebut diunggah ke dalam aplikasi Repositori Perpustakaan Universitas, dan/atau portal GARUDA**

* Hak yang tidak terbatas hanya bagi satu pihak saja. Pengajar, peneliti, dan mahasiswa yang menyerahkan hak non-eksklusif kepada Repositori Perpustakaan Universitas saat mengupload hasil karya mereka masih memiliki hak copyright atas karya tersebut.

** Hanya akan menampilkan halaman judul dan abstrak. Pilihan ini harus dilampiri dengan penjelasan alasan tertulis dari pembimbing I/A dan diketahui oleh pimpinan fakultas (dekan/prodekan).

Demikian pernyataan ini saya buat dengan sebenarnya.

Salatiga, 16 Juni 2016

1956

Mengetahui,

Prof. Dr. GUSTI Astika, M.A

Tanda tangan & nama terang pembimbing I

PAWINDA TIFFANY PUTRI

Tanda tangan & nama terang mahasiswa

Anne Indrayanti Timotius, M.Ed

Tanda tangan & nama terang pembimbing II

VOCABULARY PROFILE OF GRADED READERS USED IN
THE FACULTY OF LANGUAGE AND LITERATURE, SATYA
WACANA CHRISTIAN UNIVERSITY

THESIS

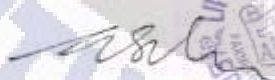
Submitted in Partial Fulfillment
of the Requirements for the Degree of
Sarjana Pendidikan

Pawinda Tiffany Putri

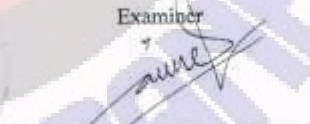
112012126

Approved by:

Supervisor


Prof. Dr. Gusti Astika, M. A.

Examiner


Anne Indrayanti Timotius, M. Ed.

INTRODUCTION

The way to learn vocabulary is not always from reading academic books. It can also be done through different ways such as reading stories or watching movies. According to Nation (2001, p.17) different genres of reading materials have different frequency profiles. The vocabulary used in the academic books and story books have different levels of difficulties. In the Faculty of Language and Literature (FLL) of Satya Wacana Christian University (SWCU), both in English Language Education Program and English Literature Program, there is a course named Extensive Reading which requires the students to read a minimum of 1000 pages in order to pass the course. Nation (2001, p.144) found that the learners will be able to increase their comprehension ability in reading text after passing a certain boundary described as vocabulary 'thresholds'. The criteria for the 'thresholds' was that the learners have to acquire 95% of vocabulary from the reading passage in order to understand the text. According to that statement, vocabulary certainly is important in reading. As mentioned above, by reading around that 1000 pages, learners' vocabulary are expected to increase.

The books used in the Extensive Reading course are graded readers which are Penguin Readers and Cambridge English Readers. The books which can be found in the SWCU library are published by Pearson Education Limited and Cambridge University Press. According to Kurniawati (2015), a lecturer of Extensive Reading course, in her syllabus design, the course aims to increase learners' interests toward reading in English. The Extensive Reading course is in

the beginning of the semester. If the academic books are used in the first semester, the learners will find English as a difficult subject. Therefore, graded readers are used. Based on this reason, profiling the vocabulary used in the graded readers is important, because those books have different levels, from level 1 up to 6. This means that the books have different levels of vocabulary difficulties. It is important to profile the graded readers used in this department because by providing an appropriate assistance for the new students will help them to acquire important vocabulary knowledge during the course.

Based on the explanation above, the research questions for this study were:

1. What is the vocabulary profile of graded readers used in the Faculty of Language and Literature, Satya Wacana Christian University?
2. What is the proportions of vocabulary that is not covered in the graded readers?
3. What is the token (word) recycling index of the graded readers?

Then, the objective of this study were to identify the vocabulary profile of the graded readers used in English Language Education Program at the Faculty of Language and Literature, Satya Wacana Christian University. Besides, this study also aimed to produce lists of vocabulary that were not covered in the graded readers and the last was to produce the token (word) recycling index of the graded readers.

By doing this research, I hoped that the learners would have information to improve their vocabulary they have learned through reading graded readers. Moreover, through these books, the learners can learn vocabulary from the beginner level to the advanced level and measure their own vocabulary knowledge. I also hoped that the result of this study could help the teacher to teach the learners about reading comprehension by knowing which words should be taught first. The more vocabulary known by the students the easier they will understand the reading passages. In that way, teachers will not spend too much time to teach all the vocabulary in the text, they can focus on what learners need.

LITERATURE REVIEW

A. Learning Vocabulary through Reading

In learning English as a foreign language, vocabulary is the biggest problem that learners should face. According to McCharty (1992), cited in Asgari and Mustapha (2011), without words to express meaning, communication in L2 will be impossible to happen in any meaningful way. That means vocabulary is everything in language learning. There are many ways to learn vocabulary, especially when the learners should learn it within the context or discourse. The learning vocabulary can be done through reading, conversation, movies, television, radio, etc (Nation, 2001). It is widely known that reading is one of the effective methods to acquire new vocabulary for the L2 learners. Though another method can be used to support learning vocabulary such as listening, imitating, etc., reading is still assumed as the best way.

Carver (2000, 2003, as cited in Grabe, 2009) claimed that vocabulary knowledge and reading comprehension has almost a perfect correlation. To support that statement, Pike (1979), cited in Grabe (2009), reported the correlations of vocabulary and the reading on TOEFL test is $r = .88$ to $r = .94$. It means that reading and vocabulary are things which cannot be separated. Fisher and Frey (2014) stated that vocabulary lies at the heart of content learning, as it serves as a proxy for learners' understanding of concepts. In reading activities, when the learners cannot understand the meaning of the words, they might not understand the content of the book and get nothing from it. In many cases learners experience difficulties in understanding the meaning of a text because it contains numbers of unknown words. They usually skip reading the sentence or the paragraph (Noor & Amir, 2009) and do not try to know the words. As a result, the learners may leave the words to be unknown and they may not increase their vocabulary knowledge. That is why vocabulary becomes an important key in learning the target language.

Learning vocabulary through reading has various types of learning activities depending on the learner's need. There are two types of reading activities, intensive reading and extensive reading. According to Nation (2008, p. 59) in intensive reading learners are engaged with the details of the text which have two goals: understanding the text and language features. This activity could help the learners develop strategies to deal with the unknown words in the reading text. Nation (2008, p. 59) found that the reading passages such as newspaper, journal articles, reading passages in the text book and so on used in intensive

reading are short, around 300-500 words long. On the other hand, Nation (2008, p. 70) stated that extensive reading involves the reading activity which focuses on the meaning of the text rather than on its language feature. Examples of extensive reading text commonly used are graded readers, novels, and other non-fiction books.

In conclusion, learners can acquire vocabulary knowledge through reading abundantly. According to Webb and Chang (2015), the more texts that are read by the learners, the more opportunities to meet the repeated words. In this study, the focus is on extensive reading activity through graded readers. If two or more graded readers are read in a week, it is expected that the learners' vocabulary knowledge will increase incidentally by reading the same words in a short period of time. Webb and Chang (2015) stated that the purpose of reading graded readers is to develop high frequency words, not to develop vocabulary for a certain topic or genre for example like vocabulary used in biology. Thus, the learners will find their vocabulary knowledge they get through graded readers is useful despite of the reading materials used later on.

B. Teaching Vocabulary through Reading

The reading activities in intensive and extensive reading have their own goals for the learners. That is why knowing L2 learner's needs become an important factor in teaching vocabulary later on. The teacher should know how much the learners are familiar with the words in order to understand the text. Nation (2001) proposed that the minimum percentage of the text coverage which

can make the vocabulary learning to occur is 95% (p. 148). This percentage of the word frequency is expected not to burden the learners. When the reading text has 80% or below the 95% of the text coverage, the text is quite difficult for the learners to understand. This amount of percentage is also intended for the non-fiction books or other reading passages for general or academic purposes. Based on Hu and Nation (2000), cited in Schmitt, Jiang and Grabe (2011), it takes 98%-99% coverage of reading for pleasure.

However, language teachers usually tend to teach their students the vocabulary which they think students will encounter. Very often, the teachers do not have enough time to teach all the vocabulary their students may need (Alhaysony, 2012). This is the case which usually happens in the traditional classroom learning activity. The teacher will directly translate difficult words into the learners' mother tongue (L1). Some people believe this is an easy way, but then actually it takes so much time and the probability of students' dependence toward the teacher may increase. Schmitt (2000), cited in Alhaysony (2012), states that there is individual differences between the learners, therefore, it will be quite difficult in deciding what vocabulary students should learn.

Based on the explanation above, there are some effective ways that the teacher can do to overcome this problem. One of them is extensive reading. Eskey (2005), cited in Li and Zhang (2009), argues that when the learners found at least just one unknown word in every twenty words, they will not feel frustrated and will continue to read. Thus, the teacher should provide appropriate text materials to meet the learner' needs on learning new vocabulary in order to achieve the

expected goal. The goals are developing learners' fluency in reading, reading comprehension, and reading for pleasure (Nation, 2008). For instance, graded readers which are included in fiction book can be used for extensive reading learning activities. The teacher can give the reading materials which are somewhat challenging but not exceeding the learners' lexical competence too much (Li & Zhang, 2009). By using graded readers as the reading materials, the learners can start reading from the level in which they are able to handle then go to the higher one. Yet, the teachers have to facilitate the students in learning vocabulary. For example, they can teach the students how to use some strategies in learning new words, like using dictionaries, guessing words, taking notes or even social strategy. Based on Alhaysony (2012), teachers are often in the position of the one who knows everything and can answer the students' question. Therefore, when the students ask about the unknown words to the teacher, they can answer question using the L1 translation, synonym, definition, or using the new word in a sentence (Alhaysony, 2012). Through those strategies, it will help the students to understand the new vocabulary easily. Besides, the teachers can guide their students to learn the words they may need more like high frequency words. As a result, the students' vocabulary learning activity through reading can be done smoothly.

C. Vocabulary Profiler

When learning and teaching vocabulary are related to reading activity, vocabulary profile has become an important issue. The issue about vocabulary profile has been discussed frequently by researchers, so it is not a new thing

anymore. Graves (2015) stated that vocabulary profile is an aggregation of word frequencies. Meanwhile, vocabulary profile or text coverage is the total number of words used in the text, minus the amount of unused words (negative words) and it is converted into a percentage (Laufer & Ravenhorst-Kalovski, 2010). If the reading activity occurs, many kinds of reading materials might be used such as newspapers, graded readers, textbooks, etc. The important point is that every genre of reading materials has different kind of vocabulary style and difficulty. Thus, profiling them is important and useful to differentiate the reading materials considering the learners' needs and reading goal. In this study, my objective was to identify the vocabulary used in graded readers. Nation (2008, p.71) reported that graded readers cover around 300-500 words for the beginner level, and 2000-2500 for the advanced level. Thus, this study aimed to find the percentage of the text coverage of graded readers, because the books have different levels, level 1 up to 6.

To reach the objective which has been set, the texts in the graded readers were analyzed using lexical profiler available at Tom Cobb's website (<http://lxtutor.ca>) (Laufer & Ravenhorst-Kalovski, 2010). The texts were copied into the vocabulary profiler to produce the analysis. Based on Nation (2001, p. 11-12) there are four levels of vocabulary in a text: the first group is *high-frequency words* which are included in *A General Services List of English Words*. It contains around 2,000 word families. The 2,000 words are divided into the first most 1,000 words frequently used (K-1) and the second most 1,000 words frequently used (K-2). The example of K-1 words are *across, after, again, and, but*, etc. Meanwhile,

the example K-2 words are *afternoon, quickly, suddenly, tomorrow*, etc. (taken from graded reader, *Twenty Thousand Leagues Under The Sea* chapter 2 by Jules Verne). The second group is the *academic words list* (AWL). This small list of words is very important for anyone using English for academic purposes. The example of AWL are *accurate, analysis, chapter, communicating, consists*, etc. (taken from the eight grader Junior High School textbook, *When English Rings A Bell* chapter X). The third group is *technical words*, the words which are used in a certain topic or certain areas. For instance the words used in biology, such as *evolution, enzymes, cancer, meiosis*, etc. Then, the fourth group is low-frequency words (off-list words), all the words which are not included in the *high-frequency words, academic words and technical words*. For example, *bathroom, benches, naughty*, etc. (taken from the eight grader textbook, *When English Rings A Bell*).

D. The Previous Study

Similar studies such as this one have been done by the researchers. First, is on graded readers of Oxford Bookworms Level 1, *White Death*, has 6, 869 running words in the book. The 6,035 tokens make 87.86% coverage of the text. Meanwhile the occurrences of proper nouns are 6.39%, therefore the proportion is close to 95% coverage. Then guessing from context technique can be used if some difficult words appear (Wan-a-rom, 2008).

Table 1 below was a research done by Nation in 2001. Nation (2001, p. 152) did a research on the frequency of vocabulary on graded readers, *Indonesian Love Story* written by Mesiter and Nation (1981), and other book titles.

The result shows that from the most frequent 2000 words of English plus names, the percentage calculated for *Indonesian Love Story* was 99%. In *The Three Little Pigs* in the Ladybird series (Elley and Mangubhai, 1981: 26) cited in Nation (2001, p. 153) the percentage found was 96.7%, and in the *Dry Days for Climbing George* written by Margaret Mahy (1988) was 93.6%. The results are quite good, because it does not go far beyond the percentage of coverage predicted. The learners will be able to read and understand the book's content.

Table 1. The percentage coverage

| Books' title | 1st 1,000 | 2nd 1,000 | Names | Remaining words |
|-------------------------------------|-----------|------------------|-----------------|-----------------|
| <i>Dry Days for Climbing George</i> | 76.8% | 11.7% (88.5%) | 5.1% (93.6%) | 6.4% (100%) |
| <i>The Three Little Pigs</i> | 78.1% | 11.1% (89.2%) | 7.5% (96.7%) | 3.3% (100%) |
| <i>Indonesian Love Story</i> | 82.7% | 8.4% (91.1%) | 7.9% (99%) | 1.0% (100%) |

In addition, Jusuf (2014) did a research in Satya Wacana Junior High School, Salatiga about the texbook they used. In the table below, she reported the result of vocabulary coverage analysis from 50% of the reading texts only which are taken from English textbook, *English on Sky* for grade 7.

Table 2. The percentage coverage of *English on Sky* for grade 7, chapter 1-7

| Chapter | K1 words (1-1000) | K2 words (1001-2000) | AWL (Academic Word List) |
|---------|-------------------|----------------------|--------------------------|
| | | | |

| | | | |
|---|--------|--------|-------|
| 1 | 68.91% | 7.00% | 0.10% |
| 2 | 83.91% | 5.59% | 0.19% |
| 3 | 79.53% | 4.57% | 3.55% |
| 4 | 80.21% | 12.53% | 0.95% |
| 5 | 83.81% | 5.17% | 0.56% |

Jusuf (2014), found that in every chapter of *English on Sky* textbook for grade 7, there are different topics discussed which influence the vocabulary used. Thus, she suggested that the teacher should have reference lists of words whether a particular word needs attention or not.

In SWCU there were some studies which have been done by researchers on vocabulary profile. They used different kinds of data, like junior high school textbook, IC textbook, etc. However, the analysis of vocabulary profile of graded readers have not been done. Thus, I decided to analyze the graded readers's vocabulary coverage.

THE STUDY

Method of Research

This study used a descriptive method in analyzing the data. According to Aggarwal (2008) cited in Salaria (2012), descriptive research is intended to collect the information which is revealing conditions or situation for the purpose of description and interpretation. Sometimes it used visual aids such graphs, charts or tables to help the reader in understanding the data. This type of research was not simply collecting and tabulating facts, but it included proper analyses,

interpretation, comparisons, identifications, etc. According to the statements, this study described the vocabulary profile used in the graded readers. The graded readers are used in the Faculty of Language and Literature, Satya Wacana Christian University. This context was chosen because both of English Language Education Program and English Literature Program use graded readers as their reading materials in a course named Extensive Reading. Then, the description of this study covered the analysis of vocabulary words which separated into *high frequency words* and *low frequency words* used in the first level of graded readers up to the sixth.

Sample

As the students in the Faculty of Language and Literature used graded readers in their first semester, then it became the samples in this study.

Table 3. The list of graded readers' title used as the sample for the analysis

| Title | Level | Total Chapters | Total pages excluding the pictures |
|--------------------------------|-------|------------------------|------------------------------------|
| <i>Twenty Thousand Leagues</i> | 1 | 7 | 14 |
| <i>Under The Sea</i> | | | |
| <i>Five Famous Fairy Tales</i> | 2 | Compilation of stories | 30 |
| <i>The Lahti File</i> | 3 | 14 | 56 |

| | | | |
|--------------------------------------|---|----|-----|
| <i>The Amsterdam Connection</i> | 4 | 16 | 72 |
| <i>In the Shadow of The Mountain</i> | 5 | 21 | 91 |
| <i>The Double Helix</i> | 6 | 15 | 102 |
| Total | - | 73 | 365 |

In Table 3 showed the list of graded readers used as the samples. The graded readers found in the Satya Wacana Christian University's library were Penguin Readers and Cambridge English Readers. These books were published by Pearson Education Limited and Cambridge University Press. I used stratified and systematic sampling technique to select the data from the sample graded readers.

As the sample of this research, I took 50% of pages each book.

Research Instrument

In order to produce the analysis of vocabulary used in graded readers, I used a tool named vocabulary profiler. This tool can be found in a website, www.lexutor.ca. It is the Tom Cobb's website available in *lexutor.ca*. There were many functions of the tool related to this study. Like, first it showed the result of vocabulary's frequency used in the text. It can also show where the words belong, *high frequency words* or *low frequency words*. It can give analysis of the negative words found in the graded readers. Besides, it can also compare the inputted sample texts.

Data Collection Procedures

The graded readers were selected as much as 50% from the whole pages due to the limitation of inputted words in the vocabulary profiler, *lextutor.ca* which is 60,000 words. Meanwhile, there were 6 graded readers being analyzed. However, the texts were cleaned up from the unnecessary components. First, numbers (including page, chapter, date, year, age, etc.) and punctuation marks were excluded from the data. Second, proper names including names of persons, places, brand of goods, and so on also were deleted from the data. The pages selected and cleaned were copied into the tool as the sample to analyze the vocabulary used in the graded readers. Actually the tool can read this kind of data, but they were excluded from analysis because those unnecessary component influenced the percentage of the analysis later on, meanwhile those components were easily recognized by the students in their reading activity.

Data Analysis Procedures

Then, from the pages which had been selected, the text were copied into the tool provided. The first step was opened the *lextutor.ca* website, then clicked the *Vocabprofiler*. After that, the *VP Classic* and the blank space were appeared. In this space the text copied and clicked *submit the window*. Then, automatically the analysis of the cumulative percentage showed. After that, the description of the vocabulary used based on the results appeared in the tool as well as the button of *negative VP* which was placed beside the K-1, K-2, AWL, and *off-list* vocabulary list. Then, clicked the *negative VP* and the analysis of words which

were not covered or the negative words appeared. Therefore, the whole data were put to show the overall result of the analysis. Then, after that, the data were put one by one from the first level up to the sixth level of graded readers to show the vocabulary profile differences occurred between the level. To find out the *token (word) recycling index*, the comparing texts were copied into *Text Lex Compare* provided in the tool.

FINDINGS AND DISCUSSIONS

The first part of this section presents the overall vocabulary profile of the graded readers, level 1 up to 6. It shows the proportions of the vocabulary frequency classifications in percentages. The second part presents the negative vocabulary profile of K-1, K-2, AWL and the Off-list word. Then, the last part shows the comparison of two graded readers considering the result of the vocabulary frequency with an extreme difference.

A. Overall vocabulary profile of graded readers, level 1 up to 6

Table 4. Cumulative percentage of overall vocabulary profile of graded readers

| Freq.Level | Families (%) | Types (%) | Tokens (%) | Cumul.token (%) |
|------------|-----------------|------------------|-------------------|-----------------|
| K-1 Words | 846 (55.19%) | 1850 (51.10%) | 41784 (88.50%) | 88.50% |
| K-2 Words | 498 (32.49%) | 823 (22.73%) | 3077 (6.52%) | 95.02% |

| | | | | |
|--------------------------|----------|-----------------|-----------------|---------|
| AWL | 189 | 297 | 817 | |
| (570 FAMS TOT: 2.570) | (12.33%) | (8.20%) | (1.73%) | 96.75% |
| Off-List Words | ? | 650 (17.96%) | 1536 (3.25%) | 100.00% |
| Total | 1533+ ? | 3620 (100%) | 47214 (100%) | - |

According to the data presented in the table, the analysis of vocabulary profile used in graded readers was based on word *families*, *types* and *token*. The term *families* or *word families* refers to the head of word. For example, *confuse* is the head word of *confused* or *confusing*. The term *type*, is the variation of form words in the same families. Therefore, *confuse*, *confused* and *confusing* are considered as the same *type*. While *figure* and *create* are considered as the different types. Last is *token*; the total number of words in the input text. For instance if the result of analysis in the tool shows, *leagues* [5], *prince* [20], *acid* [4] and *wallet* [3], the number of token is 32. That is why *token* shows the highest number in the table presented above.

The cumulative percentage in the table shows that the use of K-1 words was 88.50% with the additional of K-2 words (6.52%), so the total of K-1 and K-2 was 95.02%. According to Hu and Nation (2000) cited in Schmitt, Jiang and Grabe (2011), it takes 98%-99% of word knowledge for the purpose of reading for pleasure. It appears that the cumulative percentage of K-1 and K-2 words, which are considered as the most frequently used words, is less than the expected word

knowledge. Meanwhile, the AWL percentage was 1.73%. The AWL which commonly appear in the academic texts, were also found in the graded readers. The total academic words found were 817 words. Then, the question that should be considered is how the AWL words should be taught in this course. The last category is the off-list words which covered 3.25%. By having around 1536 words included in the off list words, the teacher and the students should consider how to handle those words. There might be some words which are useful for the students to know.

B. Negative Vocabulary Profiles of The Graded Readers

In this section, the negative words of K-1, K-2, AWL and the Off-list word will be presented. Negative words or negative vocabulary is the words or vocabulary which are not found in the graded readers. The experts have created the list of K-1, K-2 and K-3 (AWL) words which is known as the New General Service List (NGSL). Then, the negative words which were not found in the graded readers were analyzed by comparing the words on the sample texts and the NGSL. The use of NGSL helped the researcher to know what words included in graded readers and what were not.

1. The negative VP of K-1

Below are the analysis of negative words of K-1 found in the samples of the graded readers. Here is the summary of the analysis of K-1.

| | |
|-------------------------|----------------|
| K-1 total word families | : 964 |
| K-1 families in input | : 835 (86.62%) |

K-1 families not in input : 130 (13.49%)

The percentage shows above is different from the cumulative percentage presented in the first table (overall vocabulary profiles) because the percentages here refer to the word families (head word). For further analysis, the total percentage of K-1 words families found in the text (in input) was 86.62% which means, the books used around 86.62 % words in the NGSL. Then, the remaining 13.49% was the words which were not found in the texts (not in input) based on the NGSL. Some words which were not found in the text (negative words) were listed below. The complete word list has been attached in the Appendix A.

| | | | |
|-------------|-----------|----------|-------------|
| ACCOUNT | COAL | DISTRICT | INDUSTRY |
| ACCOUNTABLE | COIN | EARTH | INFLUENCE |
| ACTOR | COLONY | EAST | IRON |
| ACTRESS | COMMITTEE | ELECT | JOY |
| ADOPT | CONTENT | EMPIRE | JUSTICE |
| APPLY | CORN | ENEMY | LATTER |
| ARISE | COST | EXPRESS | LAUGHTER |
| ARMY | COTTON | FELLOW | LITERATURE |
| ASSOCIATE | COUNCIL | FLOW | LORD |
| BATTLE | COUNT | FOREST | LOSS |
| BRANCH | CROWN | FORTH | MANUFACTURE |
| BRIDGE | DECLARE | GAIN | MASTER |
| CAPITAL | DEFEAT | GATHER | MEMBER |

| | | | |
|-----------|------------|--------|-----------|
| CASTLE | DEPARTMENT | HEAVEN | MERE |
| CHARACTER | DESERT | HONOUR | MILE |
| MINER | MONDAY | MOTOR | NECESSITY |
| MINISTER | MOON | MRS | NEIGHBOUR |
| MISTER | MORAL | NATIVE | NOBLE |

2. The negative VP of K-2

Then the negative VP of K-2 found in the samples of the graded readers were presented below.

| | |
|---------------------------|----------------|
| K-2 total word families | : 986 |
| K-2 families in input | : 498 (50.51%) |
| K-2 families not in input | : 489 (49.59%) |

Based on the NGSL the words found to be used in graded readers (in input) was 50.51%. Then the remaining 49.59% was the words which were not found (not in input) in the sample texts. Some words which were not found in the text (negative words) were listed below. The complete word list has been attached in the Appendix B.

| | | | |
|------------|----------|---------|--------|
| ABSOLUTELY | ASTONISH | BILLION | CAPE |
| ACCUSTOM | AVENUE | BIRTH | CART |
| ACHE | AXE | BLADE | CATTLE |

| | | | |
|-------------|---------|--------|----------|
| ADVERTISE | BAGGAGE | BLESS | CAVE |
| AEROPLANE | BAKE | BLIND | CEREMONY |
| AFFORD | BALANCE | BOAST | CHALK |
| AGRICULTURE | BAND | BOLD | CHEAP |
| AIM | BARBER | BORDER | CHEAT |
| AIRPLANE | BARELY | BOUND | CHIMNEY |
| ALTOGETHER | BARREL | BOWL | CLAY |

3. The negative VP of AWL (K-3)

The last was the negative VP of K-3. The following analysis was showed that the words found in the sample texts (in input) based on NGSL was 33.22%. Meanwhile the words which were not found (not in input) in the graded readers was 66.96%.

| | |
|---------------------------|----------------|
| K-3 total word families | : 569 |
| K-3 families in input | : 189 (33.22%) |
| K-3 families not in input | : 381 (66.96%) |

Below was the list of some negative words found in the graded readers as the sample texts. The complete word list has been attached in the Appendix C.

| | | | |
|-------------|------------|--------------|------------|
| ABSTRACT | ANNUAL | CATEGORY | COMPATIBLE |
| ACCESS | ANTICIPATE | CEASE | COMPILE |
| ACCUMULATE | APPEND | CHART | COMPLEMENT |
| ACKNOWLEDGE | ARBITRARY | CIRCUMSTANCE | COMPONENT |

ADAPT ASPECT CITE COMPOUND
 ADEQUATE ASSEMBLE CIVIL COMPREHENSIVE
 ADJACENT ASSIGN CLARIFY COMPRISE
 ADMINISTRATE ATTAIN CLASSIC CONCEIVE
 ADVOCATE ATTRIBUTE COMMODITY CONCURRENT
 AGGREGATE CAPACITY COMMUNITY CONDUCT

C. The block frequency output of off-list words

Below is the list of *off-list* words. The term *off-list* means all the words which are not included in the *high-frequency words (K-1 and K-2)*, *academic words (AWL)* and *technical words*. Here is the block frequency output of *off-list* words. The complete list has been put in Appendix D.

Table 5. Block frequency output of off-list words

| RANK | FREQ | COVERAGE | | WORD |
|------|------|----------|------------|------------|
| | | individ | cumulative | |
| 1. | 35 | 2.32% | 2.32% | LABORATORY |
| 2. | 29 | 1.92% | 4.24% | HELIX |
| 3. | 29 | 1.92% | 6.16% | PRINCE |
| 4. | 26 | 1.72% | 7.88% | HYDROGEN |
| 5. | 23 | 1.52% | 9.40% | PRINCESS |
| 6. | 18 | 1.19% | 10.59% | MAGIC |
| 7. | 17 | 1.13% | 11.72% | GUY |
| 8. | 16 | 1.06% | 12.78% | ATOMS |
| 9. | 15 | 0.99% | 13.77% | GIANT |
| 10. | 15 | 0.99% | 14.76% | HELICAL |

| | | | | |
|-----|----|-------|--------|------------|
| 11. | 14 | 0.93% | 15.69% | CHEMISTRY |
| 12. | 13 | 0.86% | 16.55% | MOLECULE |
| 13. | 13 | 0.86% | 17.41% | PHOSPHATE |
| 14. | 12 | 0.79% | 18.20% | GENE |
| 15. | 12 | 0.79% | 18.99% | PROTEIN |
| 16. | 11 | 0.73% | 19.72% | ALPHA |
| 17. | 11 | 0.73% | 20.45% | GENETICS |
| 18. | 11 | 0.73% | 21.18% | MOLECULES |
| 19. | 11 | 0.73% | 21.91% | SUBMARINE |
| 20. | 11 | 0.73% | 22.64% | UNDERWATER |

The *off-list* words found do not mean that the words are no longer important. There may be words in *off-list* that might be useful for the students. The list above has been divided per 10 words from the frequently appear in the text up to the less frequent. Therefore, knowing those words might be useful for the students when they were reading the graded readers. For teaching could select words that the students might have to know such as words that appear in daily conversation; *guy, prince, giant, underwater, etc.*

D. The comparison of vocabulary frequency across levels of graded readers

Table 6 shows the comparison of the vocabulary frequencies of the graded readers, level 1 up to 6.

Table 6. The comparison of VP across levels graded readers

| | K-1 words | K-2 words | AWL | Off-list words |
|---------|-----------|-----------|-------|----------------|
| Level 1 | 92.96% | 3.67% | 0.35% | 3.01% |
| Level 2 | 93.19% | 4.80% | 0.00% | 2.01% |

| | | | | |
|---------|--------|-------|-------|-------|
| Level 3 | 90.37% | 7.07% | 0.50% | 2.06% |
| Level 4 | 90.41% | 6.08% | 0.92% | 2.59% |
| Level 5 | 90.86% | 6.11% | 0.97% | 2.06% |
| Level 6 | 84.70% | 6.90% | 3.69% | 4.71% |

As seen in the Table 6, an interesting finding was an extreme percentage appeared in the AWL frequency level of graded reader level 2 which was 0.00%. From 6 graded readers, level 2 taken was the book entitled *Five Famous Fairy Tales*. It told about nonsense stories intended purely for reading for pleasure. Therefore, the appearance of AWL was few in this level. Yet, the use of K-1 words for this level was the highest; 93.19%. The example words used were *beautiful, brother, country, loved, etc.*

Meanwhile, the frequency of AWL in the graded reader level 6 was the highest; 3.69%. The book's title was *The Double Helix*. The book told about how DNA was created. So there were many words used in biology, such as *protein, atomic, theory, etc.* The example of AWL were *abandon, academy, accomodation, achieve, etc.* While the sample of level 6 taken showed the highest percentage of AWL, the K-1 words had the lowest percentage among other readers; 84.70%. There were many difficult words to understand considering that these books are intended for the first semester of college students. Yet, it was acceptable because the purpose of reading the graded readers was also to increase students' ability in their vocabulary learning.

E. The text comparison of graded readers, level 2 and 6

This is the last part of the analysis. Based on the difference of AWL percentage between level 2 and 6, then the two graded readers have been analyzed in more detail. The texts were copied into the *Text Lex Compare*. The result shows the *token recycling index* of the graded readers being compared. *Recycling index* is the ratio between words that are shared by the two input texts and the total number of words in the second text. Based on that data, the shared words and the unshared words or known as the unique words between two texts were showed. Table 7 below is a small part of the result of the analysis. The complete table has been put in Appendix E.

TOKEN Recycling Index: (10866 repeated tokens : 16081 tokens in new text) = 67.57%

FAMILIES Recycling Index: (325 repeated families : 1459 families in new text) = 22.28%

Table 7. Token recycling index of graded readers, level 2 and 6

| | | | |
|---|---|---|---|
| <u>Unique to first</u> 750 tokens 123 families | <u>Shared</u> 10866 tokens 325 families | <u>Unique to second</u> 5215 tokens 1134 families | |
| 001. you 93 002. king 47 003. can 41 004. dog 32 005. prince 32 006. queen 26 007. gold 25 008. cry 23 | 001. the 1011 002. be 690 003. to 586 004. i 485 005. of 433 006. this 429 007. he 408 008. a 394 009. have | <u>Freq first (then alpha)</u> 001. base 77 002. model 61 003. however 60 004. structure 54 005. helix 51 006. which 49 | <u>Same list Alpha first</u> 001. abandon 2 002. able 11 003. abroad 1 004. absence 2 005. absent 1 006. absolute 3 007. academy |

| | | | |
|---------------------|-------------------|-----------------------|-----------------------|
| 009. princess 22 | 304 | 007. might 47 | 1 |
| 010. magic 17 | 010. in 287 | 008. possible 44 | 008. accent 1 |
| 011. city 16 | 011. and 266 | 009. chain 39 | 009. accept 15 |
| 012. tree 15 | 012. not 182 | 010. laboratory 36 | 010. accident 1 |
| 013. coat 14 | 013. we 171 | 011. even 32 | 011. accommodate 3 |
| 014. donkey 12 | 014. would 158 | 012. same 32 | 012. accompany 1 |
| 015. catch 11 | 015. for 145 | 013. much 30 | 013. according 1 |
| 016. bird 10 | 016. with 134 | 014. almost 29 | 014. accurate 3 |
| 017. kill 10 | 017. on 127 | 015. molecule 29 | 015. accuse 1 |
| 018. angry 9 | 018. it 116 | 016. pair 29 | 016. achieve 5 |
| 019. blue 9 | 019. at 110 | 017. several 29 | 017. acid 7 |
| 020. die 9 | 020. do 91 | 018. bond 28 | |
| 021. fat 9 | 021. about | 019. fact 28 | |
| 022. giant 9 | | 020. week 28 | |

Table 7 above shows that the words which were only found in the first sample text (unique to first) were 123 words. Meanwhile, the words which were only found in the second sample text (unique to second) were 1134 words. The last was for the term shared which means that the two texts have similar words used, and it had 325 words shared.

For further analysis, the analysis of comparison shows that the token recycling index was 67.57%. It means; as much as 67.57% of words in the two graded readers were similar (shared). Meanwhile, as much as 32.43% of words found in the graded reader level 6 were the new or unique words. Therefore, the comprehensibility of this level was acceptable considering it was the highest level, so it should have a little much of new vocabulary than the other levels.

CONCLUSION

The purpose of this study were to identify the vocabulary profile of the graded readers used in the Faculty of Language and Literature, Satya Wacana

Christian University, to produce lists of vocabulary that were not covered in the graded readers and to produce the token (word) recycling index of the graded readers. The investigation of graded readers, level 1 up to 6 has shown that the word coverage of the graded readers samples were 88.50% for K-1, 6.52% for K-2, 1.73% for AWL and 3.25% for Off-List word.

Based on NGSL, the major finding is that the words not found in the graded readers' were 13.49% of K-1, 49.59% of K-2, and 66.96% of AWL. Besides, this study has also shown that the comparison of level 2 and 6 of graded readers shows that there were some shared words which was 67.75% and unshared words (unique) which was 32.43%.

The results of this study indicate that the use of graded readers in Extensive Reading course was quite appropriate to the course. Although the total of high frequency words (K-1 & K-2) was less than the expected word knowledge for reading for pleasure which was took 98%-99% (Hu & Nation, 2000, as cited in Schmitt, Jiang & Grabe, 2011), but it passed the minimum threshold mention by Nation (2001, p. 144) which was 95.02%. From the teaching perspective it was appropriate to use such reading materials; graded readers, because it contains many new vocabulary wrapped in many fiction stories. From the learning perspective, using graded readers brought less pressure for the learners to learn English as the foreign language. Yet, it was a good suggestion that the learners especially the college students to have a reference list of vocabulary coverage to decide whether a particular word is worth the attention or not.

The last thing being highlighted was the study did not evaluate the technical words used in the graded readers. Therefore, the question raised by this study is “What is the technical word used in the graded readers or another sample text being analyzed?” Then, further studies regarding the use of the technical words in the text would be interesting.

ACKNOWLEDGEMENTS

There were so many people I would like to say my gratitude for. First of all, my greatest gratitude was for Allah SWT who has always granted my wishes and let me finish my thesis on time. My gratitude was also for my parents and my family who have always supported me no matter what happens. I also would like to offer my sincere gratitude to Prof. Dr. Gusti Astika, M. A., for his guidance in reviewing my thesis so I could complete my thesis. Fourth, was to my second reader, Anne Indrayanti Timotius, M. Ed who has helped me to develop my thesis. My last gratitude was intended to Ratih Chandra who has helped me in collecting my data, to all of my friends especially Rachma Maulia, Hannasista, Nathasya who always made me laugh and listen patiently to all of my unnecessary talk whenever I felt depressed. Thank you so much.

REFERENCES

- Alhaysony, M. (2012). Vocabulary discovery strategy used by Saudi EFL students in an intensive english language learning context. *International Journal of Linguistics*. Vol. 4, (2), 518-535. doi: 10.5296/ijl.v4i2.1724.
- Asgari, A., & Mustapha, G. B. (2011). The type of vocabulary learning strategies used by ESL students in University Putra Malaysia. *English Language Teaching*. Vol. 4, (2), 84-90. doi: 10.5539/elt.v4n2p84.
- Fisher, D. & Frey, N. (2014). Content area vocabulary learning. *The Reading Teacher*. Vol. 67, (8), 594-599. doi: 0.1002/trtr.1258.
- Fuller, A. W. (n.d). Unpublished thesis. *Vocabulary analysis of a graded reader*. Colorado State University.
- Grabe, W. (2009). *Reading in a Second Language Moving from Theory to Practice*. New York : Cambridge University Press.
- Graves, D. (2005). A Publication of the Jane Austen Society of North America. *Vocabulary Profiles of Letters and Novels of Jane Austen and her Contemporaries*, 26 (1). Retrieved March 03, 2016, from <http://www.jasna.org/index.html>.
- Hu, H. M. (2013). The effects of word frequency and contextual types on vocabulary acquisition from Extensive Reading: a case study. *Journal of Language Teaching and Research*. Vol. 4, (3), 487-495. doi: 10.4304/jltr.4.3.487-495.
- Jusuf, Y. A. (2014). *Vocabulary profile in English textbooks used in Satya Wacana junior high school*. Unpublished thesis. Faculty of Language and Literature, Satya Wacana Christian Univeristy.

- Kurniawati, A. (2015). Syllabus Extensive Reading semester 1/2015-2016. Faculty of Language and Literature, Satya Wacana Christian University.
- Laufer, B., & Kalovski, G. C. (2010). Lexical threshold revisited: lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22 (1), 15-30.
- Li, L., & Zhang, X. (2009). Teaching vocabulary in Chinese Universities; what and how?. *Asian Social Science*, 5 (7), 126-129.
- Martinez, R., & Murphy, V. A. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *Tesol Quartely*, 45 (2), 267-290.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening. *The Canadian Modern Language Review/ La Revue Canadienne Des Langues Vivantes*, 63 (1), 59-82.
- Nation, I. S. P. (2008). *Teaching Vocabulary Strategies and Techniques*. Boston : Heinle, Cengage Learning.
- Noor, M. N., & Amir, Z. (2009). Exploring the vocabulary learning strategies of EFL learners. *Language and Culture: Creating and Fostering Global Communities. 7th International Confronce by the School of Studies and Linguistics Faculty of Social Sciences and Humanities*, 313-327. Retrieved March 10, 2016, from <https://www.researchgate.net/>
- Salaria, N. (2012). Meaning of the term descriptive survey research method. *International Journal of Transformations in Business Management*, 1 (6), 1-7.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*. Vol. 95, (1), 26-43. doi: 10.1111/j.1540-4781.2011.01146.x0026-7902/11/26-43.

Wan-a-rom, U. (2008). Comparing the vocabulary of different graded reading schemes. *Reading in a Foreign Language*, 20 (1), 43-69.

Webb, S., & Chang, A. C-S. (2015). Second language vocabulary learning through Extensive Reading with audio support: how do frequency and distribution of occurrence affect learning?. *Language Teaching Research*. Vol. 19, (6), 667-686. doi: 10.1177/1362168814559800.

