INTRODUCTION

There are now many questions asked about textbooks for the teaching of languages. While research shows that the 2000 high frequency words in the English language cover 87% of running words in a text, material developers are wondering how textbooks are crowded with words (the Form 1 book in this study used 4730 different words!) while missing out on many words from the 2000 high frequency word list. On the other hand, while words from the 2000 high frequency list do appear, they are usually not repeated and this is against the perspective of language learning as Thornbury (2002) claims that students remember words which are repeated seven times over spaced intervals.

In previous research using corpus-based procedures, the researchers worked on dead textbooks from the previous cycle and the forensic study (Mukundan and Aziz, 2008) revealed for the first time weaknesses in the development of textbooks, chief amongst which were in the area of lexical loading and distribution patterns. In this present study, the researchers attempted to analyze the load and distribution efficiency of the 2000 high frequency words in the textbooks.
Second language learners need to know the 3,000 high frequency words of the language (Waring and Nation, 1997), as knowing these words enables them to begin reading authentic texts (Nation, 1990; Schmitt, Schmitt and Clapham, 2001). However, knowledge of the first 2,000 most frequent words in the language is sufficient enough to allow learners to access to approximately 87% of any ordinary text (Nation, 1990).

High frequency words are generally the most familiar words and they make up the majority of tokens in any discourse (Schmitt, 2000). Higher frequency words tend to be short and learners are likely to meet these words more frequently (Thornbury, 2002). It is important for language learners to know these words in order to function effectively in English. Once the high frequency words are acquired, learners can then independently learn the less frequent words through a combination of intentional and incidental learning (Schmitt, 2000). Using these words, learners are also able to “make accurate guesses about meanings of the remaining less frequent words which are likely to be unknown” (Schmitt, 2000).

The most frequent 2,000 headwords from West’s General Service List (GSL) seem to be the most cited initial goal for second language learners as it accounts for at least 80%-85% of the words on any page of any book on any subject matter (Waring and Nation, 1997; Nation and Newton, 1997; Schmitt, 2000). The GSL is “a set of 2,000 words selected to be the greatest general service to learners of English” (http://jbaumann.com). However, the list does not contain the most common 2000 words although during the selection, frequency count was one of the factors considered (Nation, 2004). Each of these words is headword representing a word family, which was loosely defined in West (http://jbaumann.com). This list has served as the basis for graded readers and other material for many years (Nation, 1993).

**OBJECTIVES OF THE STUDY**
The main aim of the study was to investigate patterns of loading, distribution, and repetiton of the 2000 high frequency words in Malaysian Secondary School English Language Textbooks (Forms 1-5). The study also attempted to identify words in the 2000 high frequency list which are not used in the entire English Language textbooks.

**RESEARCH QUESTIONS**
The questions this study seeks to answer are as follows:

- What are the general characteristics of vocabulary loading in Malaysian Secondary School English Language textbooks (Forms 1-5)?
- To what extent are words in the 2000 high frequency list covered in Malaysian Secondary School English Language textbooks (Forms 1-5)?
- How often are words in the 2000 high frequency list being recycled in Malaysian Secondary School English Language textbooks (Forms 1-5)?
- Which words in the 2000 high frequency list are not being used in the entire Malaysian Secondary School English Language textbooks (Forms 1-5)?
- How often are randomly selected words from the 2000 high frequency list being distributed within and across the entire Malaysian Secondary School English Language textbooks (Forms 1-5)?

**SIGNIFICANCE OF THE STUDY**
This study will benefit textbook developers in the following ways:

- Material developers will develop a greater awareness of vocabulary load and distribution patterns and have greater control of words and how they should be presented using concordance software. This will benefit the writing of new textbooks in the future.
- Materials developers will know that books written in an ad hoc manner will not even
take into consideration some important aspects such as the existence of the 2000 high frequency word list. Future textbook developers will also have a greater awareness with regards to this.

METHODOLOGY
The study employed a purposive sampling method. The English language textbooks for Form 1 to Form 5 from the Klang Valley schools (Central Zone) were chosen for the study. These books are from the newest cycle of textbooks which were introduced in 2003, and are still used at the present time. All pages in the textbooks were scanned and saved as image files. These image files were converted to text files to enable WordSmith 4.0 to analyze the vocabulary used in these textbooks. Two WordSmith tools, WordList and Concord tools, were mainly used in this study. The study used West’s GSL as a reference for the 2000 most frequent words in analyzing the patterns of loading, distribution, and repetition of these words in Malaysian Secondary School English Language Textbooks (Forms 1-5).

RESULTS AND DISCUSSION
The presentation and discussion in this section are based on the research questions stated earlier. The first part deals with the general statistics of vocabulary loading in Malaysian Secondary School English language textbooks for Forms 1-5. The second part deals with the coverage of the 2000 high frequency words in Malaysian Secondary School English Language textbooks (Forms 1-5). The third part deals with the repetition of the 2000 high frequency words in Malaysian Secondary School English Language textbooks (Forms 1-5), while the fourth part deals with words in the 2000 high frequency list which are not used in the entire Malaysian Secondary School English Language textbooks (Forms 1-5). The final part deals with the distribution of 10 randomly selected words from the 2000 high frequency list within and across the entire Malaysian Secondary School English Language textbooks (Forms 1-5).

Part 1: General Characteristics of Vocabulary Loading in Malaysian Secondary School English Language Textbooks (Forms 1-5)
Table 1 shows the total number of running words (tokens) and the total number of different words (types) found in the textbooks. There is a gradual increase in the total number of tokens from the Form 1 textbook to Form 5 textbook. There is also a gradual increase in the total number of types found in the Forms 1-3 textbooks. However, a notable increase is seen in the total number of types in the Form 4 textbook compared to the Form 3 textbook. The Standardized Type/Token Ratio (STTR) was used to measure the density level of textbooks. Textbooks with higher percentage of STTR indicate that the textbooks have more types being introduced for every 1000 tokens in the textbooks. The results presented in Table 1 show that Form 5 textbook has the highest density level compared to the other four English language textbooks. At this level, students are assumed to be ready to handle a larger number of words, therefore more types are introduced. However, the Form 1 textbook does not have the lowest density level. In fact, its density level is higher than the textbooks for Forms 2 and 3. The finding also shows that the Form 1 textbook introduces more types than the Form 2 and Form 3 textbooks, making it relatively more difficult than the other two textbooks in terms of the vocabulary load in the textbooks. Form 1 students have to handle more vocabulary load than the Form 2 and Form 3 students, which should not be the case. Apparently, the Form 2 textbook has the lowest density level, making it the least difficult textbook in terms of the vocabulary load in the textbook.

Part 2: The Coverage of the 2000 High Frequency Words in Malaysian Secondary School English Language Textbooks (Forms 1-5)
A passive knowledge of these words provides familiarity of nearly nine out of ten words in most written texts (Thornbury, 2002), while Nation (1990) claims that the 2000 high frequency
Words cover 87% of the running words in a text. Results of the analysis show that 27.0% of the types in Form 1 textbook (1278 out of 4730 types) appear in the 2000 high frequency list. This covers 63.9% of the words in the list. Meanwhile, 26.7% of the types in Form 2 textbook (1264 out of 4738 types) appear in the 2000 high frequency list. This covers 63.2% of the words in the list. 25.8% of the types in Form 3 textbook (1369 out of 5308 types) appear in the 2000 high frequency list. It covers 68.5% of the words in the list. 19.9% of the types in Form 4 textbook (1549 out of 7788 types) appears in the 2000 high frequency list. This covers 77.5% of the words in the list. 19.4% of the types in Form 5 textbook (1553 out of 7994 types) appear in the 2000 high frequency list. It covers 77.7% of the words in the list. 12.6% of the types in the entire secondary school textbook (1861 out of 14732 types) appear in the 2000 high frequency list. It covers 93.1% of the words in the list. Having been exposed to 93.1% of the 2000 high frequency words at the end of Form 5, students are assumed to be able to comprehend 81.0% of texts on any general topics.

Table 2 shows the number and percentages of the 2000 high frequency words in Malaysian Secondary School English Language textbooks (Forms 1-5). It can be concluded that the textbooks have prepared the students with the knowledge of all the 2000 high frequency words in the high frequency list as this would show awareness on the part of material developers that learners who leave school after five years of secondary education have to be exposed to the entire list of 2000 high frequency words.

### TABLE 1
The total number of tokens and types in Malaysian Secondary School English Language Textbooks (Forms 1-5)

<table>
<thead>
<tr>
<th>Form 1</th>
<th>Form 2</th>
<th>Form 3</th>
<th>Form 4</th>
<th>Form 5</th>
<th>All Textbooks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens in the Textbooks</td>
<td>45,105</td>
<td>49,497</td>
<td>60,038</td>
<td>75,154</td>
<td>81,420</td>
</tr>
<tr>
<td>Types in the Textbooks</td>
<td>4,730</td>
<td>4,738</td>
<td>5,309</td>
<td>7,788</td>
<td>7,994</td>
</tr>
<tr>
<td>Standardized Type/Token Ratio (STTR)</td>
<td>40.11</td>
<td>39.67</td>
<td>39.94</td>
<td>42.58</td>
<td>42.21</td>
</tr>
</tbody>
</table>

### TABLE 2
The number and percentage of the 2000 high frequency words in Malaysian Secondary School English Language Textbooks (Forms 1-5)

<table>
<thead>
<tr>
<th>Form 1</th>
<th>Form 2</th>
<th>Form 3</th>
<th>Form 4</th>
<th>Form 5</th>
<th>All Textbooks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of types in the textbooks</td>
<td>4730</td>
<td>4738</td>
<td>5309</td>
<td>7788</td>
<td>7994</td>
</tr>
<tr>
<td>Number of words found in the 2000 high frequency word list</td>
<td>1278</td>
<td>1264</td>
<td>1369</td>
<td>1549</td>
<td>1553</td>
</tr>
<tr>
<td>Percentage of words appearing in the 2000 high frequency word list</td>
<td>63.9%</td>
<td>63.2%</td>
<td>68.5%</td>
<td>77.5%</td>
<td>77.7%</td>
</tr>
<tr>
<td>Percentage of types in textbook appearing in the 2000 high frequency words</td>
<td>27.0%</td>
<td>26.7%</td>
<td>25.8%</td>
<td>19.9%</td>
<td>19.4%</td>
</tr>
</tbody>
</table>

Part 3: Repetition of the 2000 High-frequency Words in Malaysian Secondary School English Language Textbooks (Forms 1-5)

Thornbury (2002) claims that students remember words which are repeated at least seven times over a spaced interval. The results show that...
71.9% of the words in the list (1438) are repeated at least seven times in the entire textbooks, 21.2% of the words (423 words) are recycled less than seven times, while 6.9% of the words (139 words) are not found in any of the textbooks. Table 3 displays the number of words being repeated in the entire set of textbooks.

These results also show that although the coverage of the words in the high frequency list is 93.1% (1861 out of 2000 words), the textbooks only recycle 71.9% of the words in the list effectively. Table 4 displays the list of words repeated less than seven times in the entire set of textbooks.

Part 4: Words in the 2000 High Frequency List Which Are Not Used in the Entire Malaysian Secondary School English Language Textbooks (Forms 1-5)

About 6.9% of the words in the high frequency list (139 words) are not used in the textbooks. Table 5 presents the words in the 2000 high frequency list which are not used in any of the textbooks.

However, derivatives and inflections of some of the words in Table 4 are found in the textbooks. Although these words are not found in the textbooks, teachers may have taught these words in order to teach the derivative and inflectional forms of the words. Table 6 shows words (their derivatives and inflectional forms included) which are not used in any of the textbooks. There are 45 (2.3%) words (including their derivatives and inflections), in the 2000 high frequency list, which are not used in the entire set of Malaysian Secondary School English Textbooks (Forms 1-5).

Part 5: Distribution of 10 Randomly Selected Words from the 2000 High Frequency List within and across the Entire Malaysian Secondary School English Language Textbooks (Forms 1-5)

Ten words were randomly selected using random number generator at GraphPad Software (www.graphpad.com/quickcalc/randomN1.cfm). These words are used as examples to illustrate the distribution patterns of words within and across textbooks and not as examples of problematic words for students to learn. The words and their ranking in the 2000 high frequency list are presented in Table 7. The word “much” is the 94th highest frequency word in the list, while “yellow” is ranked 1277th rank in the 2000 high frequency list. The Concord Tool of WordSmith 4.0 was used to plot the distribution patterns of the selected words.

Table 8 shows how frequent the words are being used in the textbooks. As a whole, students encounter these words more than seven times within the five years of schooling. There are, however, some problems with regards to the repetition of these words across the entire period of secondary school education. For instance, the word ‘art’ which is ranked 277th in the 2000 high frequency list, is rarely used in the Forms 1, 2, 3, and 5 textbooks. The word is recycled well only in the Form 4 textbook. The word ‘yellow’ is ranked 1277th in the list. This word has not been recycled well in the entire set of textbooks. In fact, it is not being used at all in the Form 3 textbook. These results show that although more than 90% of the 2000 high frequency words are covered in the textbooks, not all the words are recycled effectively within the textbooks to ensure learning.

<table>
<thead>
<tr>
<th>No. of Occurrence in Form 1-5 Textbooks (times)</th>
<th>No. of Words</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 times and more</td>
<td>1438 words</td>
<td>71.9%</td>
</tr>
<tr>
<td>Less than 7 times</td>
<td>423 words</td>
<td>21.2%</td>
</tr>
<tr>
<td>Zero</td>
<td>139 words</td>
<td>6.9%</td>
</tr>
</tbody>
</table>

TABLE 3
The number of words that are repeated fewer than seven times in the entire set of Malaysian Secondary School English Language Textbooks
<table>
<thead>
<tr>
<th>abroad</th>
<th>cheer</th>
<th>eager</th>
<th>gray</th>
<th>miserable</th>
<th>Prove</th>
<th>sincere</th>
<th>tend</th>
</tr>
</thead>
<tbody>
<tr>
<td>absolute</td>
<td>chief</td>
<td>ease</td>
<td>grind</td>
<td>misery</td>
<td>pump</td>
<td>skirt</td>
<td>tent</td>
</tr>
<tr>
<td>actor</td>
<td>Christmas</td>
<td>eastern</td>
<td>guard</td>
<td>modest</td>
<td>punish</td>
<td>slavery</td>
<td>thorough</td>
</tr>
<tr>
<td>admission</td>
<td>church</td>
<td>efficiency</td>
<td>guilt</td>
<td>motion</td>
<td>pupil</td>
<td>slide</td>
<td>thread</td>
</tr>
<tr>
<td>admit</td>
<td>circular</td>
<td>elder</td>
<td>hang</td>
<td>multiply</td>
<td>pure</td>
<td>slight</td>
<td>threaten</td>
</tr>
<tr>
<td>advance</td>
<td>clay</td>
<td>elect</td>
<td>haste</td>
<td>murder</td>
<td>push</td>
<td>slip</td>
<td>thunder</td>
</tr>
<tr>
<td>advertise</td>
<td>clever</td>
<td>election</td>
<td>heaven</td>
<td>musician</td>
<td>qualify</td>
<td>slope</td>
<td>ticket</td>
</tr>
<tr>
<td>affair</td>
<td>coal</td>
<td>enemy</td>
<td>hesitate</td>
<td>mystery</td>
<td>quarral</td>
<td>snow</td>
<td>tie</td>
</tr>
<tr>
<td>agent</td>
<td>coat</td>
<td>entertain</td>
<td>holy</td>
<td>nail</td>
<td>quarter</td>
<td>soap</td>
<td>tight</td>
</tr>
<tr>
<td>airplane</td>
<td>commerce</td>
<td>envelope</td>
<td>hunt</td>
<td>neglect</td>
<td>queen</td>
<td>solemn</td>
<td>toe</td>
</tr>
<tr>
<td>alike</td>
<td>compete</td>
<td>envy</td>
<td>ideal</td>
<td>nest</td>
<td>railroad</td>
<td>somehow</td>
<td>ton</td>
</tr>
<tr>
<td>allowance</td>
<td>completion</td>
<td>equal</td>
<td>imaginary</td>
<td>noble</td>
<td>recognition</td>
<td>soul</td>
<td>tool</td>
</tr>
<tr>
<td>altogether</td>
<td>confuse</td>
<td>essence</td>
<td>inch</td>
<td>northern</td>
<td>reflection</td>
<td>soup</td>
<td>toward</td>
</tr>
<tr>
<td>amuse</td>
<td>confusion</td>
<td>everybody</td>
<td>inquire</td>
<td>nowhere</td>
<td>refresh</td>
<td>spare</td>
<td>towel</td>
</tr>
<tr>
<td>annoy</td>
<td>connection</td>
<td>examine</td>
<td>insect</td>
<td>operate</td>
<td>relation</td>
<td>spell</td>
<td>trade</td>
</tr>
<tr>
<td>apart</td>
<td>conscience</td>
<td>exception</td>
<td>instant</td>
<td>operator</td>
<td>relieve</td>
<td>spin</td>
<td>translate</td>
</tr>
<tr>
<td>appoint</td>
<td>convenience</td>
<td>excessive</td>
<td>instrument</td>
<td>oppose</td>
<td>remark</td>
<td>spirit</td>
<td>translation</td>
</tr>
<tr>
<td>approve</td>
<td>corn</td>
<td>excite</td>
<td>insurance</td>
<td>opposition</td>
<td>remedy</td>
<td>spirit</td>
<td>trap</td>
</tr>
<tr>
<td>arise</td>
<td>cottage</td>
<td>excuse</td>
<td>intention</td>
<td>organ</td>
<td>representative</td>
<td>spirit</td>
<td>treasury</td>
</tr>
<tr>
<td>arrest</td>
<td>cow</td>
<td>expense</td>
<td>interference</td>
<td>ought</td>
<td>reproduce</td>
<td>stair</td>
<td>tremble</td>
</tr>
<tr>
<td>arrow</td>
<td>crack</td>
<td>explosion</td>
<td>interrupt</td>
<td>package</td>
<td>reputation</td>
<td>stamp</td>
<td>trial</td>
</tr>
<tr>
<td>attempt</td>
<td>crash</td>
<td>extend</td>
<td>jaw</td>
<td>pad</td>
<td>reserve</td>
<td>steady</td>
<td>tribe</td>
</tr>
<tr>
<td>attraction</td>
<td>creep</td>
<td>extension</td>
<td>joint</td>
<td>pan</td>
<td>resign</td>
<td>steam</td>
<td>trick</td>
</tr>
<tr>
<td>autumn</td>
<td>criminal</td>
<td>extensive</td>
<td>jump</td>
<td>parent</td>
<td>resist</td>
<td>steep</td>
<td>tune</td>
</tr>
<tr>
<td>avenue</td>
<td>crop</td>
<td>extraordinary</td>
<td>justice</td>
<td>passenger</td>
<td>retire</td>
<td>steer</td>
<td>twist</td>
</tr>
<tr>
<td>awake</td>
<td>crown</td>
<td>fade</td>
<td>knee</td>
<td>pattern</td>
<td>reward</td>
<td>stem</td>
<td>union</td>
</tr>
<tr>
<td>axe</td>
<td>curtain</td>
<td>faint</td>
<td>knock</td>
<td>pause</td>
<td>roar</td>
<td>stir</td>
<td>unite</td>
</tr>
<tr>
<td>bargain</td>
<td>curve</td>
<td>faith</td>
<td>ladder</td>
<td>permanent</td>
<td>roast</td>
<td>stock</td>
<td>universe</td>
</tr>
<tr>
<td>bathe</td>
<td>custom</td>
<td>fan</td>
<td>lamp</td>
<td>pet</td>
<td>rob</td>
<td>stomach</td>
<td>urge</td>
</tr>
<tr>
<td>beam</td>
<td>customary</td>
<td>fancy</td>
<td>lean</td>
<td>pink</td>
<td>root</td>
<td>stove</td>
<td>urgent</td>
</tr>
<tr>
<td>beard</td>
<td>damp</td>
<td>fasten</td>
<td>lessen</td>
<td>pipe</td>
<td>royal</td>
<td>straw</td>
<td>veil</td>
</tr>
<tr>
<td>bedroom</td>
<td>dare</td>
<td>fault</td>
<td>lid</td>
<td>pity</td>
<td>rub</td>
<td>strengthen</td>
<td>vessel</td>
</tr>
<tr>
<td>beg</td>
<td>debt</td>
<td>feather</td>
<td>lip</td>
<td>plain</td>
<td>ruin</td>
<td>stretch</td>
<td>violence</td>
</tr>
<tr>
<td>bend</td>
<td>declare</td>
<td>fence</td>
<td>load</td>
<td>pleasure</td>
<td>rust</td>
<td>strike</td>
<td>virtue</td>
</tr>
<tr>
<td>bind</td>
<td>defeat</td>
<td>fix</td>
<td>loan</td>
<td>polish</td>
<td>sacred</td>
<td>stroke</td>
<td>voyage</td>
</tr>
<tr>
<td>bite</td>
<td>delay</td>
<td>flame</td>
<td>lock</td>
<td>political</td>
<td>sake</td>
<td>stuff</td>
<td>wage</td>
</tr>
<tr>
<td>blame</td>
<td>delight</td>
<td>flesh</td>
<td>loose</td>
<td>politician</td>
<td>salary</td>
<td>stupid</td>
<td>wait</td>
</tr>
<tr>
<td>bleed</td>
<td>deliver</td>
<td>float</td>
<td>lord</td>
<td>politics</td>
<td>salesman</td>
<td>substance</td>
<td>wander</td>
</tr>
<tr>
<td>boast</td>
<td>dependence</td>
<td>fold</td>
<td>lump</td>
<td>possess</td>
<td>scarce</td>
<td>suit</td>
<td>warmth</td>
</tr>
<tr>
<td>border</td>
<td>descend</td>
<td>fool</td>
<td>male</td>
<td>postpone</td>
<td>scenery</td>
<td>summer</td>
<td>wax</td>
</tr>
<tr>
<td>branch</td>
<td>deserve</td>
<td>forgive</td>
<td>mankind</td>
<td>pound</td>
<td>scratch</td>
<td>supper</td>
<td>weapon</td>
</tr>
<tr>
<td>brass</td>
<td>destructive</td>
<td>fork</td>
<td>manufacture</td>
<td>powder</td>
<td>screw</td>
<td>surround</td>
<td>west</td>
</tr>
<tr>
<td>brick</td>
<td>dig</td>
<td>forth</td>
<td>marriage</td>
<td>praise</td>
<td>seed</td>
<td>suspect</td>
<td>whip</td>
</tr>
<tr>
<td>broad</td>
<td>dip</td>
<td>fortune</td>
<td>mass</td>
<td>pray</td>
<td>seize</td>
<td>swallow</td>
<td>whistle</td>
</tr>
<tr>
<td>bunch</td>
<td>disappoint</td>
<td>frequency</td>
<td>mechanic</td>
<td>preference</td>
<td>seldom</td>
<td>sweat</td>
<td>width</td>
</tr>
<tr>
<td>bundle</td>
<td>distant</td>
<td>friendship</td>
<td>mechanism</td>
<td>prejudice</td>
<td>servant</td>
<td>sweep</td>
<td>wine</td>
</tr>
<tr>
<td>businessman</td>
<td>distinguish</td>
<td>frighten</td>
<td>melt</td>
<td>prevention</td>
<td>sew</td>
<td>swell</td>
<td>wing</td>
</tr>
<tr>
<td>cape</td>
<td>disturb</td>
<td>gay</td>
<td>merchant</td>
<td>print</td>
<td>shell</td>
<td>swing</td>
<td>winter</td>
</tr>
<tr>
<td>caution</td>
<td>dive</td>
<td>gentleman</td>
<td>mercy</td>
<td>probable</td>
<td>shield</td>
<td>sympathetic</td>
<td>wisdom</td>
</tr>
<tr>
<td>cautious</td>
<td>dot</td>
<td>govern</td>
<td>mere</td>
<td>profession</td>
<td>shine</td>
<td>sympathy</td>
<td>wreck</td>
</tr>
<tr>
<td>cave</td>
<td>dozen</td>
<td>hay</td>
<td>mild</td>
<td>program</td>
<td>shoot</td>
<td>tax</td>
<td>yard</td>
</tr>
<tr>
<td>century</td>
<td>dull</td>
<td>heap</td>
<td>mile</td>
<td>prompt</td>
<td>shower</td>
<td>taxi</td>
<td>yield</td>
</tr>
<tr>
<td>charm</td>
<td>duty</td>
<td>grave</td>
<td>mineral</td>
<td>proposal</td>
<td>shut</td>
<td>telegraph</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 5
Words in the 2000 high frequency list that are not used in the entire set of Malaysian Secondary School English Textbooks (Forms 1-5)

<table>
<thead>
<tr>
<th>accord</th>
<th>confess</th>
<th>flavor</th>
<th>lodge</th>
<th>scatter</th>
</tr>
</thead>
<tbody>
<tr>
<td>accuse</td>
<td>confession</td>
<td>fond</td>
<td>membership</td>
<td>scrape</td>
</tr>
<tr>
<td>accustom</td>
<td>critic</td>
<td>forbid</td>
<td>moderate</td>
<td>separation</td>
</tr>
<tr>
<td>adopt</td>
<td>crush</td>
<td>freeze</td>
<td>neighbor</td>
<td>shave</td>
</tr>
<tr>
<td>angle</td>
<td>cultivate</td>
<td>funeral</td>
<td>neighborhood</td>
<td>simplicity</td>
</tr>
<tr>
<td>anyhow</td>
<td>curl</td>
<td>furnish</td>
<td>nut</td>
<td>soften</td>
</tr>
<tr>
<td>arch</td>
<td>curse</td>
<td>garage</td>
<td>objection</td>
<td>stain</td>
</tr>
<tr>
<td>astonish</td>
<td>decisive</td>
<td>governor</td>
<td>omit</td>
<td>straighten</td>
</tr>
<tr>
<td>awkward</td>
<td>decrease</td>
<td>grease</td>
<td>owe</td>
<td>suck</td>
</tr>
<tr>
<td>backward</td>
<td>defendant</td>
<td>guest</td>
<td>ownership</td>
<td>suspicion</td>
</tr>
<tr>
<td>barrel</td>
<td>defense</td>
<td>harbor</td>
<td>particle</td>
<td>swear</td>
</tr>
<tr>
<td>basis</td>
<td>depth</td>
<td>hatred</td>
<td>peculiar</td>
<td>tempt</td>
</tr>
<tr>
<td>behavior</td>
<td>desire</td>
<td>heal</td>
<td>plaster</td>
<td>tender</td>
</tr>
<tr>
<td>blade</td>
<td>devil</td>
<td>hire</td>
<td>plow</td>
<td>theater</td>
</tr>
<tr>
<td>boundary</td>
<td>dine</td>
<td>honor</td>
<td>preach</td>
<td>theatrical</td>
</tr>
<tr>
<td>calculation</td>
<td>disapprove</td>
<td>idle</td>
<td>propose</td>
<td>thumb</td>
</tr>
<tr>
<td>carriage</td>
<td>dismiss</td>
<td>imitation</td>
<td>qualification</td>
<td>tire</td>
</tr>
<tr>
<td>center</td>
<td>dollar</td>
<td>immense</td>
<td>rabbit</td>
<td>tray</td>
</tr>
<tr>
<td>certainty</td>
<td>drum</td>
<td>inquiry</td>
<td>rail</td>
<td>tube</td>
</tr>
<tr>
<td>civilize</td>
<td>earnest</td>
<td>insult</td>
<td>reduction</td>
<td>victory</td>
</tr>
<tr>
<td>classification</td>
<td>educator</td>
<td>insure</td>
<td>religion</td>
<td>weaken</td>
</tr>
<tr>
<td>classify</td>
<td>elsewhere</td>
<td>interfere</td>
<td>republic</td>
<td>weave</td>
</tr>
<tr>
<td>clothe</td>
<td>empire</td>
<td>joke</td>
<td>ribbon</td>
<td>widow</td>
</tr>
<tr>
<td>collar</td>
<td>employee</td>
<td>kick</td>
<td>rival</td>
<td>wipe</td>
</tr>
<tr>
<td>collector</td>
<td>explosive</td>
<td>kiss</td>
<td>rug</td>
<td>witness</td>
</tr>
<tr>
<td>colony</td>
<td>favor</td>
<td>kneel</td>
<td>saddle</td>
<td>worship</td>
</tr>
<tr>
<td>color</td>
<td>favorite</td>
<td>latter</td>
<td>satisfactory</td>
<td>wrist</td>
</tr>
<tr>
<td>complicate</td>
<td>fellowship</td>
<td>liberty</td>
<td>satisfy</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 6
Words (their derivatives and inflectional forms included) which are not used in the entire Malaysian Secondary School English Textbooks (Forms 1-5)

| accustom  | colony  | governor | ownership | scrape  |
| angle  | curse  | decisive | idle | plaster  |
| anyhow  | decisive | idle | immense | plaster  |
| astonish  | devil | immense | liberty | redaction |
| awkward  | hearst | latter | liberty | redaction |
| barrel  | elsewhere | liberty | membership | simplicity |
| basis  | empire | membership | nut | simplicity |
| civilize  | fellowship | objection | saddle | simplicity |
| collar  | garage | owe | scatter | simplicity |
Jayakaran Mukundan and Anealka Aziz

Figs. 1-10 show the dispersion plots of the selected words within the entire set of textbooks. Words like much, study, country, and article are distributed well within and across the textbooks. Nevertheless, words like quite, black, science, and discussion are distributed well only in certain textbooks. Meanwhile, words like art and yellow are the most poorly distributed words within and across the textbooks.

**CONCLUSION**

Generally, the textbooks have large loads of vocabulary for each Form based on the results of the analysis. The results also conclude that the textbooks have not adequately prepared the students with the knowledge of the 2000 high frequency words effectively, as only 71.9% (1438) of the high frequency words are effectively introduced and repeated, despite the high coverage of 93.1% (1861 out of 2000 words). Meanwhile, about 6.9% of the words in the high frequency list (139 words) are not at all used in the textbooks. However, derivatives and inflections of some of these words are found in the textbooks. 2.3% of the 2000 high frequency words (46 words that take into account their derivatives and inflections) are never used at all in the entire set of Malaysian Secondary School English Textbooks (Forms 1-5). In terms of the distribution patterns of 10 randomly selected words from the 2000 high-frequency list, the

### TABLE 7
Ten randomly selected words and their ranking within the 2000 high frequency list

<table>
<thead>
<tr>
<th>No.</th>
<th>Words</th>
<th>Ranking in the 2000 High Frequency List (1-2000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>much</td>
<td>94</td>
</tr>
<tr>
<td>2</td>
<td>country</td>
<td>217</td>
</tr>
<tr>
<td>3</td>
<td>study</td>
<td>231</td>
</tr>
<tr>
<td>4</td>
<td>art</td>
<td>277</td>
</tr>
<tr>
<td>5</td>
<td>quite</td>
<td>405</td>
</tr>
<tr>
<td>6</td>
<td>black</td>
<td>586</td>
</tr>
<tr>
<td>7</td>
<td>science</td>
<td>653</td>
</tr>
<tr>
<td>8</td>
<td>discussion</td>
<td>832</td>
</tr>
<tr>
<td>9</td>
<td>article</td>
<td>976</td>
</tr>
<tr>
<td>10</td>
<td>yellow</td>
<td>1277</td>
</tr>
</tbody>
</table>

### TABLE 8
Distribution of words within Malaysian Secondary School English Textbooks (Forms 1-5)

<table>
<thead>
<tr>
<th>Words</th>
<th>Form 1</th>
<th>Form 2</th>
<th>Form 3</th>
<th>Form 4</th>
<th>Form 5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>much</td>
<td>30</td>
<td>33</td>
<td>48</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>2</td>
<td>country</td>
<td>12</td>
<td>24</td>
<td>7</td>
<td>37</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>study</td>
<td>16</td>
<td>13</td>
<td>23</td>
<td>11</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>art</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>quite</td>
<td>7</td>
<td>9</td>
<td>13</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>black</td>
<td>7</td>
<td>7</td>
<td>10</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>science</td>
<td>33</td>
<td>3</td>
<td>9</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>discussion</td>
<td>4</td>
<td>X</td>
<td>10</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>article</td>
<td>19</td>
<td>7</td>
<td>37</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>yellow</td>
<td>2</td>
<td>2</td>
<td>x</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Loading and Distribution of the 2000 High Frequency Words in Malaysian English Language Textbooks

Fig. 1: The use of ‘MUCH’ in Malaysian Secondary School English Language Textbooks for Forms 1-5

Fig. 2: The use of ‘COUNTRY’ in Malaysian Secondary School English Language Textbooks for Forms 1-5

Fig. 3: The use of ‘STUDY’ in Malaysian Secondary School English Language Textbooks for Forms 1-5

Fig. 4: The use of ‘ART’ in Malaysian Secondary School English Language Textbooks for Forms 1-5
<table>
<thead>
<tr>
<th>Fig. 5: The use of ‘QUITE’ in Malaysian Secondary School English Language Textbooks for Forms 1-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 6: The use of ‘BLACK’ in Malaysian Secondary School English Language Textbooks for Forms 1-5</td>
</tr>
<tr>
<td>Fig. 7: The use of ‘SCIENACE’ in Malaysian Secondary School English Language Textbooks for Forms 1-5</td>
</tr>
<tr>
<td>Fig. 8: The use of ‘DISCUSSION’ in Malaysian Secondary School English Language Textbooks for Forms 1-5</td>
</tr>
</tbody>
</table>
results show that although more than 90% of the 2000 high frequency words are covered in the textbooks, not all the words are distributed effectively within the textbooks to ensure learning. Therefore, material developers need to be cautioned that high coverage of the 2000 high frequency words does not ensure effective learning especially when the words are poorly distributed within and across textbooks.

REFERENCES


