Vocabulary Size and Vocabulary Learning Strategy among Chinese Undergraduates at Universiti Putra Malaysia

Yasir Bdaiwi Jasim Al-shujairi
yasir.bdaiwi.jasim1990@gmail.com
Al-Mustansiriya University, Iraq

Arifur Rahman
arifurrahman70@yahoo.com
Times University Bangladesh, Bangladesh

Ma Yuepeng
amandacai0329@gmail.com
Universiti Putra Malaysia, Malaysia

Abstract
While studies have investigated the strategies used by Chinese English as a Foreign Language (EFL) learners to learn new words and the relationship with their current vocabulary knowledge, there is a lack of studies that examine the vocabulary learning strategies (VLSs) and vocabulary sizes of Chinese students studying abroad in the L2 context. In view of the increasing number of Chinese students studying in ESL communities, this study aims to determine the vocabulary size of Chinese undergraduates at a high-ranking public university in Malaysia – Universiti Putra Malaysia (UPM) – and to discover the extent to which they employ a range of VLSs. Accordingly, 30 students enrolled in various degree programmes were randomly selected to take part in this study. The vocabulary size test (Schmitt, Schmitt, & Clapham, 2001) and a vocabulary level questionnaire based on Schmitt’s (1997) taxonomy of VLSs were adopted as the data collection instruments. The findings showed that, on average, the participants have a “Threshold” vocabulary size at the 3,000 and the 5,000-word level, which according to...
Schmitt et al. (2001) enables learners in reading different authentic texts. However, Nation and Beglar (2007), Schmitt and Schmitt (2014), Coxhead, Nation, and Sim (2015) suggested that non-native undergraduates studying in English as a medium of instruction (EMI) colleges and universities require a vocabulary size of 9,000-word family. The findings of this research raised important concerns that require our attention on the Chinese undergraduates studying in EMI context are at risk due to their insufficient vocabulary size. Moreover, results also revealed that the participants used direct VLSs such as memory and cognitive strategies the most. Taken together, the findings of this research provide insights into undergraduates’ overall inadequate vocabulary size, and how different VLSs were employed to cope with academic demands.

Keywords: Chinese EFL students, vocabulary learning strategies, vocabulary size

1. Introduction
To master any language, it is essential to learn as much vocabulary as possible because perceiving and producing a text depends on vocabulary knowledge. The amount of vocabulary that learners possess is considered to be a crucial factor in language acquisition because vocabulary composes the largest meaningful part of any language (Carter, McCarthy, and Channell, 1988). According to Wilkins (1972), a little can be conveyed without having enough knowledge of grammar but nothing can be conveyed without vocabulary. In short, it is hard to communicate with other people without enough vocabulary knowledge. Vocabulary plays a significant role in the acquisition of both a first (L1) and a second (L2) language. However, learning new words in L2 is considered one of the most difficult processes in human development (Gupta & MacWhinney, 1997). In fact, limited knowledge of words is considered being an obstacle that could slow down the process of learning among L2 students (Zhihong, 2000). The acquisition of enough vocabulary knowledge is one of the main challenges faced by the foreign language and second language learners because they need an adequate amount of vocabulary to comprehend what they read (Teng, 2014). Moreover, the size of a language lexicon grows over time as new words are added. However, the vocabulary of L2 students does not increase at the same pace as the addition of new vocabulary to the existing lexicon (Coady and Huckin, 1997).
Over the past three decades, a good deal of studies have sought to determine how learners learn new words effectively and how best to teach students learning another language. Evidence suggests that the employment of effective vocabulary learning strategies (VLSs) is one of the most important factors for successful vocabulary acquisition. Oxford (2003) stated that strategies of learning a second language are methods that learners utilize to improve their own L2 learning. These vocabulary learning strategies might involve cognitive, metacognitive, motivational or social activities. Schmitt (1997) defined a VLS as a process by which the information is acquired, stored, retrieved and utilized. In the same vein, Cameron (2001) defined a VLS as “activities that learners do to help themselves to comprehend and memorize words.” According to Nation (2001), a VLS comprises a set of general strategies of language learning. Therefore, it can be discerned that VLSs can play a key role in helping students learn new vocabulary of a target language.

With regard to students studying at the university level, White, Graves, and Slater (1990) stated that L2 students who are about to start their undergraduate studies and have limited English vocabulary will suffer long-lasting effects from this deficiency, including poor academic achievement as well as a lower standard of English proficiency. In China, students learn English as a foreign language from the age of eight or nine onward (Bolton, 2010). The English language is also the medium of instruction for many undergraduate programmes in China and elsewhere. Thus, Chinese EFL learners need to be able to recognize, understand, and use a considerable amount of English words to perform well at the university level at home and abroad.

Several previous studies have examined vocabulary size in relation to VLS among Chinese EFL learners at the university level in China (Gu, 2010; Teng, 2015; Wei, 2007; Wu, 2008). However, it seems that no previous study has looked at the vocabulary knowledge and the VLSs used by Chinese students studying abroad. As the English language is the medium of instruction for most of the programmes at tertiary level in Malaysia, Chinese students are required to obtain a particular level of English courses conducted at an English language centre at the university before they can enrol in their undergraduate studies. Thus, Chinese undergraduate students studying in Malaysia are assumed to be more proficient in the English language than their counterparts studying in China. Accordingly, this study examines the vocabulary size and use of a range of VLSs among Chinese undergraduate students at a high-ranking public university in Malaysia in order to answer the following research questions:
1. What is the vocabulary level of Chinese undergraduates at Universiti Putra Malaysia?

2. What are the most and least commonly applied VLSs among Chinese undergraduates at Universiti Putra Malaysia?

2. Vocabulary Size and University Success

Vocabulary has long been recognized as one of the language-related factors that contribute to academic success (Birkenhead, 2011). Also, contemporary vocabulary research has confirmed the general intuitive feeling that many teachers and researchers have long harboured about vocabulary: the more words that are known and used well, the better. For instance, Elder and von Randow (2008) stated that: “The size of a student’s lexicon functions as an effective predictor of academic success” (p. 177). Similarly, Douglas (1996) stated that students who excel in their studies are the ones with “an apt, varied, and accurate written vocabulary” (p. 205). In more recent studies, vocabulary size has been observed to be the best indicator of overall academic achievement (Cloate, 2016; Daller and Phelan, 2013; Harrington & Roche, 2014). Further, one of the latest studies in the Malaysian context, which was conducted by Rahman (2017), found that second-year tertiary students’ vocabulary size could contribute as much as 25% to their overall cumulative grade point average. This finding is similar to that of Harrington and Roche (2014). Milton (2013) also reported that ESL learners’ vocabulary knowledge plays a fundamental role in achieving better performance in English language proficiency. In addition, using a combination of different tests, Daller and Phelan (2013) reported similar conclusions and pointed out that students’ ability to use vocabulary accurately seems to be a reliable predictor for their academic success.

Nation (2001) examined the number of words a reader needs in order to be able to comprehend an academic text and found that an L2 tertiary student needs a vocabulary size of 9,000-word family. Not surprisingly, studies have also shown that comprehension improves as more words are acquired. However, the ideal threshold for unassisted reading comprehension was reported to be 98% (cf.; Hu & Nation, 2001; Nation, 2001). For this level of comprehension, learners would need to have about a 4,000-word vocabulary; this would consist of 2,000 high-frequency general service words, 570 words on the academic word list (AWL) and about 1,000 technical words, on top of proper nouns and 11 low-frequency words (Hunt, & Beglar, 2005 p.
From this, we can see that the vocabulary burden on L2 learners in an academic environment is a high one.

One of the compelling reasons for measuring vocabulary size is to check whether learners possess adequate vocabulary for some specific purposes such as reading a novel, newspapers, watching movies, and understanding everyday conversations. Also, the results derived from a vocabulary size test can offer further compelling evidence whether students should focus on gaining high-frequency vocabulary, academic words, or low-frequency words. Kremmela and Schmitt (2016) argued that these tests can be used for many purposes, ranging from a proxy for language proficiency in general, to a diagnostic to determine whether learners possess the lexical resources to be functional in the four skills of Reading, Writing, Listening, and Speaking. Moreover, the estimation of learners’ vocabulary size can facilitate the comparison of different groups of students and the assessment of continuing vocabulary development, or offer “one kind of goal for learners of English as a second or foreign language” (Nation, 2013, p. 522).

2.1 Past Studies on the Vocabulary Level of Chinese University Students

A considerable amount of literature has been published on the vocabulary level of Chinese undergraduates. In a large longitudinal study, Fan (2003) investigated the vocabulary level of 1,076 undergraduates studying at different universities in Hong Kong and found that the students’ vocabulary size decreases respectively in the low frequency vocabulary level: 2,000 words (95%), 3,000 (86%), 5,000 (68%), and 10,000 (36%) words which seems to be a common trend among EFL and ESL learners. Another study by Li and MacGregor (2010) on 128 participants and Kirchner (2013) offered probably the most comprehensive measurement of Chinese tertiary students’ vocabulary level in China, showing a similar finding (downward trends with respect to frequency level).

On the contrary, a more recent study by Liu (2016), which was conducted among 108 first-year undergraduate students in China, found a mean vocabulary size of 27% in 2,000, 16% in the 3,000, 11% in the 5,000, and 9.89% at the academic-word level which is comparatively lower than the other studies (Fan, 2003; Li and MacGregor, 2010; Kirchner, 2013). This inconsistency can be explained by the fact that the participants of Liu’s (2016) study were fresh undergraduates whereas, Fan (2003), Li and MacGregor (2010) and Kirchner (2013) studies took
the sophomores. A similar finding was also observed among 156 diploma-level students in the Faculty of Architecture, Planning and Surveying of a public university in Malaysia showing a lower level of vocabulary size among diploma students. These findings, which are summarized in Table 1, show that the results for Chinese tertiary students’ vocabulary size are varied. Therefore, there is a need to measure these students’ vocabulary size.

Table 1. Past Studies on Students’ Vocabulary Size

<table>
<thead>
<tr>
<th>Author</th>
<th>Participants</th>
<th>VL2,000</th>
<th>VL3,000</th>
<th>VL5,000</th>
<th>VL10,000</th>
<th>Academic-word level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li &amp; MacGregor (2010)</td>
<td>128 Chinese undergraduate students</td>
<td>97%</td>
<td>92%</td>
<td>77%</td>
<td>30%</td>
<td>94%</td>
</tr>
<tr>
<td>Fan (2001)</td>
<td>1,076 undergraduates of different universities in Hong Kong</td>
<td>95%</td>
<td>86%</td>
<td>68%</td>
<td>36%</td>
<td></td>
</tr>
<tr>
<td>Kirchner (2013)</td>
<td>70 Chinese first-year undergraduates</td>
<td>90%</td>
<td>65%</td>
<td>43%</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Manan (2016)</td>
<td>156 diploma-level Malaysian students</td>
<td>41%</td>
<td>36%</td>
<td>34%</td>
<td>34%</td>
<td></td>
</tr>
<tr>
<td>Liu (2016)</td>
<td>108 freshers in Jiangsu University of Technology, China</td>
<td>27%</td>
<td>16%</td>
<td>11%</td>
<td></td>
<td>9.89%</td>
</tr>
</tbody>
</table>

2.2 Past Studies on Vocabulary Learning Strategies

According to Oxford (1990), strategies for vocabulary learning can be categorized into two main types: direct and indirect. The direct strategies involve learning the language directly and include cognitive strategies, memory strategies and compensation strategies. When using cognitive strategies, learners tend to perceive and produce a new language by only different means.
Examples of cognitive strategies are deductive reasoning and summarizing. In contrast, indirect strategies promote vocabulary learning without a direct focus on the target language (Oxford, 1990, p. 135). There are three main types of indirect strategies: social, metacognitive, and affective. Social strategies help learners to learn new vocabulary through interacting and communicating with people while metacognitive strategies help learners to control their own cognition by using planning, evaluating, and monitoring processes. In the case of affective strategies, learners regulate their attitudes, emotions, and motivation. Moreover, Nation (2001) argues that an ample amount of vocabulary can be learned by applying different VLSs and that various strategies prove useful for students of different language levels. Moreover, Oxford (1990) sets the scoring system of both indirect and direct strategies suggesting that values below 2.5 as “low strategy use”, scores between 2.5 to 3.5 are “medium strategy use” and the scores above 3.5 are deemed as “high strategy use”.

In the past few years, several studies in China (e.g., Gu & Johnson, 1996; Peacock & Ho, 2003; Wu, 2008; Gu, 2010; Lou, 2014; Teng, 2015; Zhang, 2011) have been conducted to investigate the VLSs and vocabulary knowledge of Chinese EFL students at different educational levels (secondary school, preparatory level, and university level). One of the most significant studies on the effects of direct and indirect VLSs is that of Teng (2015), who conducted a study among 145 Chinese undergraduates in China. His results showed that Chinese EFL students use more direct strategies such as memory and cognitive than indirect strategies such as social and metacognitive. Thus, direct strategies such as the application of memory strategies, which are related to remembering and retrieving new information; seem to be commonly used by Chinese EFL learners.

### 2.2.1 Memory Strategies

Memory strategies involve facilitating learners storing and retrieving new information (Teng, 2015; Fan 2015). Examples of memory strategies are using groups of unrelated words and grouping words according to categories like synonyms or common themes. Kafipour, Yazdi, Soori, and Shokrpour (2011) examined VLSs and their relationship with the EFL vocabulary size of 238 Iranian university students. They found that memory strategies were the most frequently used type of VLS, with a mean score of 3.01. A similar finding was observed in the Jordanian context in Al-Kasawneh (2012)’s analysis of the various VLSs used by Jordanian undergraduate
students at the Jordan University of Science and Technology, where the memory strategy was identified as the most commonly used strategy among Jordanian undergraduates with a mean score of 2.76. In a study based in Morocco, Ghouati (2014) investigated the VLSs of 60 Master of Arts students studying English at the School of Arts & Humanities. The findings of his study were also in line with the findings of the above-mentioned studies. In his study, memory strategies were the second most common type of strategy used by the students, with a percentage of 27.69%. However, Komol and Sripetpun (2011) in their investigation of the VLSs adopted by Thai second-year university students in Thailand, found that memory strategies were among the least used VLSs for learning EFL. Despite the findings of Komol and Sripetpun (2011), it seems that memory strategies are commonly employed by EFL learners of different nationalities.

2.2.2 Cognitive Strategies

Cognitive strategies are somewhat similar to that of memory strategies. In this, they do not focus on manipulative mental processing; instead, they rely on repetition and mechanical means to study vocabularies. Word lists flash cards, and taking notes are considered as cognitive strategies. Kafipour et al. (2011) found that cognitive strategies were the least commonly employed VLS among Iranian university students with a mean score of 2.96. Similarly, Al-Khasawneh (2012) also found these strategies were the second least used among Jordanian students with a mean score of 2.68 mean score. However, the use of cognitive strategies was found to be moderate among Moroccan (Ghouati, 2014) and Chinese (Lou, 2014) EFL learners with a percentage of 22.04% and a mean score of 3.30, respectively. In contrast, Komol and Sripetpun (2011) showed that these strategies were highly used by Thai students, with a mean score of 3.20.

2.2.3 Compensation Strategies

Examples of compensation strategies are guessing intelligently while reading. This can allow learners to use the language despite their large gaps in linguistic knowledge. Learners apply these strategies directly to learn new linguistic items. Recent studies have reported that the use of compensation strategies among EFL learners is quite often. For instance, Peacock and Ho (2003) examined the VLSs used by tertiary-level students in Hong Kong and the correlation with vocabulary knowledge. They found that compensation strategies were among the strategies most
commonly employed by students in learning new words. Similarly, Seffar (2005 as cited in Kablani, 2006) conducted a study in the Moroccan context among 100 university students and discovered that compensation strategies were the most frequently used type of strategy.

2.2.4 Social Strategies

According to Teng (2015), social strategies facilitate students when learning new words through interaction with others. One example of social strategies is asking classmates and colleagues for the meaning of a particular word. The use of social strategies was found to be “medium strategy use” in Kafipour et al.’s (2011) study in the Iranian context, with a mean score of 2.97. On the other hand, this type of strategy was the most frequently used by university students in learning new words in Al-Khasawneh’s (2012) study in Jordan with a mean score of 3.03. However, social strategies were among the least used by Chinese (Lou, 2014) and Moroccan (Ghouati, 2014) EFL learners. These latter two findings are parallel with that of Komol and Sripetpun (2011), who showed that social strategies were the least employed type of VLS among Thai EFL learners. Therefore, Arab, Chinese, and Thai EFL learners may not be fully aware of social strategies for learning new vocabulary.

2.2.5 Metacognitive Strategies

Metacognitive strategies are strategies used by learners to control and assess their learning. Examples of metacognitive strategies are reading books, watching movies, and interacting with native speakers to get maximum exposure to language. Efficient use of time and knowing when to actively study a new word are also useful metacognitive strategies. Some studies (e.g., Kafipour et al., 2011; Komol & Sripetpun, 2011; Lou, 2014) found a high frequency use of metacognitive strategies in learning new words, whereas some other studies (e.g., Al-Khasawneh, 2012; Ghouati, 2014) found a low frequency of use. Kafipour et al. (2011) revealed that Iranian EFL learners rely heavily on metacognitive strategies for learning new vocabulary. Similarly, Lou (2014) showed that these strategies were the most frequently utilized among Chinese students (mean = 3.57). Similarly, Komol and Sripetpun (2011) found that these strategies were found to be common among Thai undergraduates (mean = 3.00). However, Ghouati (2014) and Al-Khasawneh (2012) found that metacognitive strategies were the least
used type of VSL among Moroccan and Iranian undergraduates, respectively. Thus there is a need for further studies to examine the use of metacognitive VLSs.

2.3 Past Studies on the Vocabulary Learning Strategies used by Chinese University Students

Several studies (e.g. Fan, 2015; Gu, 2010; Teng, 2010) were conducted to examine the VLSs employed by Chinese undergraduates. In a study based in China, Gu and Johnson (1996) investigated the VLSs of 850 Chinese sophomores at Beijing University and found that the students apply several VLSs for learning new words. They found that self-initiation and selective attention, which are both metacognitive strategies, are the most common strategies employed by Chinese students. In the same line, Wenyu (1998) examined the VLSs of Chinese students at Beijing Normal University. They found that Chinese undergraduates applied a number of VLSs including metacognitive and cognitive strategies. In addition, they showed that VLSs are strongly correlated with both the quality and quantity of English vocabulary size. Based on the two above-mentioned studies, it could be argued that Chinese students at university level rely more on metacognitive strategies in learning new words.

Gu (2010) examined the changes in the VLSs adopted by Chinese L2 students in Singapore during a 6-month language programme by using used the vocabulary level questionnaire (version 5) developed by Gu and Hu (2003), which is a seven-point Likert-scale questionnaire. Gu (2010) noticed changes in the various VLSs used by the Chinese EFL learners after 6 months of study. He also found that looking in the dictionary and contextual guessing were the most common strategies applied by Chinese L2 undergraduates, while visual repetition and remembering semantically related words were the least used strategies.

Except for Gu’s (2010) study, most previous studies have investigated Chinese students studying in China and who do not use English in real-life situations outside the classroom. To the best of our knowledge, no other research has been conducted on Chinese students pursuing degrees abroad (i.e., outside China). Therefore, the present study investigates the use of VLSs among Chinese undergraduate students enrolled in various undergraduate degree programmes at Universiti Putra Malaysia (UPM) in Malaysia. Chinese undergraduates in Malaysia use the English language not only inside the classroom, but also outside because they need to interact
with colleagues, lecturers and friends in real-life situations. Hence, it is assumed that these students have a better opportunity to learn a larger amount of words. Besides, they might use a wide range of VLSs, including affective and social strategies, to develop their vocabulary knowledge. It is hypothesized that their vocabulary size will be positively affected by interacting with their peers or friends and that they might be more motivated to learn the language due to being immersed in a different study environment.

3. Methodology

This research employed a quantitative approach to collect data using a questionnaire and a test. The results obtained from the questionnaires were analysed using Statistical Package for the Social Sciences (SPSS) statistical software.

3.1 Participants

A total of 38 Chinese second-year undergraduates studying at UPM took part in this research. They were randomly selected from different degree programmes at different faculties (e.g., Business Management, Communication, and Human Sciences). All the participants were previously students at the ELS English Language Centre at UPM. The ELS institution consists of nine levels, where levels 101 to 103 denote beginners; levels 104 to 106 lower moderate, level 107 moderate and levels 108 and 109 advanced. In order to be allowed to enrol for academic study at UPM, students have to pass level 107. Therefore, the proficiency level of the participants can be classed as moderate because it is given that they have all passed the 107 level.

In China, English is taught and learned as a foreign language (EFL), whereas in Malaysia, it is a second language. Therefore, it was assumed that all the participants had practised the language and used it in real-life situations. They were given the VLS questionnaire and vocabulary size test (VST) together. Brief verbal instructions in Chinese (Mandarin) were given to students in order to clarify how the questionnaire and the test should be answered. Eight of the questions were excluded from the study as they were incomplete or had not been taken seriously by the students (e.g., some students had just ticked answers without reading the questions and others skipped some items). The remaining 30 questions were used for this research. First, they were crossed-checked and then the data was entered into SPSS for the purpose of analysis.
3.2 Instrumentation

Two instruments were used in this study. One was the vocabulary size test developed by Schmitt et al. (2001, p. 63). There are two versions of this test. The present study adopted the second version because it was established based on a new AWL Schmitt et al. (2001, p. 63). The second version consists of five vocabulary levels: 2,000, 3,000, 5,000, 10,000, and academic-word levels. There are 30 vocabulary items in groups of 10 for every word level. In each group, the answer of six options is given in which three of them are distracters for each item. For every vocabulary item, the respondents were instructed to select only one answer option. This test is considered one of the most commonly used tests to estimate learners’ vocabulary size, and a number of past studies (Qian, 2002; Teng, 2015; Xiang & Fulcher, 2007) has adopted it.

The Cronbach-α reliability coefficient for the questionnaire used in the present study was estimated as 0.91, which Cohen and Holliday (1982 mentioned in Bryman and Cramer 2002, p. 174) is high. A bilingual version (Chinese–English translation) of the test was used in this study in order to prevent any misunderstanding caused by a lack of understanding of grammar among the participants. Elgort (in press) states that distributing a bilingual version of a test can result in 10% higher scores. These higher scores are most likely due to participants being able to see a translation of the terms into their own language, which avoids difficulty in interpreting the grammar of English definitions and makes it easier to comprehend the questions. Note that in a bilingual version of this type of test, the test items are not translated into the first language; rather, first language (translated) single words or phrases are provided for every L2 choice. Therefore, the choices for every test item are first language (Mandarin) synonyms of the target language (English) words.

Other instrument that was utilized in this research was a five-point Likert-scale VLS questionnaire. The answering options on the scale ranged from “Never” to “Always”. The questionnaire has been used in several previous studies (Gu & Johnson, 1996; Teng, 2015; Zhang, 2011; Zhang, 2009). The data was calculated to have a Cronbach’s alpha of 0.90, which indicates a high level of reliability (Bennett & Allen, 2012). The questionnaire was a bilingual questionnaire, i.e., it was translated into the Chinese language (Mandarin) so that the test items would be as comprehensible as possible for the test takers. The VLS questionnaire comprised two main sections. Section 1 was designed to elicit general demographic information about the participants such as gender and age. Section 2 contained 40 items, which included statements
about the use of various VLSs. During the analysis, the 40 VLSs were combined into 10 main
categories of direct and indirect strategies.

3.3 Data Analysis Procedure

In this study, IBM SPSS 23.0 was used to analyse the data. First, the participants’ answers to the
VLS questionnaire were changed into scores, where 1 = “Never”, 2 = “Rarely”, 3 =
“Sometimes”, 4 = “Usually” and 5 = “Always”. Then, the scores for the vocabulary test and the
questionnaire were analysed by calculating descriptive statistics and applying an independent t-
test. The first research question was answered by measuring the mean and standard deviation for
each vocabulary level. According to Phakiti (2015), a paired-samples t-test have a widespread
popularity in determining whether there is a statistically significant difference between two sets
of means taken from the same respondents. Thus, to answer research question two, along with
the mean and standard deviation, the significance (p) value was measured to determine whether
there was any difference in the employment of direct and indirect strategies among the
participants.

4. Findings and Discussion

The findings are presented below, according to the research questions. The accompanying tables
provide the mean (M), standard deviation (SD), and the Pearson’s correlation (r) between
vocabulary size (VS) and vocabulary learning strategies (VLSs).

Question 1: What is the vocabulary size of Chinese undergraduates at Universiti Putra
Malaysia?

In order to answer the first research question, it was necessary to calculate the vocabulary size
for each word level. The descriptive statistics (M and SD) of the undergraduate Chinese students’
VST results are provided in Table 2. Schmitt et al. (2001) set the cut-off point as 24 over 30
items (80%) for obtaining each level. Thus, taking a minimum of 80% as the cut-off point for an
average vocabulary level, the analysis showed that the participants had an inadequate vocabulary
size at the 2,000 (M = 50.33); 3,000-word vocabulary size (mean = 70.81); 5,000-word
vocabulary size (mean = 64.43); 10,000-word (M = 23.53); and academic-word (M = 54.04). The
overall mean of 57.60 indicates that the average vocabulary size of the Chinese undergraduates at
Vocabulary size and vocabulary learning strategy

UPM is below the cut-off point (80%) suggested by Schmitt et al. (2001). The standard deviations of the VST scores for the 5,000-word level and the academic-word level indicate that the students had relatively more heterogeneous academic vocabulary size ($SD = 25.280$) than those of 5,000-word level ($SD = 20.793$) vocabulary size. It has been argued that a vocabulary size of at least 5,000 words is necessary for reading advanced and academic texts (Hirsh & Nation, 1992). Thus, Chinese undergraduate students in UPM have achieved that level of ability. Table 3 shows the means and standard deviations of the participants’ total scores for the VST at each word level.

Overall, the findings of the present study indicate that the students’ vocabulary size at the 3,000, 5,000, and academic-word levels shows an upward trend, which is somewhat surprising given the fact that other research studies (Fan, 2001; Kirchner, 2013; Li & MacGregor, 2010; Liu, 2016; Manan, 2016) showed a downward trend. With regard to the mean vocabulary size for the 2,000-word level, the vocabulary size (50%) observed in this investigation is far below that observed by Li and MacGregor (2010), Fan (2001), and Kirchner (2013), who reported a greater vocabulary size at the 2,000 level of 97%, 95%, and 90%, respectively. The fact that the participants of this study were university students who were more frequently encountered/exposed to 3000-10,000 word level rather than high frequency (2000) level can explain this lower vocabulary size at 2000 word family level. They had to read different academic textbooks that are written in sophisticated language.

The subjects in the current research outperformed the participants in the Manan (2016) and Liu (2016) studies. One possible explanation for this difference might be that the participants in the current study were second-year undergraduates and had already completed foundation English language courses at the ELS English Language Centre in UPM, whereas the participants in the studies conducted by Manan (2016) and Liu (2016) were diploma students and freshers respectively. In the case of 3,000- and 5,000-word levels, the current findings showed that the participants had a smaller vocabulary size (64% and 70%) than that observed by Li and MacGregor (2010) and Fan (2001), who found a higher vocabulary size among Chinese undergraduates. In contrast, the participants of the current study scored much higher in the 3,000- and 5,000-word levels than those in the studies by Kirchner (2013), Manan (2016), and Liu (2016).
As for the 10,000-word level, the result of the present study is nearly the same as that reported in Li and MacGregor (2010), Fan (2001), and Manan (2016), but not Kirchner (2013) who found a much lower vocabulary size (15%) among Chinese first-year undergraduates. As for the Chinese undergraduate students’ academic-word level, the value of 54% observed in the current investigation is far below that of the result of 94% observed by Li and MacGregor (2010) yet far greater than that of 9.89% reported by Liu (2016). This inconsistency may be due to Li and MacGregor’s (2010) participants were pre-university students, whereas the participants in the current study were second-year undergraduates. To make sense of these differences in the results regarding Chinese undergraduate students’ vocabulary size, it is crucial to know the requisite amount of words that are needed for ESL or EFL learners studying at tertiary level, especially where English is the medium of instruction and the language is used in university textbooks.

Evidence suggested that in order to comprehend daily spoken English, a student needs 95% word coverage which requires about 4,000-word vocabulary size (Adolphs & Schmitt, 2004) or a 98% word coverage that is about 6,000 to 7,000 words (Nation, 2006). Similarly, according to Webb and Rogers (2009a), 3,000 most frequently words are essential to achieve 95% coverage in English TV shows and 7,000 most frequently used words allows the learner to gain 98% coverage (Webb & Rodgers 2009b). However, according to Nation (2007) and Schmitt (2008), a vocabulary size of 9,000-words is needed to comprehend academic texts at the tertiary level. Therefore, it can be argued that the vocabulary size of Chinese undergraduates still falls below the ideal size reported in the literature (Nation, 2007; Webb & Rodgers, 2009).

<table>
<thead>
<tr>
<th>Vocabulary word level</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,000-word level</td>
<td>30</td>
<td>50.33</td>
<td>20.608</td>
</tr>
<tr>
<td>3,000-word level</td>
<td>30</td>
<td>70.81</td>
<td>28.443</td>
</tr>
<tr>
<td>5,000-word level</td>
<td>30</td>
<td>64.43</td>
<td>20.793</td>
</tr>
<tr>
<td>10,000-word level</td>
<td>30</td>
<td>23.53</td>
<td>16.829</td>
</tr>
<tr>
<td>Academic vocabulary</td>
<td>30</td>
<td>54.04</td>
<td>25.280</td>
</tr>
<tr>
<td>Vocabulary level test</td>
<td>30</td>
<td>57.60</td>
<td>15.937</td>
</tr>
</tbody>
</table>
Question 2: What are the most and the least common VLSs applied by Chinese undergraduates at UPM?

In order to answer research question two, the means and standard deviations of the 30 eligible VLS were calculated to determine the VLSs that were most and least frequently by Chinese undergraduates at UPM. The results are provided in Table 3 below.

Table 3: Descriptive Statistics of Test Takers’ Usage of VLSs.

<table>
<thead>
<tr>
<th>Direct strategies</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory strategies</td>
<td>30</td>
<td>3.77</td>
<td>1.21</td>
</tr>
<tr>
<td>Cognitive strategies</td>
<td>30</td>
<td>3.56</td>
<td>1.32</td>
</tr>
<tr>
<td>Compensation strategies</td>
<td>30</td>
<td>2.82</td>
<td>1.12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indirect strategies</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social strategies</td>
<td>30</td>
<td>3.14</td>
<td>1.36</td>
</tr>
<tr>
<td>Metacognitive strategies</td>
<td>30</td>
<td>1.74</td>
<td>1.21</td>
</tr>
<tr>
<td>Affective strategies</td>
<td>30</td>
<td>1.73</td>
<td>1.12</td>
</tr>
</tbody>
</table>

The scoring system that is developed by Oxford (1990) reported that the scores below 2.5 as “low strategy use”, scores between 2.5 to 3.5 are “medium strategy use” and the scores above 3.5 are deemed as “high strategy use”. As can be seen from Table 3, the Chinese undergraduates at UPM tended to employ direct VLS more often than indirect strategies when learning new lexical words. Memory strategies were reported as the most frequently used strategies by these Chinese students (mean = 3.77) followed by cognitive strategies (mean = 3.56). This indicates that the Chinese undergraduate students spent a long time memorizing words intentionally. In addition, the extent of their use of cognitive strategies proved that they produced words by utilizing different methods. The results also showed that among the indirect strategies, social strategies were used very frequently. This possibly means that the Chinese students learn new lexical words by interacting with others. As mentioned above, Malaysia is an ESL country, so when students come from China, which is an EFL country, to study in Malaysia they must learn to interact with
local and foreign students for study and/or friendship purposes. This may cause them to use social strategies more in learning new words and developing their vocabulary knowledge. On the other hand, the participants tended to use metacognitive and affective (mean = 1.73) strategies less frequently (mean = 1.74 and 1.73, respectively). The infrequent use of these two strategies indicates that the Chinese students at UPM did not self-monitor or self-evaluate their learning of new vocabulary and did not drive their own motivation and emotions in developing their knowledge of vocabulary.

In the next step of the analysis, a paired sample t-test was applied to the data in order to find out whether there was a significant difference between the use of direct and indirect VLSs. The findings are given in Table 4.

<table>
<thead>
<tr>
<th>Type of strategy</th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>3.38</td>
<td>1.18</td>
<td>6.14</td>
<td>68</td>
<td>.000</td>
</tr>
<tr>
<td>Indirect</td>
<td>2.20</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The findings in Table 4 show that there was a significant difference in the application of direct and indirect VLSs (p < 0.001). This finding validates the results presented in Table 3, i.e., that the Chinese undergraduates at UPM tended to use memory and cognitive strategies more than other types of VLS and that they relied on these strategies when learning new words. It is also worth mentioning that Chinese undergraduates at UPM also used social strategies that are a form of an indirect strategy and involve interacting and communicating with others to learn new words. This differs slightly from using memory strategies; while using memory strategies is purely intentional; learning new words through social strategies is mostly unintentional. In other words, students might acquire some words without being aware of them.

In general, this result is in agreement with that reported in other studies in the same area that link the use of direct VLSs with overall vocabulary knowledge. It is particularly encouraging to compare the finding with that found by Teng (2015), who showed that Chinese EFL students tended to use direct strategies more than indirect strategies. With regard to memory strategies, the present findings corroborate the research of Kafipour et al. (2011) and Sahbazian (2004),
who demonstrated that memory strategies were the most frequently employed category among Iranian and Turkish undergraduates respectively. Furthermore, Al-Khasawneh (2012), in a Jordanian university, Ghouati (2014) in a Moroccan university and Teng (2015) in a Chinese university context found a similar practice among undergraduates of using memory strategies more than other strategies. A possible explanation for this might be that the students singled out memory strategies for learning vocabulary because these strategies are simple to be employed and do not require much of mental activity and mental processing (Kafipour et al., 2012). Another possible explanation is that they were used due to the popularity of rote learning among teachers and students (Sahbazian, 2004). On the other hand, the finding of the current study contradicts that of Zhang (2011), Ming Wei (2007), Schmitt (1997), Gu and Schmitt (1997), and Johnson (1996), who found that memory strategies were the VLS least employed by undergraduates. Moreover, Amirian and Heshmatifar (2013) argue that “postgraduate students as more successful learners who tend to employ a wide range of vocabulary learning strategies rather than just memorization and rote learning” (p. 640).

As for the utilization of cognitive strategies to learn new vocabulary, the participants of the current study had a higher average score compared to the participants in the studies conducted by Kafipour et al. (2011), Al-Khasawneh (2012), Ghouati (2014), Lou (2014), and Komol and Sripetpun (2011). The finding of the present study is in line with the view of Oxford (1990 cited in Ghouati, 2014), who pointed out that cognitive strategies were commonly used strategies among L2 learners. In addition, in contrast to the studies of Gu and Johnson (1996), Al-Khasawneh (2012) and Ghouati (2014), the participants of the present study ranked cognitive strategies higher than metacognitive strategies. As for compensation strategies, the participants in the current study ranked them as the third most-used type of strategy (“medium strategy use”) for learning new vocabulary. This finding conflicts with a number of China-based research studies (e.g., Bedell & Oxford, 1996; Goh & Foong, 1997) that were conducted among Chinese undergraduates and that showed that compensation strategies were among the most common strategies used by EFL learners at tertiary level. The present finding also differs from the results in Chang (1992), Goh, and Kwah (1997), who conducted studies among Chinese students studying in Singapore and the USA, respectively, and who both found that compensation strategies were the most commonly used.
In addition, it can be observed from the results of the current study that among the three categories of direct strategy, Chinese undergraduates studying in a foreign country where English is the medium of instruction preferred the memory strategy the most and the compensation strategy of vocabulary learning the least. As mentioned earlier, both the direct and indirect strategies of vocabulary learning were found to be popular among the Chinese undergraduates who participated in the present study. However, the indirect strategies rated the second position. Furthermore, among the three indirect strategies, social strategies were practised the most and affective strategies the least. This finding is in line with that of Wharton (2000), who stated that Chinese undergraduates studying in a foreign context such as Singapore used social strategies more frequently than other types of strategy. However, the results of the current research are not in line with some previous studies (Ghouati, 2014; Komol & Sripetpun, 2011; Lou, 2014) where social strategies were found to be the least frequently used type of VLS. This is possibly due to variations in the educational background of the participants. It is also worth mentioning that the Chinese students who employed a wide range of social strategies studied the language in Singapore, where English is considered as a second language (Kachru, 2006).

Students living and studying in an English as a second language-speaking country such as Malaysia where the use of English is widespread, were impeded by their very limited proficiency in English and due to the necessity to communicate, Chinese students had to use social strategies more frequently (Rao, 2014). In this study, metacognitive strategies were the second least popular category of strategy used by Chinese undergraduates. This result differs from earlier studies (Kafipour et al., 2011; Kalajahi & Pourshahian, 2012; Komol & Sripetpun, 2014; Lou, 2014) but is consistent with Al-Khasawneh (2012) and Ghouati (2014). The present findings also differ from Fan (2003) and Nemati, (2008) studies, who reported that high achievers are inclined to use metacognitive strategies more than other strategies.

5. Conclusion
The present study investigated Chinese overseas students’ vocabulary size and VLS at University Putra Malaysia (UPM) by using the vocabulary size test (Schmitt et al., 2001) and a questionnaire based on Schmitt’s (1997) taxonomy of VLSs. The results revealed that Chinese learners generally reached the 5000-word level in the size of vocabulary and they mostly adopted memory strategies as direct strategies and social strategies as indirect strategies. Hence, Chinese
undergraduates at UPM employed a variety of direct and indirect strategies to facilitate their English vocabulary learning. The findings would be of interest to Chinese undergraduates studying in an EMI context where an adequate vocabulary size (9,000) is prerequisite to comprehend academic text and class lectures. EMI lecturers are encouraged to measure their undergraduates’ vocabulary size since vocabulary size has been identified as the potential contributor to academic achievement (Milton & Treffer-Daller, 2013; Yixin and Daller (2014; Roche & Harrington, 2013) as well as overall English language proficiency (Stæhr, 2008; Milton, 2013; Hacking, Rubio & Tschirner, 2019). In light of the results of the present research, it is recommended metacognitive and compensation strategies are highlighted in the teaching of EFL learners’ to improve their vocabulary learning. Also, it is very important for English language teachers to provide more training on how lesser used strategies such as metacognitive and compensation strategies can also be used. By raising the students’ awareness of the use of these low-ranking strategies, students may improve their vocabulary learning for specific learning and their English proficiency in general. In addition, the present study provided valuable data on Chinese EFL learners’ vocabulary size and VLSs in the second language teaching context of Malaysia and also addressed the importance of learning vocabulary in second/foreign language learning. It is hoped that the results of this study will be of benefit not only to English language learners, but also to English language teachers in their teaching practice, especially in relation to the curriculum of vocabulary learning.

Nevertheless, it should be noted this is a case study on a limited sample of Chinese undergraduates at UPM, so it should be regarded as an initial study on the vocabulary learning process. Therefore, it is recommended that further research investigates the methods of acquiring vocabulary knowledge among Chinese EFL learners studying various degree programmes. Moreover, future studies could also analyse data from a larger sampling in order to obtain more accurate and generalizable data.
References


---

**About the Authors**

**Yasir Bdaiwi Jasim Al-shujairi** is an academic staff at the Department of English, College of Arts, Al-Mustansiriya University. His areas of research interest include academic writing, genre analysis, corpus linguistics and computational linguistics.

**Arifur Rahman** is working as a Lecturer & Coordinator in the Department of English at Times University Bangladesh. He has completed his BA (Hons) in English Language & Literature. He did his first MA Degree in English Language at Khulna University, Bangladesh. Recently he has completed his 2nd MA in English Language by research from the University of Putra Malaysia. Arifur Rahman does research in the field of Vocabulary Size; Vocabulary Learning Strategy, Academic Achievement, English Language Proficiency; Tertiary Education; Second Language Acquisition; English-Medium Instruction (EMI); Language Testing, Language Dominance, STEM Education

**Ma Yuepeng** is a doctoral candidate at Department of English, Faculty of Modern Languages and Communications in Universiti Putra Malaysia. She is currently studying Discourse analysis in UPM.