November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

Developing a Vocational Engineering Word List: Bridging the Gaps of Secondary and Tertiary Education

Ng Yu Jin
College of Energy Economics and Social Sciences,
Universiti Tenaga Nasional
Jalan Ikram-UNITEN
43000, Malaysia
Author Email: yujin@uniten.edu.my

Chong Seng Tong
College of Energy Economics and Social Sciences,
Universiti Tenaga Nasional
Jalan Ikram-UNITEN
43000, Malaysia
Author Email: stchong@uniten.edu.my

Loh Kee Shyuan
Fuel Cell Institute
Universiti Kebangsaan Malaysia
43600, Bangi, Malaysia
Author Email: ksloh@ukm.edu.my

Ahmad Zufrie Abd Rahman
Malaysian Examinations Councils
Persiaran 1, Bandar Baru Selayang
68100, Batu Caves, Selangor
Author Email: ahmadzufrie@gmail.com

Sivadass Thiruchelvam
College of Engineering
Universiti Tenaga Nasional
Jalan Ikram-UNITEN
43000, Malaysia
Author Email: sivadass@uniten.edu.my

Article Info
Volume 81
Page Number: ***3232 - 3242***
Publication Issue:
November-December 2019

Abstract:
There is a meaningful need to have an engineering corpus in Malaysia to bridge the vocabulary gap from secondary to tertiary engineering education. Thus, using the prescribed Malaysian KBSM vocational engineering textbooks, the development of Malaysia's very own vocational syllabus engineering corpus was facilitated which gave way to the creation of Engineering Word List (EngWL). The word list consists of the properties those of technical and semi-technical engineering lexis. The objectives of the study were summarized as(1) To develop a pedagogic engineering corpus; (2) To create a specific Engineering Word List (EngWL) from the developed corpus. Content analysis (corpus linguistics approach) was used as the research design for the word list creation process. For the purpose of text analysis and concordance, WordSmith Tools Version 5.0 and RANGE were used. As a result, a corpus with the size of 391,505 words (15,619 types) and the EngWL with 841 word families (1,704 types) were created. Hence, lexical products can be utilized by other researchers in the effort of bridging the gaps of secondary and tertiary education.

Article History
Article Received: *5 March 2019*
Revised: *18 May 2019*
Accepted: *24 September 2019*
Publication: *16 December 2019*

Keywords: *Engineering language; engineering English Word List; engineering lexis; engineering corpus; engineering vocabulary.*

1. Introduction

To compete in the current job market, Malaysians, especially students need to achieve certain level of proficiency which can be challenging for many (Tourres, 2011). According to the Star newspaper, Malaysian undergraduates' English standard

November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

is still not up to the required band (The Star, 2011, November 7). Perhaps, the interference and influence of the first language among students causes the deterioration of the English proficiency standards. According to Lee at al. (2012), in a multilingual country like Malaysia , lower proficiency learners tend to adopt code-switching in order to fully communicate their intended messages. In addition, their anxiety level, while they orally exchange their thoughts could be overwhelming especially where evaluation was concerned (Chan, Ain Nadzimah Abdullah &Nurkarimah Binti Yusof, 2012). At university level, skills in writing for specific purposes for any discipline was an uphill task for many undergraduates (Mariam Mohamed Nor et al., 2012) in which the evidences show Malaysian students may lack the functional and academic-related vocabulary. According to Gilmore and Millar (2018), to improve the odds for the students to learn and acquire the intended lexis to function well especially in writing and or even speaking, the introduction of the selected target vocabulary or lexical items to student in the right context could be the answer. A corpus of any kind represents the real language in terms of its 'trend' and pattern usedin the target discourse (Nartey &Mwinlaaru, 2019) and the development of a corpus and word list would provide all the vital vocabulary to fulfill the students' need in academia and even in the society (Gilmore & Millar, 2018; Todd, 2017). In the field of science, when students know more of the required words for science discourse, the more prepared they are to solve their given tasks (Menon &Mukundan, 2012).

In this study, the developed corpus is pedagogic due to the nature of textbooks which incorporated lots of pedagogical characteristics especially the nature of the language. The nature of the language being studied is limited to that of enhancing the students' linguistic capabilities in using the target vocabulary as required to comprehend the engineering concepts in English language. The researchers aim to develop a pedagogic corpusin engineering discourse context as well as a specialized word list to contribute to pedagogy for both secondary and tertiary level. In addition, it could provide valuable lexical guidelines for English for Engineering Purposes (EEP) textbook writers in choosing the right words to be included in their materials to enhance better understanding for various levels. For teachers, the specific word list can serve the immediate application for EEP pedagogy itself. In short, having the right amount of receptive vocabulary would provide ideal understanding and vocabulary production for students especially in the context of English for Specific Purposes (ESP) (Lewis, 1993; 1997).

In various studies carried out, development of field specific word list is required for numerous meaningful applications. The Student Engineering English Corpus (SEEEC) of two million running words in size was created by Mudraya (2006). SEEEC consists of 1,260 word families with 8850 types of sub-technical words for technical engineering students to learn better, bridging the linguistic gap found in the target students. Ward (2009) developed a basic engineering word list (BEL) of 299-word types consisting non-technical words using tertiary level engineering materials for Thai learners with low English proficiency level to help them read engineering specialist textbooks. However, these two corpora could not be incorporated into our syllabus as the Malaysian ESP students would have known most of the basic knowledge of English variations and derivations such as the comparable adjectives and plural-singular. Most importantly, the two corpora could not represent the language use in Malaysian context especially in secondary vocational schools or even tertiary the level education.

Good background knowledge in a specialized subject is important. Learning the technical words can be challenging as their meaning is very specific to a particular discourse (Nation, 2001). By comprehending the discourse-specific words would allow learners to be associated into their discourse community much more effectively (Menon, 2009). The target word list is believed to be a good guide for both students and teachers who require specific engineering words to function well in the EEP classrooms. Lastly, the word list serves as an important bridging agent for learners to become professional engineers or further their studies at tertiary level in the field of engineering.

The objectives of the study are below:

(1) To develop a pedagogic engineering corpus;

(2) To create a specific Engineering Word List (EngWL) from the developed corpus.

2. Method

The process of data collection in this study commenced by scanning the texts which were already converted into a Tagged Image File (TIF) format. There were then saved and transferred to the Optical Character Recognition (OCR) software which converted all the TIF files into text (.txt) files. Nowadays, scanners are sophisticated as they come with built-in OCR software that made the conversion of scanned files into text (.txt) files easier. The text files were then saved and renamed according to the respective textbooks after they were checked manually for errors. Minimal rectification and correction were made to the scanned pages. It was ensured that the scanned texts did not lose their originality. Finally, the text files were analyzed using the software (WordSmith Tools 5.0 and RANGE).The data collection procedure in this study adapts and adopts procedures carried out by previous researchers in identifying essential vocabulary for pedagogic purposes (Coxhead & Hirsh, 2007; Wang, Liang & Ge, 2008; Mukundan& Ng, 2012; Ng et al., 2013; Gilmore & Millar, 2018). As for compiling the pedagogic mini corpus of written English in Malaysian engineering textbooks, the researcher followed the procedure carried out by Mukundan and Aziz (2006; 2007), Menon and Mukundan (2010; 2012).

2.1. Word List Development Process and Word Selection Criteria

The process engages the comparison with a larger reference corpus (at least 5 times bigger) method (Nation, 2001; Berber-Sardinha, 2002; McEnery, Xiao &Tono, 2006), where the chosen texts were analyzed against a larger corpus to determine the essential or pedagogic words needed. Since the operational definition of technical words is not widely and empirically researched, one way to determine such technical vocabularies is to compare their frequency and range in a specialized text with their frequency in a much greater corpus (Chung & Nation, 2003; 2004). Due to the large size of the British National Corpus (BNC) of 100 million tokens, which is far bigger than the target corpus, the BNC was selected as the reference corpus. In addition, the BNC is the ideal reference corpus because "the acrolectal version of Malaysian English is

November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

similar to that of British English" (Menon, 2009, p.14). There are several studies which used the BNC as the chosen reference corpus (see Mukundan& Ng, 2012; Al-Mahrooqi et al. 2011), just to name a few.

Through the function of 'Keyword' in the WordSmith 5.0 programme, once the corpus has been compared with the reference corpus holistically, a raw list of words could be found, taking into consideration only to accept the positive 'keyness' words in the lists. The next criteria to be considered are those that the finalized list of words should not contain any of the words in major word lists like the General Service List (GSL) (West, 1953) and the Academic Word List (AWL) (Coxhead, 2000).The nature of the specialized words should be technical and semi-technical words and thus the English function words must be removed using the RANGE software (Heatley, Nation & Coxhead, 2002).The next step would involve the comparison analysis with the on-line McGraw-Hill Dictionary of Engineering (2008) and the hard-cover version of the McGraw-Hill Dictionary of Engineering (2003) due to the expertise found in McGraw-Hill companies with the wide publication of engineering materials. Should a word from the filtered list matched the searched token in McGraw-Hill Dictionary of Engineering (2003; 2008), then it was categorized as technical engineering word; the rest of the essential words were considered as semi-technical words if they did not match the search. At the final process, the researchers consulted two experts in the filed who were university lecturers in engineering to arrive at the technicality or semi-technicality of some words. The two experts received their PhD degree in the field of engineering and possessed at least 5 years of teaching experience. After the consultation, the word list was compiled.

The process is summarized as below:

I. The frequency of the words should be high enough to be regarded as specialized engineering vocabulary and be compared with a larger reference corpus (British National Corpus).

II. The words should be closely associated to the field of engineering. Thus, the lexis found need to be those found in the General Service List (GSL) (West, 1953) and Academic Word List (Coxhead, 2000) which are considered the benchmark of general and academic English respectively.

III. For technical words, a specialized word identified should match the technical words found in the McGraw-Hill Engineering Dictionary (2003; 2008). Otherwise, it should be considered as a semi-technical engineering word until further cross-checking is executed.

IV. The identified technical and semi-technical words are validated by two experts to be categorized as technical or semi-technical words (inter-coder reliability is accessed). The words which do not meet the mentioned criteria should be eliminated from the analysis.

The final EngWL product consisted of salient engineering word families extracted from the 'keyness' concept. The EngWL words would not have any general or academic use that is expected to be learned by the students prior to learning the technical materials. Thirdly, the EngWL consists of both technical and semi-technical vocabularies, which seem to be the 'gap filling' word list to the existing word lists of the GSL and AWL when students demand higher text coverage to comprehend their

November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

engineering texts. The cross-referencing with the technical engineering dictionary of McGraw-Hill (2003; 2008) was regarded as the preliminary segregation stage in the categorization of the engineering words. The final decision regarding the word selections was made by the two experts in the field of engineering (coders). There were some words suggested by the experts to be removed from the EngWL because they were not technical in nature. Instead, they just functioned as 'general' or low-frequency words

3. Results

Using the 'WordList' function of WordSmith Tools 5.0 (Scott, 2008) ', the lexical properties of the corpus to meet the first objective of the study with the information on the tokens (running words) and types. The created corpus consists of eight Malaysian upper secondary school textbooks in engineering, totaling 391,505 words, which makes it the first corpus of this nature that created in Malaysia. From the corpus, the word list can be produced. This is the first time that such a word list is created for engineering pedagogic purposes and the process of coming up with the word list required the use of RANGE (Heatley, Nation & Coxhead, 2002), apart from the WordSmith Tools 5.0. Firstly, the created corpus file must be saved in 'wordlist' format (.lst). Then, it has to be loaded into the 'keyword' portal of the WordSmith software. The reference corpus which is used for the comparison of corpora is the BNC which is already pre-loaded in WordSmith Tools 5.0 'wordlist' format (.lst). The 'wordlist' file can be easily found on the World Wide Web via the website address 'www.lexically.net'. Both files were loaded into the 'keyword' portal to identify the positive keyness words, when compared to the reference corpus. Positive keyness signals that a particular word in the target corpus is 'unusually' frequent compared to its frequency in the reference corpus. Only positively 'keyed' words are included in the next filtering step in creating the word list. The positive keyness measurement was analyzed using the 'log-chi' formula pre-programmed in the 'keyword' function of the software.

Initially, there were 2,269 types of words from the analysis after removing the non-word characters from the WordSmith 5.0. All the words were 'positive keyness' words, which are unusually frequent as compared to the frequency found in the BNC words. Prior to running the analysis, the default setting of the software was set at retaining 'positive keyness' words. These extracted words were more specific and unique to the target corpus. At this stage, the words were not categorized into word families or lemmatized and filtered accordingly. The word list creation process went through several more stages.

1. Removal of the GSL and AWL words found in the keywords

The extracted words were then inserted to the RANGE programme to remove the GSL and AWL words. At this step, the RANGE programme functioned as 'filtering' software which segregates the words which were similar to that of the GSL and AWL lexis. In the RANGE programme, the GSL word lists are labelled as 'basewrd 1'and 'basewrd2' in the programme folder, each referring to the first list of 1,000 GSL words and the second list of 1,000 GSL words respectively. 'Basewrd 3' represents the AWL word families and the inflected forms which were also available in the software. The list had 570word families and 3,107

November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

word derivatives as prepared by Coxhead (2000).

The filtering process continued but at this point of the research, similar ‘keywords’ found in the GSL and AWL were identified. Only the remaining words from the “Not in the lists” were retained for further analysis. In other words, the remaining ‘keywords’ did not have any shared words or types with those of the GSL and AWL words and derivatives.

2. Comparing keywords with technical dictionary and seeking experts’ opinion

After removing the GSL and AWL words from the analyzed keywords, the retained potential 1354 words (not according to word families) for the list were compared with the on-line McGraw-Hill Dictionary of Engineering, Version 1.0 (2008) and the McGraw-Hill Dictionary of Engineering, 2nd ed. (2003). The approach was adopted from a study by Ng et al. (2013) on creating the Engineering Technology Word List (ETWL). Due to the McGraw-Hill’s long involvement in engineering materials publication, the dictionaries published by the company were selected as authoritative references to the subject matter. The on-line dictionary was one of the pioneering on-line portals for users to obtain the meaning of Engineering terms. Cross referencing with the hard copy of McGraw-Hill Dictionary of Engineering, (2003; 2008) was needed to ensure that the online uploaded database was sufficient and up to date.

Word-to-word examination was carried out to determine whether specific words from the filtered list of engineering keywords appeared in the engineering dictionary. If they did appear, the words were then accepted into the building blocks of the engineering word list and were regarded as technical engineering words. However, the final decision was only made after getting the input from the two coders of this study. If these words did not appear in the engineering dictionary, they were considered as semi-technical words, unless further clarification from the two experts was obtained. Then, two experts from Universiti Tenaga Nasional, which is an engineering-based university, were consulted to see whether or not some of the chosen technical, semi- technical and ‘doubtful’ semi-technical keywords should be eliminated or retained. These words were identified as words that may not be technical or exclusive to the field of engineering. A majority of the acronyms, abbreviations, proper nouns and apparent compounds were eliminated as these word forms are believed to be easily acquired or learned when learners are exposed to them while learning technical lessons (Hsu, 2014). The inter-coder reliability for the coders was found to be high at $p < 0.001$ at 95% confidence interval (Cohen, 1960) with the Kappa Value of, K = 0.82 and high percentage of agreement of 93%. Table 4.17 indicates the results of the inter-coder reliability.

Table 1: Summary of the Inter-Coder Reliability of the Categorization of the Technical & Semi-Technical Words in the Engineering Corpus

Corpus/Word Class	Number of Words Identified	Agreement	Disagreement	Kappa Value, K	Percentage of Agreement (%)

Engineering	1354	1259	95	0.82	93
Technical	105	57	48		
Semi-Technical	832	791	41		
To be Eliminated from the List	417	411	6		

Based on the high inter-coder reliability, it shows that the boundaries between the technical, semi-technical and unwanted words were clear to the coders. The disagreement between the coders were deliberated and moderated before deciding on the final selection. Then, the McGraw-Hill Dictionary of Engineering (2003; 2008), Oxford Dictionary of Science (2010) and Oxford Advanced Learner's Dictionary (2010; 2013) were presented to the coders for further reference and discussion before reaching the final decision. Through using the three dictionaries, the coders were satisfied with the deliberation and decided upon the finalized categorization and came to a 100% agreement.

4. Discussion

A majority of large English written texts consist of more than 2000 most frequently used general words of English (Nation, 2001). An arising concern would be that the percentage of general English words varies across disciplines. According to Nation and Waring (1997), the 2,000 most frequently used English words lead to the text coverage of 78 to 92% in any written texts. However, Coxhead (2000) discovered that the 2,000 most general words provided a much lower text coverage which ranges from 70.7 to 79.1% in her four academic sub-corpora (arts, commerce, law and science). The lowest text coverage was found in the sub-corpus of Science. Ng et al. (2013) reported that only 69.4% of GSL words were found in the text coverage of the engineering technology corpus studied. The dissimilar levels of text coverage found in the corpus of different disciplines warrant more field-specific word lists. The next best text coverage of running words in academic texts would apparently come from the 570 words of the AWL (Coxhead, 2000), encompassing 10% of the text coverage.

Moving from the general and academic words coverage to field-specific word list coverage requires specialized corpus creation. The additional text coverage can bridge the lexical gap for learners in various fields. Konstantakis (2007) created the Business Word List (BWL) without those properties existed in the GSL and AWL. The text coverage provided by the BWL was 2.79% in the Business-related corpus. Similarly, Coxhead and Hirsh (2007) developed the Science-Specific Word List, covering 3.79% of the text coverage in the corpus. In the other related study, Mukundan and Ng (2012) developed the Nursing Education Word List (NEWL) and it proved to be significant with the text coverage of 9.9%. Hence, the text coverage of the EngWL needs to be examined to see whether it can also provide significant coverage to the target corpus. In the present study, the coverage of EngWL for the KBSM **e**ngineering textbooks could be

November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

computed using the RANGE. However, some encoding needed to be done before loading the EngWL to the programme. The encoding included categorizing the EngWL words into word families and its inflected forms and adding '0' next to every word in the EngWL. Then, the inflected forms of each word family should be 'spaced' by pressing the 'tab' button on the computer keyboard.

The analysis of the coverage of the EngWL for the target **e**ngineering textbooks was run using the RANGE software and the results are tabulated in Table 2. The results showed significant text coverage in the sub-corpora of KBSM **e**ngineering textbooks, ranging from 9.42% to 12.07%. From the table, it can be inferred that the percentage of text coverage of EngWL is rather high. Hence, it can definitely serve as a valuable word list to educators and learners in the field of **e**ngineering. It was also found that the EngWL recorded the highest text coverage of Electrical and Electronics (12.07%), followed by Mechanical Engineering (11.60%), Engineering Technology (9.92%) and Civil Engineering (9.42%). despite the differences in the percentage of text coverage, the words from the EngWL were considered to be evenly distributed among the **e**ngineering subjects analyzed, with over 700 matching types except for Engineering Technology (1015 matching types). However, Engineering Technology comprised the greatest number of tokens amongst the four **e**ngineering subjects and this also explains the high number of matching words.

Table 2: EngWL Text coverage in the Sub-corpora

Engineering Subject (KBSM) Form Four and Form 5	Total Tokens	Total Types	EngWL matching words (tokens)	EngWL matching words (types)	EngWL Coverage (%)
Civil	97227	6654	9,158	733	9.42
Electrical and Electronics	84410	6201	10,187	771	12.07
Engineering Technology	115775	7958	11,482	1015	9.92
Mechanical	94093	6206	10,911	755	11.60

EngWL is a word list consisting of 841 word families and 1704 types of words which provide a significant word coverage of 10.65% in the overall engineering texts. It is essentially the highest text coverage that a word list can provide in the field of EEP. The high text coverage provided by the EngWL of the engineering texts narrows the gap of the text coverage of 'unknown' vocabulary for students which facilitate better transition into tertiary engineering programmes.

5. Conclusion

This corpus-based study set out to develop a pedagogic engineering corpus from the KBSM upper secondary engineering textbooks. The data collected

November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

were empirically analyzed using the WordSmith Tool 5.0 and RANGE software, which brought about a series of concordance outputs and tables. The word list, EngWL was created from the target corpus which seems to have significant implications for learners in the field of engineering. The word list is believed to enhance learners' technical or specialized vocabulary which they need to understand engineering texts better. As a result, learners can cope better in comprehending specific details intended for them.

The output of this research provided significant findings and tentative conclusion as follow:

i. A specific engineering corpus was created, consisting 391,505 tokens which could be used for further research or immediate classroom use.

ii. The 841 word families of the Engineering Word List (EngWL) accounted for 10.65% of the entire engineering corpus. This percentage was higher than any other specific (engineering) word lists which have been created and published previously. Most importantly, it was for the first time in the Malaysian secondary education system that an engineering word list was created as the recommended guide for learners in order to comprehend engineering texts better.

6. Acknowledgement

The authors are grateful to UNITEN UNIIG Grant, for the funding of this project.

References

1. Al-Mahrooqi, R., Al – Busaidi, S., Mukundan, J., Ahour, T. & Ng, Y.J. (2011). Can the essential lexicon of geology be appropriately represented in an intuitively written EAP module? English Language and Literature Studies, 1(1), 50-66.
2. Berber-Sardinha, T. (2002). Comparing corpora with WordSmith Tools: How large must the reference corpus be? Retrieved 13 November 2011 from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.7923
3. Chan, Swee Heng and Ain Nadzimah Abdullah, and Nurkarimah Binti Yusof, (2012) Investigating the construct of anxiety in relation to speaking skills among ESL tertiary learners. 3L: The Southeast Asian Journal of English Language Studies, 18 (3), 155-166.
4. Chapman, K., Zuhrin Azam, A., Loh, F.F., Kulasagaran, P., Wong, P. M., Kang, S.C.,…Farah Fazanna, Z. (2011, November 5). PM has been consulted over PPSMI, says Muhyiddin. The Star Online. Retrieved from http://thestar.com.my/news/story.asp?file=/2011/11/5/nation/9846317andsec=nation
5. Cheng, N. (2013, January 28). Groups want independent findings on PPSMI included in education blueprint. The Star Online. Retrieved from http://thestar.com.my/news/story.asp?file=/2013/1/28/nation/20130128150105&sec=nation
6. Choy, S. C., & Troudi, S. (2006). An investigation into the changes in perceptions and attitudes towards learning English in a Malaysian college. International Journal of Teaching and Learning, 18(2), 120-130.
7. Chung, T. M. & Nation, P. (2003). Technical vocabulary in specialised texts. Reading in a Foreign Language, 15, 103-116.
8. Chung, T.M. & Nation, I.S.P. (2004). Identifying technical vocabulary. System 32(2), 251-263.
9. Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.
10. Coxhead, A. & Hirsh, D. (2007). A pilot science word list for EAP. Revue Française de Linguistique Appliqueé XII (2), 65-78.
11. Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34(2), 213-238.
12. Gill, S.K. (2005). Language policy in Malaysia: Reversing direction. Language Policy, 4(3), 241-260.
13. Gilmore, A. & Millar, N. (2018). The language of civil engineering research articles: A corpus-based approach. English for Specific Purposes 51, 1–17

14. Habibah Elias, Rahil Mahyuddin, Nooreen Noordin, Maria Chong Abdullah, & Samsilah Roslan (2009). Self-efficacy beliefs of at-risk students in Malaysian Secondary Schools. The International Journal of Learning, 16(4), 201-209.
15. Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. Available online: http://www.vuw.ac.nz/lals/staff/Paul_Nation
16. Holme, R. (2006). English for science or science for English: Designing content-based materials for language teaching. In J. Mukundan (Ed.), Readings on ELT materials (pp. 17-34). Serdang, Malaysia: Universiti Putra Malaysia Press.
17. Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. English for Specific Purposes, 33 (1), 54-65.
18. John, E. & Aniza Damis. (2008, September 7). Study reveals policy's flaws. New Straits Times, pp. 6-7.
19. Kang, S.C. (2014, January 5). Uproar over premature halt to PPSMI policy. The Star Online. Retrieved from http://www.thestar.com.my/News/Nation/2014/01/05/Uproar-over-premature-halt-to-PPSMI-policy/
20. Kon, Y. H. (2005). Teaching efficacy beliefs of primary school teachers in using English to teach Mathematics and Science, Jurnal IPBA, 3(2), 45-49.
21. Konstantakis, N. (2007). Creating a business word list for teaching business English. Estudios de Lingüística Inglesa Aplicada, 7, 79-102.
22. Lee, Y. L., Ng, Y.J., Chong, S.T. & Mohd Ariff Ahmad Tarmizi (2012). Code switching in Sepet: Unveiling Malaysians' communicative styles. British Journal of Arts and Social Sciences (BJASS), 6(2), 166-181.
23. Leech, G. (1991). The state-of-the-art in corpus linguistics. In K.Aijmer, & B.Altenberg (Eds.), English corpus linguistics, studies in honour of Jan Svartvik (pp.8-29). London, UK: Longman.
24. Leech, G. (2002). The importance of reference corpora. Hizkuntza-corpusak. Oraina eta geroa, 10(24/25), 1-11.
25. Lewis, M. (1993). The lexical approach, the state of ELT a new way forward. Boston: Thomson Publishing.
26. Lewis, M. (1997). Implementing the lexical approach, putting theory into practice. Boston: Thomson Publishing.
27. Mariam Mohamed Nor, Ng, Y.J., Lee, Y. L., Chong, S.T. & Mohd Ariff Ahmad Tarmizi (2012). "What they can see, they can write, right?" Global Journal of Human Social Sciences (GJHSS), 12 (10), Version 1.0, 64-72.
28. McEnery, T., Xiao, R., & Tono, Y. (2006). Corpus-based language studies. An advanced resource book. London: Routledge.
29. McGraw-Hill Dictionary (Digital Edition 1.0). (2008). Retrieved from http://mcgraw-hill-dictionary-of-engineering.software.informer.com/download/.
30. McGraw-Hill Dictionary of Engineering (2nd ed.). (2003). New York, NY: McGraw-Hill Companies.
31. Menon, S. & Mukundan, J. (2010). Analysing Collocational Patterns of Semi-Technical Words in Science Textbooks. Pertanika Journal of Social Sciences and Humanities, 18 (2), 241-258.
32. Menon, S. & Mukundan, J. (2012). Collocations of high frequency noun keywords in prescribed science textbooks. International Education Studies, 5 (6), 149-160. doi:10.5539/ies.v5n6p149.
33. Menon, S. (2009). Corpus-based analysis of lexical patterns in Malaysian secondary school Science & English for Science & Technology textbooks. Unpublished doctoral dissertation, Serdang: Universiti Putra Malaysia, Malaysia.
34. Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. English for Specific Purpose, 25, 235-256. doi:10.1016/j.esp.2005.05.002
35. Mukundan, J. & Ng, Y.J. (2012). Development of a technical nursing education word list (NEWL). International Journal of Innovation in English Language Teaching and Research, 1(2), 105-124.
36. Mukundan, J. (2004). A composite framework for ESL textbook evaluation. Unpublished doctoral dissertation, Serdang: Universiti Putra Malaysia, Malaysia.
37. Mukundan, J. and Aziz, A. (2006). Automatic and semi-automatic processes of WordSmith 3.0 as a textbook evaluation instrument: A case study. TEFLIN Journal, 17 (2), 200-219.

November-December 2019
ISSN: 0193-4120 Page No. 3232 - 3242

38. Mukundan, J. and Aziz, A. (2007). A forensic study of vocabulary load and distribution in five textbooks within an entire Malaysian secondary school English language programme (Forms 1-5). Pertanika Journal of Social Sciences and Humanities, 15 (2), 183-198.
39. Nation, I.S.P. & Waring, R. (1997). Vocabulary size, text coverage & word lists. In N. Schmitt & M.McCarthy (Eds.), Vocabulary: Description, Acquisition & Pedagogy (pp. 6-19). Cambridge, UK: Cambridge University Press.
40. Nation, I.S.P. (2001). Learning vocabulary in another language. Cambridge, UK: Cambridge University Press.
41. Nartey, M.&Mwinlaaru, I.N. (2019). Towards a decade of synergizing corpus linguistics and critical discourse analysis: A meta-analysis. Corpora, 14(2), in press
42. Ng, Y. J., Lee, Y.L., Chong, S.T. Nurhanis Sahiddan, Philip, A. Noor Hafiza Nor Azmi & Mohd Ariff Ahmad Tarmizi. (2013). Development of the engineering technology word list for vocational schools in Malaysia. International Education Research, 1(1), 43-59.
43. Ng, Y.J., Chong, S.T., Mariam Binti Mohamed Nor, Mohd Ariff Ahmad Tarmizi &Alif Fairus Nor Mohamad (2012). Corpus based analysis of the TOEFL course books: What are the words we should teach our students? International Review of Social Sciences and Humanities (IRSSH), 3(2), 152-160.
44. Noraini Idris, Loh, S. C., Norjoharuddeen Mohd. Nor, Ahmad Zabidi Abdul Razak and Rahimi Md. Saad. (2007). The professional preparation of Malaysian teachers in the implementation of teaching and learning of Mathematics and Science in English, Eurasia Journal of Mathematics, Science & Technology Education, 3(2), 101-110.
45. Oxford Advanced Learner's Dictionary (8th ed.). (2010). Oxford, UK: Oxford University Press.
46. Oxford Advanced Learner's Dictionary (online version). (2013). Retrieved from http://oald8.oxfordlearnersdictionaries.com/
47. Oxford Dictionary of Science (sixth edition). (2010). Oxford, UK: Oxford University Press.
48. Pandian, A. & Ramiah, R. (2004). Mathematics and science in English: Teacher voice. The English Teacher, 33, 50-61.
49. Pavithiraa, R. (2014, January 6). Students in PPSMI to continue learn in English. The Edge Online. Retrieved from http://www.theedgemalaysia.com/political-news/270096-malay-press-students-in-ppsmi-cohort-to-continue-to-learn-in-english.html/
50. Scott, M. (2008). WordSmith Tools version 5. Liverpool: Lexical Analysis Software.
51. Singh, P., Arba Abdul Rahman & Teoh, S.H. (2010). Languages and mathematics achievements among rural and urban primary four pupils: A Malaysian experience. Journal of Science and Mathematics Education in Southeast Asia, 33 (1), 65-85
52. Sopia Md Yassin, Marsh, D., Ong, E.T., & Lai, Y.Y. (2009). Learners' Perceptions towards the Teaching of Science through English on Malaysia: A Quantitative Analysis. International CLIL Research Journal, 1 (2), 54-69.
53. The Star. (2011, November 7). Don: English standard of undergraduates still not up to par, p.6.
54. Todd, R.W. (2017). An opaque engineering word list: Which words should a teacher focus on? English for Specific Purposes, 45, 31-39.
55. Tognini-Bonelli, E. (2001). Corpus linguistics at work: Studies in corpus linguistics, v. 6. Amsterdam: John Benjamins.