International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 14, Issue 1, 2020

Essential Technicality Aspects (ETA) in Engineering Language Collocations

Y.J. Ng[a*], S.T. Chong[b], S. Thiruchelvam[c], M.F. Chow[d], J. Karthikeyan[e],
[a,b,c,d]Universiti Tenaga Nasional, Selangor, 43000, Malaysia, [e]Vellore Institute of Technology, Tamil Nadu, 632014 India, Email: [a*]yujin@uniten.edu.my

In the field of engineering studies or English for Engineering Purposes (EEP), noun compound or technical word collocation pose a significant challenge to learners due to the arbitrariness of the language. Understanding how collocation patterns or noun strings could appear provide more concrete understanding for learners in the engineering discourse community. This study adopts the principle that 'chunks of language', which are less predictable than those of idioms, should be identified in order to facilitate better pedagogy. The analysis carried out in this study was limited to the 'immediate two-word noun compounds or lexical collocation', believed feasible to the learners. The analysis adopts the frequency-based approach in identifying the meaningful collocations which take the word or lexical coverage of the co-existing words into consideration. In other words, the 'mechanism' of how the words work together or co-exist to bring about the meaning is the core emphasis of this study.

Key words: *Collocation, Noun Compound, Engineering Language, Dictionary Skills, Words Technicality.*

Introduction

Collocation can be defined as the nature of two or more words which co-exist in a discourse, and are lexically fixed or that syntactically belong to a specific extent (Tognini-Bonelli, 2001; Schmitt, 2000). Collocation is additionally expressed as co-selected words (Lewis, 1997); clusters of words (Scott, 1996); idioms (Sinclair, 1991) and prefabricated language (Pawley & Syder, 1983). Studies on the acquisition of prefabricated language have shown that it establishes a significant effect on language learning (Schmitt, 2000; Nattinger & De Carrico, 1992) despite the criticism by Krashen and Scarcella (1978). However, Nattinger and De Carrico (1992) contended that most of the languages that people are exposed to are repetitive and predictable, allowing students to produce 'innovative' new expressions. The concept of

language acquisition through collocation is well documented in the 'idiom principle' suggested by Sinclair (1991).

The collocation study is an extended study of Sinclair's (1991) 'idiom principle'. That is, the meanings in a language are fabricated by chunks of language that can be predictable without fixed sequences of morphemes (Sinclair, 1991). Sinclair adds on describing collocations as the non-random nature principles of idioms with an open-ended string of choices or combinations to express the intended meanings. A learner is expected to interpret a part of language used as a series of individual items in the form of chunks. According to Sinclair (1991), the co-existed word in a common phrase has a fixed or specific meaning associated to the entire phrase rather than the individualised words being coded and decoded as single entities. On the contrary, lexis and grammar should be viewed as separate entities (Ma & Qian, 2020). Linguistics experts advocate the sheer importance of teaching and learning collocations in classrooms and at higher education institutions (Flowerdew, 2012; Menon & Mukundan, 2012; Ward 2007; Hill, 2000; Hunston & Francis, 2000). The combinations in collocations are fundamentally predictable as Hill (2000) contends the learning of the combinations of language chunks would facilitate better comprehension of a subject matter which ranges from non-technical to technical language.

On a related note, Hoey (2005) puts forward the notion of 'lexical priming' which shares the same essence that collocation in fact prompts students to comprehend textual record more systematically. Hunston and Francis (2000) suggest that retention and recall of prefabricated units of a language help improve the memory power of the students while learning. Collocation facilitates thinking and efficient communication as the brain memory captures the prefabricated chunks of words better (Pawley & Syder, 1983). Gledhill (2000) did a content analysis study on cancer research articles and reported that the predictability of the meaning collocated word was largely dependent on what was acceptable and appropriate within the discourse community in expressing certain terms or ideas. Similar theories are stated in Hoey's (2005) 'the lexical priming' and Lewis's (1993; 2001), 'the lexical approach' concepts and framework.

The theory embraces the notion that an individual word links a particular meaning to words or phrases. The linked meaning is based on the intrinsic meanings as well as the previous context which is often encountered in a learning process. Hoey (2005) added that students within the similar discourse community will produce or share the same ideas on certain collocations. In his corpus, he categorised it into four main aspects of semantic association, which comprise 90% of all the adjectives that modify the keyword 'consequence'. The corpus included 95 million words of 'Guardian' newspapers, 3 million words of written texts and 230,000 words of spoken data retrieved from the British National Corpus (BNC). With the percentages in descending order, the main four semantic categories were 'logical consequence' (59%), 'doleful consequence' (15%), 'serious consequence' (11%) and 'surprising consequence'

(6%). Through lexical analysis, there exists a realisation that lexical items or lexicon are the core to language use. Lewis's (2000) "Lexical Approach" stressed the essence of vocabulary or lexis in language pedagogy, where the principle clearly defines that language consists of "grammaticalised lexis, not lexicalised grammar" (p.51) and "lexis is the core or heart of the language" (p.89). Lewis emphasised that grammar is limited by lexical choices. The building blocks of language learning are mainly the knowledge of lexis and word combinations (Lewis, 1993; 2001). This is in contrast to the standard view that language primarily consists of structured grammar and then vocabulary (Sinclair, 1991; Ma & Qian, 2020). The Lexical Approach asserts that "language consists of chunks which, when combined, produce coherent text" (Lewis, 2001 p.7).

The principle of the Lexical Approach is that grammar is not as important as lexis which needs to be learnt through multi-word units such as collocations. Collocations can pose problems to non-native speakers, who have a good command of grammar but are unable to remember words and collocations accurately (Nattinger & De Carrico, 1992; Lewis, 2001; Ward, 2007). The Lexical Approach highlights the prominence of authentic or real language input which is the real language that is 'used' (corpus-based). Lewis (2001) proposes that instead of focusing on rules and accuracy, the generalisability and the various ways in which a language can be used should be emphasised. Hill (2000) states that approximately 70% of what people express in words, comprehend, write and listen are found in certain prefabricated expressions. The frequency-based approach describes the existence of collocated words that are based on the frequency measures (Nesselhauf, 2004). Moon (1998) considers all the co-occurrences as collocations; however, Kennedy (1998) and Nesselhauf (2005) agree on the co-occurrence of only more than once in a corpus as the crucial criterion. The most frequent collocations were found to be useful because students have more chance to encounter them in their language learning (Durrant, 2008; Shin & Nation, 2008; Ward, 2007). Thus, this study suggests the uses of the frequently occurring specialised engineering words as the keyword for the collocation, by engaging the concordance Software of WordSmith Tools 5.0 (Scott, 2008).

Collocation is also widely known as noun compound. Noun compounds are widespread in science and technology texts and other ESP areas, often causing problems for non-native learners (Master, 2003; Trimble, 1985). The difficulty in understanding noun compounds is due to the inability in decoding them. Learners might make out the meaning of the two words 'dish' and 'cloth' but may not decode that "the collocation 'dishcloth' means a cloth used for drying" (Master, 2003, p.3). Furthermore, the frequent types of noun compounds would normally comprise only two bases and be classified as 'simple' noun compounds (Trimble, 1985). However, even some 'simple' noun compounds do not follow the convention of decoding or paraphrasing the meaning of noun compounds, as shown in Table 1.

Table 1: The samples of 'simple' noun compounds with the meanings of words (see Trimble, 1985, p.133)

Noun Compounds	Meaning or definition
Metal shaft	A shaft made of metal
Metal spring	A spring made of metal
Metal cutter	Not a cutter made of metal, but an instrument used to cut metal

Thus, the research aims to identify the meaningful relationship between the various word structures or collocations and their meanings in the noun compound using identified engineering related words. These noun compound forms are derived from the highest frequency keywords from the concordance software.

Methodology

Concord is a concordance programming using Disk Operating System (DOS), Text Only (.txt), American Standard Code for Information Interchange (ASCII) or American National Standards Institute (ANSI) text files found in WordSmith Tools 5.0. By selecting a search word, the Concord will run a search on all the chosen text files, and then present a concordance display and provide information about the word families. It demonstrates the ways or how frequent words are distributed and recycled. For instance, the words 'play', 'plays', 'played', 'playing', 'player' and 'players' are displayed when searching the keyword 'play'. The tool then shows how the word 'play' is utilised in various ways throughout the text, which can be saved for further use. This enables the identification of word collocates and common phrases, known as word clusters. Concord enables the researchers to determine patterns in language use (Mukundan, 2009). In this study, Concord was used to identify the most frequent collocations found in the engineering corpus. The result was later used in determining patterns and semantic relations of the noun compounds. 'Concord' allows the researchers to find certain word(s) in all the text files chosen, presenting a concordance display and giving access to information about collocation. This function was used in the study to identify the collocations which in turn translated into the essential noun compounds used in the field of engineering. The approach used to identify the phrase collocation (noun compounds) is the frequency-based approach in which a collocation is the co-occurrence of words at a certain range, based on frequency (Hunston, 2002; Nesselhauf, 2004). Pertaining to manageability and feasibility issues, this research only identifies the 2-word noun compounds using some engineering related words to highlight its word or vocabulary connection.

The noun compounds were identified and extracted by looking up each noun phrase collocations in the dictionaries, namely, the McGraw-Hill Dictionary of Engineering (2003; 2008) and Oxford Advanced Learner's Dictionary (2010; 2013) to arrive at the appropriateness level of the noun compounds combinations. The 2-word phrase collocations were then

classified as noun compounds if they could be found with their intended meaning in either of the dictionaries. The meaning of the noun compounds identified is checked to see if they contain engineering-oriented meanings such as 'control circuits' which means 'a circuit that controls a controlled system that is directly measured or some functions of a machine, device, or piece controlled'. Only engineering-oriented compound nouns, such as 'logic array', 'inductance coil' and 'resonant capacitor' were analysed as this would identify the phraseology which is typical of engineering English. The McGraw-Hill Dictionary of Engineering (2003; 2008) was used as it contains occurrences from real language (authentic use in engineering). The Oxford Advanced Learner's Dictionary (2010; 2013) was also used to identify the engineering terms with some detailed general English explanations, which could be understood by non-engineers. In addition, the dictionary was used to ascertain whether some engineering terms were missing from the more technical McGraw-Hill Dictionary of Engineering (2003; 2008). Each selected noun compound was then analysed to look at the position of the keyword in the collocation and to identify the lexical patterns or phraseology found. The extended meanings and the technicality levels of the noun compounds were also examined.

Results and Discussion

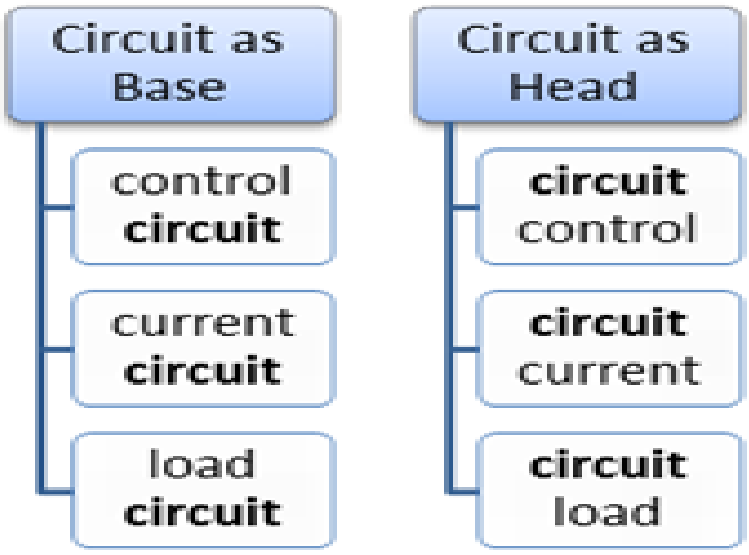
It is often debated that inclusion of salient high frequency items of a target language in the instructional materials provide learners with the best learning dividend (Nation, 2007; 2001). This is because the probability of coming across these words in the future is substantially high (Flowerdew, 2012; Mudraya, 2006; Nation & Waring, 1997; Romer, 2010; Ward, 2009). For example, Mudraya (2006) identified the 100 most frequent word families in the developed wordlist of the Student Engineering Word List. Similarly, Gardner and Davies (2007) presented the top 100 English phrasal verbs in the BNC. Shin and Nation (2008) coded the most frequent collocations in spoken English, and Ward (2009) listed the 299 most frequent basic engineering words. Furthermore, Martinez and Schmitt (2012) provided the most frequent phrasal expressions list. Alif Fairus Nor Mohamad and Ng (2013) offered the top 100 most frequently used vocabulary in nursing and Yang (2015) created a nursing word list to bridge the vocabulary gap for the nursing students. Using the similar concept and approach, keywords with high frequency appearing in disaster risk reduction framework (Sendai Framework), were identified as well (Thiruchelvam et al., 2019). This study viewed noun compounds as a major area of discussion, as it creates problems for engineering students. This is due to their language arbitrariness, as the apparent combinations take on extended and normally, technical meaning. In identifying collocations, a keyword may maintain the same position. However, the diverse noun compounds formed by the same keyword may have different meanings.

In this study, the word 'circuit' was selected and analysed in order to highlight the language arbitrariness of the noun compounds in this field. The compounds 'control circuit', 'current circuit' and 'load circuit' - all with the same syntactic combinations of 'noun + noun' (N + N)

- holds the keyword 'circuit'. In the context of this study, the word 'circuit' means "the complete path of wires and equipment along which an electric current flow" (Oxford Advanced Learner's Dictionary, 2013). Similarly, the meaning of the word 'current' does not refer to 'of the present time'. Rather, it relates to 'the flow of electricity'. Looking at the three combinations of compounds, students are supposed to apply Hoey's (2007) 'lexical priming theory', and realise that these compounds are a type of circuit or have the properties of 'controlling', 'loading' and 'electrical flow'. The theory embraces the notion that an individual links particular meaning to words or phrases based on the intrinsic meanings as well as the previous contexts which are often encountered in a learning process. Hoey (2005) also stated that students or members of the same program or discourse community will share identical priming.

The students might guess the same meaning as the earlier meaning when circuit was at the base position. From this study, it is interesting to note that the words 'control', 'current' and 'load' can take positions that is, the head and base of the keyword 'circuit'. In other words, 'circuit control', 'circuit current' and 'circuit load' literally existed and they carried out different functions when compared with the former sets of noun compounds discussed. The combination of structures of the identified keyword 'circuit' is shown in Figure 1. Thus, students are partially accurate in their 'priming' endeavour because they need to work out the specific meaning in the midst of the language arbitrariness found in the noun compounds. In order to identify the meaning of the noun compounds the researcher used the three dictionaries listed earlier as a reference. In addition, the researcher consulted experts for their opinion to moderate and reach a consensus, as some of the meanings of the noun compounds could not be found in the dictionaries.

Figure 1. The combinations of 'circuit' as the key word of the noun compounds



The specific and brief meaning of each noun compounds with 'circuit' being the base word is given in Table 2. The meaning and limitation of these noun compounds are abstract, and students need very specific details to fully comprehend the concepts.

Table 2: Noun Compounds with 'Circuit' as Base

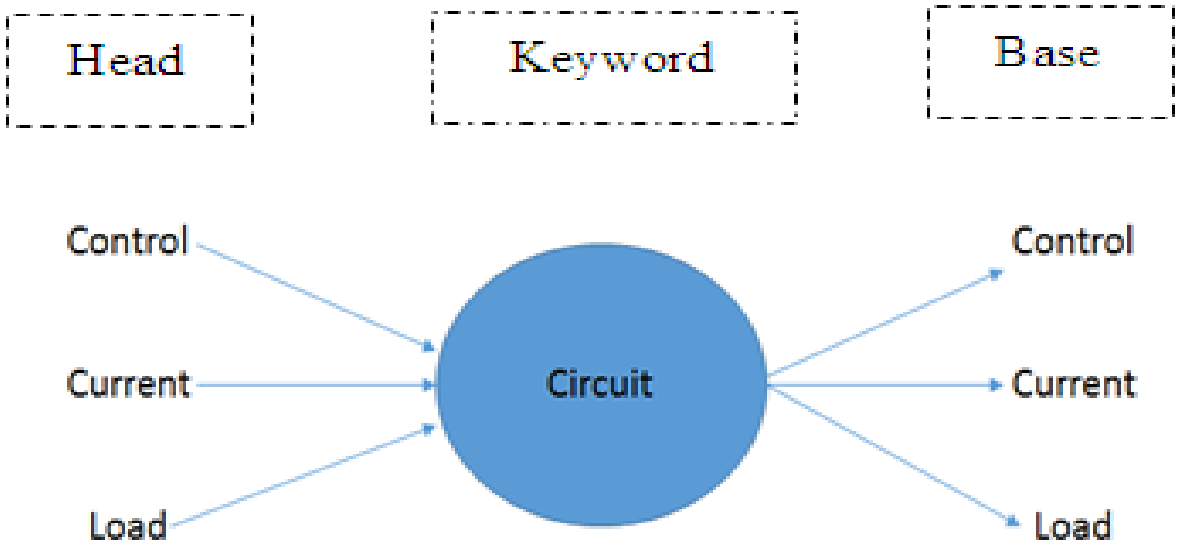
Noun compound of circuit	Meaning or definition
1. Circuit control	1. The act of controlling the circuit (with specific panels), it refers to the level of intelligence or sensitivity of the circuit to execute the predetermined task.
2. Circuit current	2. The flow of electric charge in the circuit (the current of the circuit, often abbreviated as *I*.)
3. Circuit load	3. An electrical component or portion of a circuit that consumes electric power in the circuit, for example light bulb or resistor.

The specific and brief meaning of each noun compounds with 'circuit' being the head word is shown in Table 3. Similarly, they carried out specific meanings in the field of engineering and they are clearly different from those given in Table 2. Figure 2 summarises the collocation patterns of the keyword 'circuit' in the target corpus, which is recognised to collocate with the identical nouns, found in the 'head' and 'base' position (but carrying out different technical meanings).

Table 3: Noun Compounds with 'Circuit' as Head

Noun compound of circuit	Meaning or definition
1. Circuit control	1. The act of controlling the circuit (with specific panels), it refers to the level of intelligence or sensitivity of the circuit to execute the predetermined task.
2. Circuit current	2. The flow of electric charge in the circuit (the current of the circuit, often abbreviated as *I*.)
3. Circuit load	3. An electrical component or portion of a circuit that consumes electric power in the circuit, for example light bulb or resistor.

Figure 2. The keyword 'circuit' with the possibilities of identical nouns appearing in the 'head' and 'base'

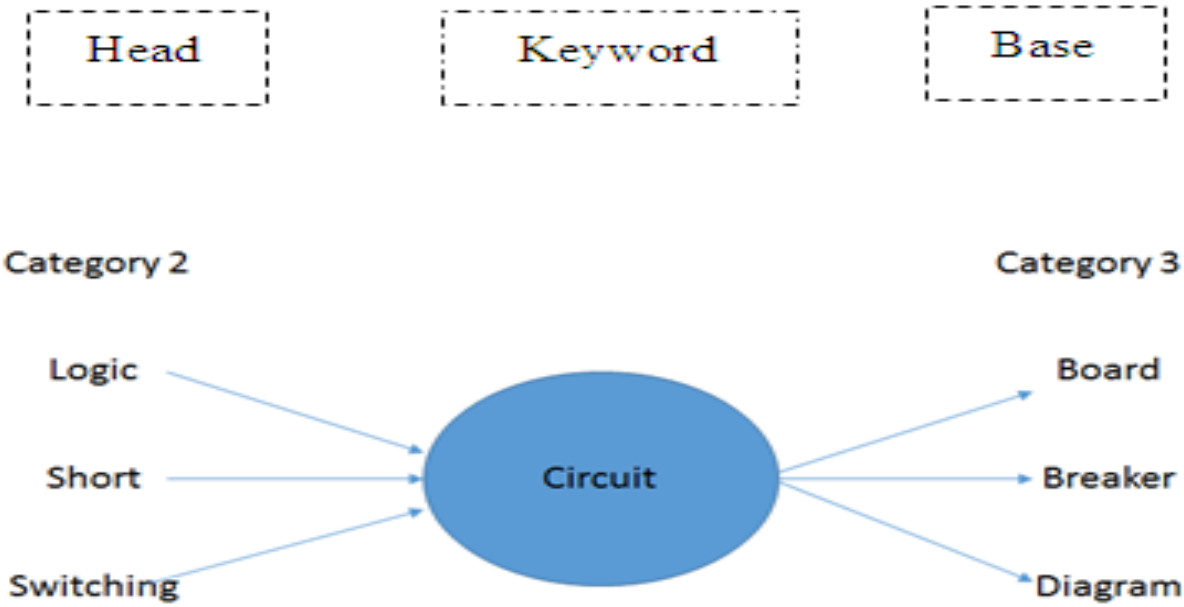


Pertaining to these findings, the 'flexibility' of the syntactical positioning (head – base) of the keyword 'circuit', in fact poses great difficulty to students due to the subjective nature of these noun compounds. There are no grammatical or definite patterns for students to follow in order to manage the learning of these noun compounds. Perhaps, unless these students are experts in the field, they can decipher the meanings of the noun compounds. This can cause problems to students who have long way to learn highly specialised lexis in the field, on top of not possessing the required academic (Coxhead & Hirsh, 2007) or high frequency vocabulary (Ward, 2009). This problem is often overlooked, and the consequences can be costly, more with the lack of familiarising with dictionary definitions. Students with higher confidence may try to create personal meaningful definitions using 'word networks' of the target subject, where they can relate to the technical subject matter (Bruce, 2011). Unfortunately, the less confident ones may ignore the specific piece of unidentified information or terms in their discourse, which may produce more learning problems for them (Arnold, 1999; Bruce, 2011; Nation, 1990; 2001).

The notion of early exposure and introduction of noun compounds to learners in specific fields is advocated because each compound requires detailed knowledge about the subject matter (Cowan, 1974; Trimble, 1985). From the findings, the most frequent 123 keywords were used as a base to identify noun compounds patterns of the engineering texts, and three apparent categories were recognised. The first category (Category 1) has been discussed above, where the forming words can be positioned as both head and base of the target keywords. Each position provides holistically different technical meanings. The second compounds category (Category 2) only exists as a pre-modifier to target keywords. Basically, this means that the forming compound word takes the head position. Finally, the third compounds category

(Category 3) are those forming words which only exist as post-modifiers of target keywords, taking the base position. Category 2 and Category 3 can be explained through using the same keyword, 'circuit', but this time it collocates with a variety of other nouns both in the head and base position as shown in Figure 3.

Figure 3. The keyword 'circuit' with the possibilities of non-identical nouns appearing in the 'head' and 'base'



In the second category of noun compounds found in this study, students would assume that the compounds 'logic circuit', 'short circuit' and 'switching circuit' would have the syntactic combination of 'noun + noun' and 'adjective (with present participle) + noun'. All three compounds could be assumed to be types of circuits, which is not entirely incorrect. However, by definition (McGraw-Hill Dictionary of Engineering, 2003; 2008), 'logic circuit' related to the computational electronics components which perform logical operation on data, and then put into a computer. It does not solely refer to the 'science of thinking' of a circuit. 'Short circuit' is defined as "a low-resistance connection across a voltage source or between both sides of a circuit or line, usually accidental and usually resulting in excessive current flow that may cause dam-age" (McGraw-Hill Dictionary of Engineering, 2003, p.496).

Based upon the obtained meaning, 'short circuit' does not literally mean 'small in height' circuit. Lastly, 'switching circuit' refers to "a constituent electric circuit of a switching or digital processing system which receives, stores, or manipulates information- in coded form to accomplish the specified objectives of the system" (McGraw-Hill Dictionary of Engineering, 2003, p.548). Perhaps, some students might guess the meaning, but do not know the specific details. The meaning of the noun compounds identified is far greater than a mere understanding

of the English grammatical rules. The combination of words takes on extended technical meanings in the field of engineering and some noun compounds can be highly technical, such as 'pneumatics circuit'. For this reason, grammatical rules of general English cannot be directly applied when determining the meaning of noun compounds in the language use of engineering.

As for the noun compounds in Category 3, such as 'circuit board', 'circuit breaker' and 'circuit diagram', the keyword 'circuit' is in the head position. Hence, it works as a pre-modifier of the base word. Assuming that grammatical rules of general English are applicable to infer the meanings of these compounds, the compounds could be understood as types of 'board' and 'diagram'. In addition, 'breaker' is often related to 'a person or thing that performs the act of breaking'. This time, the transfer of grammar rules can be applied with some degree of accuracy. However, definitions or characteristics of the noun compounds should be noted by students, so that they can easily understand subject matters. Students should realise that a 'circuit diagram,' is indeed a type of diagram, but the diagram should contain standardised engineering symbols with suitable interconnection arrangement. This is often referred to as schematic diagram as well (McGraw-Hill Dictionary of Engineering, 2003). In addition, a 'circuit board' should be a board which has a characteristic of integrated chips with a variety of interconnecting electronic components, whose back is printed with electrically conductive pathways.

Finally, a 'circuit breaker' is "an electromagnetic device that opens a circuit automatically when the current exceeds a predetermined value" (McGraw-Hill Dictionary of Engineering, 2003, p.102). Although students may be able to guess the grammatical pattern of these compounds, it should be noted that these noun compounds can be defined and understood with specific details or characteristics. In the case of 'circuit breaker', students might guess that it is a device to disable the circuit. Nevertheless, students need to realise that it is not just any device but an 'electromagnetic' one. It only 'breaks' the circuit given the condition that the current (electricity flow) exceeds the 'break-point'.

Furthermore, when it is in compound form, non-technical words such as 'board', 'diagram' and 'breaker' have been transformed into more technical words with extended meanings, expressing 'a single idea noun'. These noun compounds now acquire extended meanings more specific to the field of engineering and more technical in hierarchy. In their analysis, Mukundan and Menon (2007) and Menon and Mukundan (2010) found that there is a tendency for nouns to acquire extended meanings when they are in compound form. In addition, there is no definite hierarchy pattern for students to refer to when identifying noun compounds. The collocation patterns of the noun compounds such as 'circuit board' and 'circuit diagram' showed that the nouns can combine 'freely' to form compound nouns, which may be technical or semi-technical in nature. Furthermore, there is no fixed set of grammatical patterns to guide students in identifying the meaning of the noun compounds. Without having a fixed grammatical

pattern, decoding the noun compounds can pose great confusion and difficulty to students in learning engineering subjects in schools (see Gilmore & Millar, 2018; Thiruchelvam et al., 2018; Green & Lambert, 2018).

Pedagogical Discussion on Noun Compounds

The complexity, arbitrariness and the extended meanings that noun compounds acquire in the engineering environment should be explicitly taught to students. It is crucial to identify the types of words learners should be exposed to and how the teachers can focus in delivering them (Todd, 2017). The researchers believe that the amount of exposure given to noun compounds in engineering textbooks is insufficient for Malaysian students. This might have resulted from the fact that limited engineering discourse or corpus-based studies were carried out in Malaysia (Ng et al., 2019). Exercises which require students to identify noun compounds should be done regularly. Students need to define the noun compounds through group or pair work or even form noun compounds by inserting various prefixes or suffixes. These noun compounds should be treated as single words and new vocabularies because extensive inference is needed along with important empirical experiences.

Students need to be educated on why grammatical rules of general English language cannot be applied to all situations to infer the meanings of compounds. This was apparent in compounds like ‘load circuit’ or ‘circuit load’, which could be confusing at times. Students need enormous exposure on noun compounds, especially those which have more technical form and require extended meanings specific to the engineering field. These noun compounds should be learned using various memory tasks which include making a list of complex noun compounds. The semantic categories were formed based on cross-reference analysis (WordSmith Tools 5.0) and the definitions given by McGraw-Hill Dictionary of Engineering (2003; 2008), Oxford Dictionary of Science (2010) and Oxford Advanced Learner’s Dictionary (2010; 2013). For more types of noun compounds found in the engineering corpus, refer to the Appendix (divided into 3 categories).

Conclusion

A wide range of noun compounds were found in the engineering corpus to be learned by students. In essence, grammatical rules of general English language cannot be used to infer meanings of noun compounds, as most of the elements in compound nouns do not retain their literal meaning and require extended meanings (Rahimi & Momeni, 2012). Many non-technical engineering words in compound form evolve into more technical words with extended meanings. These extended meanings were more specific to the given engineering field and more technical in hierarchy. However, there was no definite pattern of hierarchy, as the

collocation patterns indicated that the nouns could combine and form new noun compounds which are semi-technical or technical in nature.

Funding

This work was supported by the **BOLD 2025 Grant Funding Universiti Tenaga Nasional (Project Code – 10436494/B/2019050)**

Acknowledgement

The researchers are sincerely thankful to the support and fund provided by iRMC, Universiti Tenaga Nasional to carry out this study under the BOLD 2025 Grant Funding Universiti Tenaga Nasional *(Ref: "Developing Energy Literacy through Language in the 21st Century Learning Project Code – 10436494/B/2019050)*

REFERENCES

Alif Fairus Nor Mohamad & Ng, Y. J. (2013). Corpus-based studies on nursing textbooks. *Advances in Language and Literary Studies, 4*(2), 21-28. doi:10.7575/aiac.alls.v.4n.2p.21

Arnold, J. (1999). *Affect in Language Learning*. Cambridge, UK: Cambridge University Press.

Bruce, I. (2011). *Theory and concepts of English for Academic Purposes*. London, UK: Palgrave Macmillan.

Cowan, J R. (1974). Lexical & syntactic research for the design of EFL reading materials. *TESOL Quarterly, 8*(4), 389-399.

Coxhead, A. & Hirsh, D. (2007). A pilot science word list for EAP. *Revue Française de Linguistique Appliqueé 12* (2), 65-78.

Durrant, P. L. (2008). *High-frequency collocations and second language learning*. (Unpublished doctoral dissertation). University of Nottingham, UK. Retrieved May 6, 2008 from http://etheses.nottingham.ac.uk/622/

Flowerdew, L. (2012). *Corpora and language education*. Hampshire, UK: Palgrave Macmillan.

Gardner, D., & Davies, M. (2007). Pointing Out Frequent Phrasal Verbs: A Corpus-Based Analysis. *TESOL Quarterly, 41*(2), 339-359.

Gilmore, A. & Millar, N. (2018). The language of civil engineering research articles: A corpus-based approach. *English for Specific Purposes 51*, 1–17

Green, C. & Lambert, J. (2018). Advancing disciplinary literacy through English for academic purposes: Discipline-specific wordlists, collocations and word families for eight secondary subjects. *Journal of English for Academic Purposes, 35*, 105-115

Hoey, M. (2005). *Lexical Priming: A new theory of words and language*. London, UK: Routledge.

Hoey, M. (2007). Lexical priming and literary creativity. In M.Hoey et al. (Eds.), *Text, Discourse and Corpora*, (pp.7-29). London, UK: Continuum.

Hunston, S. (2002). *Corpora in applied linguistics*: Cambridge University Press.

Hunston, S., & Francis, G. (2000). *Pattern grammar. A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.

Krashen, S. & Scarcella, R. (1978). On routines and patterns in language acquisition and performance. *Language Learning, 28*, 283 – 300.

Lewis, M. (1993). *The lexical approach, the state of ELT a new way forward*. Boston: Thomson Publishing.

Lewis, M. (1997). *Implementing the lexical approach, putting theory into practice*. Boston: Thomson Publishing.

Lewis, M. (2000). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 126-154).Hove: Language Teaching Publications.

Lewis, M. (2001). *Implementing the lexical approach. Putting theory into practice.* Hove, England: LTP.

Ma, H. & Qian, M. (2020). The creation and evaluation of a grammar pattern list for the most frequent academic verbs. *English for Specific Purposes, 58*, 155 – 169.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied linguistics*, *33*(3), 299-320.

Master, P. (2003). Noun Compounds and Compressed Definitions. *English Teaching Forum 41*(3), 2-9.

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.

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.

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford, UK: Oxford University Press.

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

Mukundan, J. & Menon, S. (2007). Lexical similarities and differences in the mathematics, science and English Language textbooks. *k@ta, 9*(2), 91-111.

Mukundan, J. (2009). *ESL textbook evaluation: A composite framework*. Köln, Germany: Lambert Academic Publishing.

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.

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.

Nation, I.S.P. (2007). The four strands. *Innovation in Language Learning and Teaching 1(*1), 1-12.

Nation, I.S.P., (1990). *Teaching and learning vocabulary*. New York, NY: Newbury House Publishers.

Nattinger, J.C. & DeCarrico, J.S. (1992). *Lexical phrases and language teaching*. Oxford, UK: Oxford University Press.

Ng, Y. J. , Chong, S.T., Loh, K. S. Ahmad Zufrie Abd Rahman, Thiruchelvam, S. (2019). Developing a Vocational Engineering Word List: Bridging the Gaps of Secondary and Tertiary Education. International Journal of Advanced Science and Technology, 28, No. 10, pp. 104-112

Pawley, A. & Syder, F.H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Richards, J.C. & Schmitt, R.W. (Eds.), *Language and communication* (pp.191 – 225). London, UK: Longman

Rahimi, M. & Momeni, G. (2012). The effect of teaching collocations on English language proficiency. *Procedia - Social and Behavioral Sciences, 31*, 37 – 42

Romer, U. (2010). Using general and specialised corpora in English language teaching: Past, Present and Future. In M.C. Campoy, B. Belles-Fortuno & M. L. Gea-Valor (Eds.). *Corpus-Based Approaches to English Language Teaching* (pp. 18-38). London, UK: Continuum.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, UK: Cambridge University Press.

Scott, M. (1996). *WordSmith Tools*. Oxford: Oxford University Press. ISBN 0-19-458984-6.

Scott, M. (2008). *WordSmith Tools version 5*. Liverpool: Lexical Analysis Software.

Shin, D. and Nation, I.S.P. (2008). Beyond single words: the most frequent collocations in spoken English. *ELT Journal 62*, 4: 339-348. doi:10.1093/elt/ccm091.

Sinclair, J.M. (1991). *Corpus, concordance and collocation*. Oxford, UK: Oxford University Press.

Thiruchelvam, S., Ng, Y.J, Chong, S.T., Norhayati Bte Mat Husind, Azrul Ghazalie, Ahmad Kamal (2019). Communicating the Keywords of Sendai Framework: A corpus-based approach. *International Journal of Innovation, Creativity and Change, 8* (7), 27-45

Thiruchelvam, S., Ng, Y.J, Chong, S.T., Azrul Ghazali & Norhayati Bte Mat Husin (2018). The Language of Civil Engineering: Corpus-based Studies on Vocational School Textbooks in Malaysia. *International Journal of Engineering & Technology, 7* (4.35), 844-847

Todd, R.W. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes, 45*, 31-39.

Tognini-Bonelli, E. (2001). Corpus linguistics at work: *Studies in corpus linguistics*, *v. 6*. Amsterdam: John Benjamins

Trimble, L. (1985). *English for Science & technology: A discourse approach*. New York, NY: Cambridge University Press.

Ward, J. (2007). Collocation & technicality in EAP engineering. *English for Specific Purposes, 6*, 18-35.doi:10.1016/j.jeap.2006.10.001

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purpose, 28*, 170-182. doi:10.1016/j.esp.2009.04.001.

Yang, M.N. (2015). A nursing academic word list. *English for Specific Purposes, 37*, 27-38.

Appendix: Noun compounds identified in the Engineering corpus

Category 1: The Forming Words can be at Both Head and Base of the Target Keyword

No.	Keywords	Similar base and head words with different meanings	
		Keyword as Head	Keyword as Base
1	Circuit	Circuit CONTROL	CONTROL Circuit
		Circuit SWITCH	SWITCH Circuit
		Circuit Load	LOAD Circuit
		Circuit Inductor	INDUCTOR Circuit
		Circuit CURRENT	CURRENT Circuit
2	Voltage	Voltage OUTPUT	OUTPUT Voltage
		Voltage SUPPLY	SUPPLY Voltage
		Voltage TERMINAL	TERMINAL Voltage
3	Concrete	Concrete REINFORCEMENT	REINFORCEMENT Concrete
		Concrete MASS	MASS Concrete
4	Valve	Valve POSITION	POSITION Valve
5	Tank	Tank SERVICE	SERVICE Tank
		Tank WATER	WATER Tank
6	Transformer	Transformer POWER	POWER Transformer
7	Electron	Electron ENERGY	ENERGY Electron
		Electron VALENCE	VALENCE Electron
8	Emitter	Emitter BASE	BASE Emitter
9	Casting	Casting MOULD	MOULD Casting
10	Tensile	Tensile STEEL	STEEL Tensile
11	Meter	Meter SQUARE	SQUARE Meter
12	Flux	Flux LEAKAGE	LEAKAGE Flux
13	Armature	Armature ROTATING	ROTATING Armature
14	Relay	Relay CONTROL	CONTROL Relay
15	Carrier	Carrier FREQUENCY	FREQUENCY Carrier
16	Binary	Binary BIT	BIT Binary

Category 2: The Forming Compound Word Takes the Head Position

Keywords	Keyword as Base	
Circuit	LOGIC	Circuit
	INTEGRATED	Circuit
	SUPPLY	Circuit
	AMPLIFIER	Circuit
	RESISTOR	Circuit
	SERIES	Circuit
	HYDRAULIC	Circuit

	PARALLEL	Circuit
	SCHEMATIC	Circuit
	SHORT	Circuit
	MAIN	Circuit
	SWITCHING	Circuit
	PNEUMATIC	Circuit
	ELECTRONIC	Circuit
	BASIC	Circuit
	BIASING	Circuit
	PRINTED	Circuit
	CURRENT	Circuit
	STARTING	Circuit
	CAPACITOR	Circuit
	DIGITAL	Circuit
	HYBRID	Circuit
	ELECTRIC	Circuit
	RECTIFIER	Circuit
	HOLDING	Circuit
	WIRING	Circuit
	MOTOR	Circuit
	EQUIVALENT	Circuit
	INTERNAL	Circuit
	OPEN	Circuit
	OUTPUT	Circuit
	PIPING	Circuit
	DIVIDER	Circuit
	FILTER	Circuit
	INPUT	Circuit
	LOAD	Circuit

Category 3: The Forming Compound Word Takes the Base Position

Keywords	Keyword as Base	
Circuit	Circuit	Breaker
	Circuit	Diagram
	Circuit	Operation
	Circuit	Board
	Circuit	Table
	Circuit	Technology