Lexical bundles in Thai Medical Research Articles

Panupan Panthong [1]
Kornwipa Poonpon[2]
Khon Kaen University, Khon Kaen, Thailand
Corresponding author's email: pt.panupan@gmail.com
Received 21 September 2019; revised 10 November 2019; accepted 21 January 2020; online 22 May 2020

Abstract

Lexical bundles, co-occurred words serving as building blocks facilitating communication, have been a focus in studies of academic written English for decades. Nonetheless, few studies highlighted this topic in academic writing for specific purpose .This research study aimed to investigate lexical bundles used in medical research articles published in the highest rank and to analyze their structure and functions .A corpus of 346 medical research articles was collected from the highest ranked medical journal published in Thailand. The Thai Med corpus consists of 801,052 tokens. AntConc software

program, N - Grams function, was used to identify four lexical bundles. An analysis of structures and functions of these lexical bundles was conducted, using a framework adapted from Biber, Conrad and Cortes (2004) and Hyland (2008a). The result revealed 67 salient lexical bundles in the corpus. Structurally, prepositional phases and noun phrases are most frequently used. Functionally, the lexical bundles were most highly used for research-oriented function, followed by text- oriented, while no bundles were used for participant-orientation . This study indicates that lexical bundles produced by L2 writers are greater than those in previous research of L1 while the functions of lexical bundles are not widely used.

Keywords: Lexical bundles, Corpus analysis, Research articles, Thai medical journals

1. Background

The fluency of linguistic production is concerned with lexical bundles (Hyland, 2008a, 2012; Lynch & Maclean, 2000). Linguistically, lexical bundles (Biber, Johansson, Leech, Conrad, & Finegan, 1999), or sometimes called 'clusters' (Handford, 2007, 2010), 'formulaic sequences'

(Wray, 2002; Schmitt, 2004), 'chunks' (De Cock, 2004; O'Keeffe, McCarthy, & Carter, 2007), 'lexicalized sentence stems' (Pawley & Syder, 1983), 'N-grams' (Stubbs, 2005) and 'multi-word sequences' (Biber et al, 2004), are the occurrence words that serve as 'building blocks' to glue the discourse, especially academic discourse (Biber et al., 1999). Lexical bundles occur by chance and come with frequency in the form of syntactic fragments e.g. *I want you to, you know what I, and do you think I* (in conversation); *that's one of the, in case of,* and *the beginning of the* (in classroom teaching); and *as a result of, the nature of the,* and *on the other hand* (in academic pose). A majority of lexical bundles in conversation appear in the casual forms while lexical bundles in academic discourse reveal a great number of phrasal fragments (Biber et al., 2004). Those fragment phrases can be bridged to understand meaning in the discourse and facilitated to shape thoughts to convey the sense of meaning (Biber et al., 1999). Lexical bundles are not necessarily idiomatic expressions, but they are extended collocations which don't provide the sense of meaning. The phrase 'kick the bucket' (Biber et al., 1999, p. 988), for example, means to pass away, and this

phrase doesn't function semantically. As a result, it isn't counted as a lexical bundle.

Beyond the body of knowledge in linguistics, phraseology has shed light on the English language and especially on applied linguistics for almost half a century. Phraseology was also established to focus on co-occurring words rather than single word units since English is regarded as phraseological language and co-occurring words occur frequently in everyday life (Biber et al., 2004). This could be drawn from the achievement of learning language in both first and second languages when the formulaic language plays an important role in spoken and written English produced by both native and non-native speakers (Wray, 1999, 2000; Wray & Perkins, 2000). In the psycholinguistic approach, formulaic language was stored in long term memory (Bollinger, 1976; Wray, 2002; Wray & Perkins, 2000). When it is heard, the discourse organizer, for example, dynamically functions to bridge the communication. Also, lexical bundles can reduce mental processing to perform productive skills like writing and speaking (Lynch & Maclean, 2000). The lexical bundles

should be set as examples for novice writers in order to shift their writing fluency (Hyland, 2008a).

Studies of lexical bundles are grounded and distributed from the boundaries of research from a general register into specific registers: namely, university register, academic register, and ESP register (Biber, 2003, 2006; Biber & Barbieri, 2007; Biber et al., 1999; Biber et al., 2004). However, when Cortes (2004) found that lexical bundles can vary across genre, the mainstream research of lexical bundles has contributed to academic fields including history and biology (Cortes, 2004), computer science (Lin & Kuo, 2014), engineering (Alquraishi, 2014; Chen, 2008), journalism (Dastjerdi & Rafiee, 2011), medicine (Abdollahpour & Gholami, 2018; Mbodj-Diop, 2016; Jalali & Moini, 2014; Jalali, Moine, & Arani, 2015), pharmacy (Grabowski, 2015), psychology (Mutiara, 2018), agriculture (Shi, 2010), EU documents (Jablonkai, 2010), and applied linguistics (Johnston, 2017).

Even though lexical bundles have been well researched in the different disciplines, most of the research involved L1 writers. These studies may not highly grab attention to explore how L2 writers achieve

their tasks. Some previous research makes a significant contribution to the knowledge of lexical bundles and also sets out the characteristic features of L2 bundles. It is noted that lexical bundles between L1 and L2 writers are different structurally and quantitatively. Bychkovska and Lee (2017), Pang (2010), and Wei and Lei (2011) agree that the L2 writers produce lexical bundles in verb phrase fragment forms in argumentative essay and theses. Additionally, the number of lexical bundles occurring in L2 writers are greater than L1 writers as shown in past work (see Bychkovska & Lee, 2017; Cortes, 2006; Pan, Rappen, & Biber, 2016; Pang, 2010; Staple, Egbert, Biber, & McClair, 2013; Wei & Lei, 2011) while others disagree profoundly that L2 lexical bundles are not as extensive and various as L1 bundles (Erman, 2009; Granger, 1998; Howarth, 1998). Few studies lay an emphasis on L2 lexical bundles in particular disciplines although lexical bundles vary from discipline to discipline (Cortes, 2004; Hyland, 2008a, 2008b).

To enhance effective communication within the discourse community, it is widely regarded that the medical profession uses English for their training and

publishing research articles, case reports, and other medical-related language production (Ferguson, 2013; Maher, 1986). Thai doctors, who are expected to be a part of free flow labor in ASEAN communities, expressed the demand on writing skills (Supanatsetakul, 2014; Suwanrot, Sausukpaiboon, Ketdao, Suebsunthorn, Prathombut, Thepwongsa, & Boonjaraspinyo, 2017), Supanatsetakul (2013) and claimed that writing courses for doctors are insufficiently developed. As can be seen, previous research in medical discourse (e.g. Grabowski, 2015; Jalali et al., 2015; Marco, 2000; Mbodj-Diop, 2016) highlighted L1 bundles in medical discourse. So far, there have been few studies of L2 writers in medical disciplines. With the importance of lexical bundles facilitating effective communication and function to glue discourse, especially in academic discourse, research on lexical bundles in medical disciplines is extremely rare. Consequently, to bridge the gap between lexical bundles in second language writing and the language used in medical disciplines, this study aims to explore and analyze both functional and structural patterns of the lexical bundles in Thai medical research

articles. The present study, therefore, examines the following research questions:

1) What are the most frequent lexical bundles found in Thai medical research articles?
2) What are structural patterns of the most frequent lexical bundles found in Thai medical research articles?
3) What are functional patterns of the most frequent lexical bundles found in Thai medical research articles?

2. Method

2.1 Thai Medical corpus

A research article corpus called the Thai Med Corpus was compiled from 364 medical research articles published from 2007 and 2017 in the highest ranked medical research journal in Thailand. The corpus included 801,052 tokens. The selection criteria for the journal was that the medical journal showed the highest index among national journals in the Tier 1 rated by Thai-Journal Citation Index (TCI) and the ASEAN Citation Index

(ACI). Besides, this journal was published bimonthly with a various range of medical knowledge, and it provides benefit for a great number of the medical practitioners in Thailand, including doctors, medical researchers, and people in related public health disciplines.

The compilation of the Thai Med Corpus involved the following steps. First, 364 Thai medical research articles published from 2007 to 2017 (available in the collection time) were downloaded and saved as .pdf files in separate folders; each folder represented one year. After that, all of the .pdf articles were transferred to .txt files. For each text file, research names, institutions, figures, acknowledgement, and citations were excluded.

2.2 Data analysis

There were three stages of data analysis: identifying lexical bundles, classifying their structural patterns, and examining functional patterns. The identification of bundles focused on four-word lexical bundles and their frequency. The selection of four-word bundles, instead of three-, five- or six-word bundles, was related to the following reasons. First, three-word bundles were considered to be

too prevalent in the corpus, and some of them overlapped and existed as a part of four-word bundles (Biber et al, 1999; Cortes, 2004; Csomay, 2013). Five- and six-word bundles were too rare to occur with their length of words (Biber et al., 1999). More importantly, the four-word bundles were of interest in a number of lexical bundle studies both in the same and different disciplines. This can lead us to discuss research findings of the present study with previous work. Apart from given reasons, many researchers (see Beng & Keong, 2014; O'Keeffe et al., 2007), for example, conducted empirical studies on lexical bundles from two to six words, and they found that two and three word bundles are numerous while five and six word bundles are rare or they don't meet the cut-off point criteria.

Apart from the length of lexical bundles, the identification of valid lexical bundles in the Thai Med Corpus came with the cut-off point criteria at 32 times with minimum five range. This cut-off point was normalized because the size of a different corpus size. For the cut-off point criteria, linguists (see Biber, 2006; Biber & Barbieri, 2007; Biber et al., 2004) offer the tentative explanation that frequency of lexical bundles at 40 times can reflect

the validity of lexical bundles to study. Likewise, the range was set to avoid the idiosyncrasy of writer preference in only one text, so this could ensure the lexical bundles were used by several writers in the same register. Therefore, this made the cut off criteria of 40 times per million with minimum five range valid and reliable to explore bundles.

To identify the four-word bundles, the computational software program 'AntConc' (Anthony, 2014) was used. In this the program, the function 'N-Grams' was used and set the cluster size and minimum frequency as well as range of lexical bundles. However, due to the limitation of computational software, lexical bundles falling into two characteristic features were excluded: those with the proper noun and with overlap of five-word bundles. Lexical bundles articulated with a proper noun e.g. *'in the Hong Kong'* (Hyland, 2008a, p. 13) were excluded because these bundles provide the sociological background which doesn't benefit L2 writers (Chen & Baker, 2016). Also, four-word bundles are a part of five-word bundles. For example, 'of this study was' and 'this study was to' are a part of five-word bundles 'of this study was to'. In this

case, these bundles can cause the redundancy of lexical bundle occurrence (Chen & Baker, 2010).

The second stage focused on classifying structural patterns of the lexical bundles. Structural taxonomy (Biber et al., 2004) was used as a theoretical framework. In this framework, three types and 17 subtypes of structural patterns were included (Table 1). The three main types of structural patterns were (1) verb phrase fragments (e.g., *It's going to, is going to be*), (2) dependent clause fragments (e.g., *what I want to, if you look at*) and (3) noun phrase and propositional phrase fragments (e.g., *one of the things, the way in which*). Each type of structural patterns included seven sub-types of verb phrase fragments, five sub-types of dependent clause fragments and five sub-types of noun phrase and prepositional phrase fragments, respectively.

The final stage focused on identifying functional patterns of the lexical bundles to explain how lexical bundles are used to refer to their meanings in discourse (Biber et al., 1999; Biber et al. 2004). Hyland's (2008a) framework on pragmatic functions of lexical bundles was used as it is a "more research-focused genre" (Hyland, 2008a, p.13). Such a framework, developed from many

previous conceptual frameworks (Biber, 2006; Biber et al., 2004; Halliday, 1994), was purposely invented for an academic written genre as opposed to Biber (2006) and Biber et al. (2004)'s framework underlying the broader range of lexical bundles in several registers e.g. everyday conversation, textbook, classroom teaching and so on. Hyland' s pragmatic functions were divided into three main categories: research-oriented, text- oriented, and participant-oriented with 11 sub-categories.

The research- oriented category helps writers to develop their writing on real situations and motions. It also introduced 5 subcategories: (1) location (in time/ place), (2) procedure, (3) qualification, (4) description, and (5) topic. Being analyzable, the statistical group, a subtype of qualification function, was suggested by Cortes (2004) whose taxonomy instituted that bundles in science- related processes are relevant to statistical groups, so this group was used to facilitate the analysis of medical discourse.

- Location (in time/ in place) indicating time and place: *at the beginning of the, at the same time, in the present study*

- Procedure describing the research methodology or objective: *the use of the, the role of the, the purposes of the*
- Qualification reporting the number and falling into two subtypes:
    - General group: *the magnitude of the, a wide range of, one of the most*
    - Statistical group: *there was no significant, did not differ significantly, there were no different*
- Description specifying qualities and properties of materials: *the size of the, the structure of the*
- Topic of the research area: *in the Hong Kong, the currency board system*

The text- oriented function was related to text organization when conveying messages or arguments. Four subcategories were introduced in this type.

- Transitional signal establishing additive or contrastive links between elements: *on the other hand, in addition to the*

- Resultative mark inferential or causative relations between elements: *as a result of, it was found that, these results suggest that*
- Structural signals text- reflexive markers which organized stretches of discourse or direct readers elsewhere in the text: *in the present study, in the next section, as shown in the fig.*
- Framing signals situated arguments by specifying limiting conditions: *in the case of, with respect to the, on the basis of, in the presence of, with the expectation of*

The participant- oriented function relied on the writer or reader in the text, and there were two subcategories.

- Stance features conveyed the writers' attitudes and evaluations: *are likely to be, may be due to, it is possible that*
- Engagement features addressed readers: *it should be noted that, as can be seen*

3. Findings and Discussion

3.1 Frequency of lexical bundles in Thai Medical research articles

The investigation of four-word lexical bundles shows 67 types of four-word lexical bundles, occurring at least 32 times and in a minimum of 5 ranges (texts), in Thai medical research articles. According to Table 2, the two most frequent lexical bundles are those with prepositional phrases (i.e., 164 and 161 times with 104 and 102 ranges, respectively). The subsequent bundles (among the top ten) are *was approved by the, is one of the, study was approved by, in the control group, in this study was, in the present study, shown in table the,* and *there was no significant*. The least frequently occurring lexical bundles, based on the criteria of at least 32 occurrences with 13 and 31 ranges, are *the quality of life* and *was used to determine.* This range can represent the texts from the several writers producing lexical bundles in the corpus; particularly, it can avoid idiosyncrasies among writers (Biber et al., 1999). The findings of frequency of lexical bundles in Thai medical research articles correspond with previous research. As can be seen, ThaiMed corpus provide 67

bundles while Mdodj-Diop (2016) found 48 bundle types at 40 times of the cut off point. This supports that previous research of lexical bundles produced by L2 writers are greater than those L1 writers (see Bychkovska & Lee, 2017; Cortes, 2004; Pan et al., 2016; Pang, 2010; Staple et al., 2013; Wei & Lei, 2011). Some scholars (e.g. Huang, 2015) claim that the greater occurrence of lexical bundles doesn't necessarily reflect the advantages of higher proficiency, yet Ellis (1996) stated that the fewer bundle types are produced by high proficient writers.

3.2 Structural patterns of lexical bundles

For the structural patterns of the lexical bundles used in the Thai Med Corpus, the structural taxonomy by Biber et al. (2004) was adopted to classify the bundles in the types and sub-types. As can be seen in Table 3, the results reveal that the largest proportion of structural patterns are noun phrases and prepositional phrase fragments (56.58%), followed by verb phrase fragments (30.94%) and dependent clause fragments (3.94%), respectively

In noun phrase and prepositional phrase fragments, out of 43 bundles the most widely found categories are prepositional phrase expressions **(**23 bundles**)** and noun

phrases with *-of* phrase fragments (15 bundles), respectively. In academic discourse, it is quite common that the noun phrase and prepositional phrase fragments are the largest proportion above other types, and L1 writers produced noun phrases in the largest proportion in the previous research (e.g. Biber et al., 2004; Hyland, 2008a, 2008b; Jalali et al., 2015; Johnston, 2017; Kwary, Ratri, & Artha, 2017; Mbodj-Diop, 2016). Nonetheless, lexical bundles in L2 writing were primarily dependent on prepositional phrases. These phrases were expressed in the highest number; this result is similar to those found in a number of studies (e.g. Chen & Baker, 2010; Wei & Lei, 2011). Therefore, this present study shares the mainstream pattern of lexical bundles in an academic discourse.

When looking at the verb-phrase fragments, the most frequently used sub-type of verb phrase fragments is the verb phrase (passive). Owing to the relation to academic disciplines, it is quite common to see verb phrases in passive forms like the previous studies (see Biber et al., 2004; Cortes, 2004; Hyland 2008a, 2008b; Kwary et al., 2017; Mbodj-Diop, 2016).

The dependent clause fragments are rare bundles in this corpus. This corresponds with the literature in that that this sub-type occurs more frequently in the spoken register, including conversation and classroom teaching, than in academic genres (Biber et al., 2004).

Apart from the proportions discussed above, it was found that in the sentence below (V67_N3), lexical bundles including *was approved by, and by the Ethics Committee* connect verb-phrase fragments and noun phrases to build a structural pattern in the academic discourse (Biber et al., 1999).

Example

- *The study protocol was approved by the Ethics Committee of the Faculty of Medicine….* (V67_N3)

3.3 Functional patterns of lexical bundles

Apart from the structural patterns, the function patterns of lexical bundles were analyzed using Hyland's framework. Since lexical bundles can be used for more than one function (Biber et al., 2004), the number of the proportions can show more numbers than 76 bundle types. Therefore, as shown in the Table 4, the result revealed 71 functional types of 68 lexical bundles found in the Thai

Med Corpus. Among this number, about 80.28% are research- oriented patterns, followed by text- oriented patterns accounting for 19.72%. Participant- oriented patterns are not found in the corpus.

According to the results, the occurrences of the functional patterns in the Thai Med Corpus are similar to those found in science- related corpora (Hyland, 2008a) and the medical research article corpus by L1 writers (Mdodj-Diop, 2016) in that the most frequently used functional patterns are research- oriented, text- oriented, and participant- oriented patterns, respectively. However, the Thai Med Corpus show no bundles in participant-oriented function.

In the following section, a closer look of lexical bundles in medical discourse of Thai medical researchers is shown in each type and each-subtype.

3.3.1 Research-oriented function

In this study, the largest proportion of lexical bundle occurrences relies on the research- oriented function. Lexical bundles take into account five sub- categories including location, procedure, qualification, description and topic. The results show that the lexical bundles found

in the corpus reflect all sub-categories under the research-oriented function. Looking closely, it is found that in the research- oriented function, prepositional phrases and noun phrases are used to specify the location (in time/ in place); this correlates with previous research (Hyland, 2008a; Mbodj-Diop, 2016). Examples include *the Faculty of Medicine, in the control group*, *length of hospital stay*, as shown below.

Examples

(1) The significant factors associated with post-operative *length of hospital stay* were type of surgery and CPB time, while mood state was not found to be significantly related to LOS (V63_N2)

(2) Completeness was confirmed anatomically *at the time of* surgery by inspecting the specimen and the resected thymic bed. (V69_N2)

(3) The subjects in the control group received routine nursing care, whereas those *in the experimental group* received routine nursing care plus guided imagery therapy. (V61_N2)

(4) They were obtained from the EDX laboratory of *the Faculty of Medicine* at Naresuan University. (V67_N3)

(5) The study samples were 118 cases performed on transbronchial biopsy and pleural biopsy specimens *at the Department of* Pathology, Siriraj Hospital, Mahidol University … (V63_N3)

For the procedure function, the lexical bundles used by L2 writers share similarities with those by native writers explored in a previous study. The bundles *at the end of the*, and *purpose of the study*, for example, are used to describe the procedure of research methodology (see examples below). Most of the bundles expressing the process are in forms of passive verb phrases as well as prepositional phrases (Hyland, 2008a; Mbodj Diop, 2016).

Examples

(6) The research protocol *was approved by the* local ethics committee for clinical research and all procedures involving each participant were conducted according to institutional guidelines. (V61_N6)

(7) *The objective of this* study was to compare the feasibility and oncologic outcomes between laparoscopic and open surgery after placement of SEMS for acute malignant colorectal obstruction. (V69_N2)

(8) Fisher's exact test *was used to compare* outcomes between levels of each diagnostic category. (V61_N3)

(9) The total of 140 cases were *included in this study* with the average age of … (V63 N4)

(10) *This study aimed to* establish normal PNIF ranges for an Asian population accounting for sex, age, weight, and height. (V67_6)

The qualification function can be further classified into general and statistical groups as mentioned in the data analysis section (Cortes, 2004). Because bundles among science-related procedures take into account a statistical approach (Cortes, 2004), the statistical groups are drawn to show how they are used in medical discourse. In the general group, the lexical bundles include, for example, *most of the patients, the amount of blood, patients in the study* (see examples 11- 13) while the statistical group mainly deals with *there was no significant,* and *were not significantly different* (see examples 14 and 15*)*. To give a clearer picture of bundles addressing qualifications in the discourse, the examples are shown as the examples below.

Examples

(11) *Most of the patients* with lymph node involvement had multiple groups (96.2%) and the most common location was mesenteric region. (V66_N3)

(12) The research indicated that *the amount of blood* loss post-operation among patients who did not experience hypovolemic shock, averaged at 1,018.86 ml. (V63_N3)

(13) All hemodialysis *patients in this study* used EPO for improving the Hb level and quality of life.

(14) In comparison of anesthesia duration and surgical satisfaction, *there was no significant* difference between both groups. (V67_6)

(15) The QOL scores in kidney disease components of the different Hb levels *were not significantly different* except in the effect … (V63_N1)

For the description function, four-word bundles are rarely found in this subtype. Those found in the corpus are *had a history of*, *is the most common,* and *was the most common*. As seen, two of them are only different in terms of tense use.

Examples

(16) Acute appendicitis *is the most common* pediatric surgical condition. The incidence … (V61_N43)

(17) In the neoplastic category, misinter-pretation of tumor type *was the most common* problem and gliomas had the highest number … (V66_N45)

(18) The benefit of early enucleation of injured eye in this study was also unable to evaluate since we had no data of patients who *had a history of* ocular trauma, but had no SO to compare. (V69_N2)

However, there were no bundles occurring in the topic function in the Thai Med corpus. According to Hyland's (2008a) function, the topic function, for example, include lexical bundles such as '*in the Hong Kong'*, so based on the exclusion criteria, such bundles were excluded. Also, it is possible that the topic function is not used as extensively as L1 writers do in the previous research since L2 writers' trend to produce fewer various bundles.

3.3.2 Text-oriented function

The lexical bundles occurring in the Thai Med Corpus reflect the text-oriented function in terms of research signal, structural signal, and transitional signal. Only one bundle type was found in transitional signals (see example 19)

while there are no bundle types in framing structures. Additionally, the lexical bundles like *was found to be, we found that the, it was found that* the are used to inform the readers about causative relations (see examples 20-23). For the structuring signal shown in example 24-25, these bundles, for instance, include s*hown in the table, have been shown in*, and *as shown in fig*. As discussed above, lexical bundles produced by L2 writers are not extensive and various (Ädel & Erman, 2012; Chen & Baker, 2010; Erman, 2009; Granger, 1998; Howarth, 1998). Compared to L1 writers of medical research articles in Mbodj-Diop's (2016) work, it can be seen that the Thai Med corpus shows a greater number of occurrence bundles but less various function. Consequently, this supported the idea that lexical bundles produced by L2 show fewer varieties (Howarth, 1998).

Examples

(19) *Poor cardiovascular fitness has been associated with a markedly increased risk of premature death from all causes and specifically from cardiovascular diseases. On the other hand, high cardiovascular fitness can have a protective effect.* (V66_N6)

(20) In this study, hyperuricemia *was found to be* associated with high BMI, CRP and low HDL in the patient group.(V67_N2)

(21) … 87.5 % of subjects in the experimental group were satisfied with *the results of the* specified exercise and 92.2% ... (V68_N6)

(22) From our study, *we found that the* wound infection after operation for non-complicated appendicitis … (V61_N4)

(23) Besides, 130 inpatients *were found to be* same day cases, with the same admission and discharge date. (V67_N5)

(24) … to those obtained from the reference *product as shown in Table 2*. The geometric mean … (V67_N3)

(25) Association between CT characteristics and the nature of adrenal masses *have been shown* in Table 1. (V65_N2)

3.3.3 Participant-oriented function

Nevertheless, lexical bundles articulated with participant-oriented function do not belong in the framing signal. As seen in science-related disciplines, the lexical bundles of this type in aforementioned research (see Hyland, 2008a; Jalali et al., 2015; Mbodj-Diop, 2016) showed the smallest proportion in the framework. To provide plausible explanations of why lexical bundles functioning as participant-oriented function are rare, this

result is discussed. First, this result is in agreement with some previous studies that lexical L2 writers produce lexical bundles less widely and frequently than L1 (Ädel & Erman, 2012; Chen & Baker, 2010; Erman, 2009; Granger, 1998; Howarth, 1998), so it should be taken into the account that lexical bundles in this type may not be seen in this frequency. Apart from that, among science-related disciplines is medical writing, and based on the conceptual framework, this function is composed of lexical bundles, namely '*are likely to be*', '*may be due to*', '*as can be seen*' and so on. Due to these bundles, Durrant (2017) laid emphasis on disciplinary variation of lexical bundles of university students, postulating that science related writing is inclined to convey facts straightforwardly rather than produce the description in soft science writing (social science) (North, 2005). For the stance bundles, Biber et al. (2004) and Biber and Barbieri (2007) also acknowledge that these bundles occur prevalently in spoken registers.

4. Conclusion

This study aims to explore the structural and functional patterns of lexical bundles among L2 medical writers. Consequently, 76 were bundles were found, and some bundles were related to the sociocultural background of the writers. L2 medical writers use specific bundles to communicate among L2's discourse community as also shown in previous research (Hyland, 2008a). Structurally, lexical bundles in L2 medical writers show repositional phrases in the highest proportion. Another point is that the bundles by L2 medical writers show fewer varieties of functions. There is evidence suggesting the L2 writers use bundles with less frequency and variety than native writers do (Ädel & Erman, 2012; Pérez-Llantada, 2014). It can be seen that lexical bundles produced by Thai writers are overused and less various of writing style. Pedagogically, lexical bundles are useful for L2 writers (Cortes, 2004; Hyland, 2008a, 2008b), and more importantly lexical bundles were claimed to be a part of syllabi of language courses (Elis, Simson-Vlach, & Maynard, 2008; Neely & Cortes, 2009; Römer, 2011). Therefore, it is necessary to encourage other kinds of

functional types with a variety of writing styles in the classroom.

Although the present study sheds light on lexical bundles in academic writing in a specific discipline, all data sets are retrieved from the same journal. Further studies can address other important aspects. As lexical bundles are suggested for effective communication (Cortes, 2013), relationships between lexical bundles and rhetorical moves in specific genres, especially medical discourse, would be worth exploring. Further studies should explore how lexical bundles are transferred in medical discourse as there is a direct transfer from the source language to target language (Cortes, 2006). This can contribute to academic writing in medical discourse.

References

Abdollahpour, Z., & Gholami, J. (2018). Building blocks of medical abstracts: frequency, functions and structures of lexical bundles. *Asian ESP Journal*, *14*(1), 82-110.

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, *31*(2), 81-92.

Alquraishi, M. A. (2014). *Lexical bundles in an advanced INTOCSU writing class and engineering texts: A functional analysis* (Unpublished doctoral thesis). Colorado State University, Fort Collins, CO.

Anthony, L. (2014). AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/

Beng, C. O. S., & Keong, Y. C. (2014). A corpus study of structural types of lexical bundles in MUET reading texts. *3L: The Southeast Asian Journal of English Language Studies*, *20*(2), 127-140.

Biber, D. (2003). Variation among university spoken and written registers: A new multi-dimensional analysis. *Language and Computers*, *46*, 47-70.

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, *5*(2), 97-116.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, *26*(3), 263-286.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371-405.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman grammar of spoken and written English. New York, NY: Longman.

Bollinger, D. (1976). Meaning and memory. *Forum Linguisticum*, *11*, 1–14

Bychkovska, T., & Lee, J. J. (2017). At the same time : Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes*, *30*, 38-52.

Chen, L. (2008). *An investigation of lexical bundles in electrical engineering introductory textbooks and ESP textbooks* (Unpublished *master's* thesis). Carleton University, Ottawa, Ontario, Canada.

Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, *14*(2), 30-49.

Chen, Y. H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, *37*(6), 849-880.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, *23*(4), 397-423.

Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, *17*(4), 391-406.

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, *12*(1), 33-43.

Csomay, E. (2013). Lexical bundles in discourse structure: A corpus-based study of classroom discourse. *Applied Linguistics*, *34*(3), 369-388.

Dastjerdi, H. V., & Rafiee, M. (2011). Corpus study of lexical bundles: journalistic discourse in focus. *Asian ESP Journal*, *7*(3), 59-78.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literature*, *2*, 225-246.

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, *38*(2), 165-193.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in second language acquisition*, *18*(1), 91-126.

Ellis, N. C., Simpson-Vlach, R. I. T. A., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*(3), 375-396.

Erman, B. (2009). Formulaic language from a learner perspective: what the learner needs to know. In B. Corrigan, H. Quali, E. Moravcsik, & K. Wheatley (Eds.), *Formulaic language* (pp. 27-50).

Amsterdam, The Natherlands: John Benjamins. http://dx.doi.org/10.1075/tsl.83.05erm

Ferguson, G. (2013). 13 English for medical purposes. In B. Paltridge & S. Starfiled (Eds.), *The handbook of English for specific purposes* (pp. 243-262). New York, NY: John Wiley and Sons.

Grabowski, Ł. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, *38*, 23-33.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 145–160). Oxford, England: Clarendon Press.

Halliday, M**. (**1994**)** *An introduction to functional grammar***.** London, England**:** Edward Arnold**.**

Handford, M. (2010). *The language of business meetings*. Cambridge, England: Cambridge University Press.

Handford, M. J. (2007). *The genre of the business meeting: A corpus-based study* (Unpublished doctoral thesis). University of Nottingham, England.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24-44. http://dx.doi.org/10.1093/applin/19.1.24

Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, *53*, 13-23.

Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4-21.

Hyland, K. (2008b). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, *18*(1), 41-62.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32*, 150-169.

Jablonkai, R. (2010). English in the context of European integration: A corpus-driven analysis of lexical

bundles in English EU documents. *English for Specific Purposes*, *29*(4), 253-267.

Jalali, Z. S., & Moini, M. R. (2014). Structure of lexical bundles in introduction section of medical research articles. *Procedia-Social and Behavioral Sciences*, *98*, 719-726.

Jalali, Z. S., Moini, M. R., & Arani, M. A. (2015). Structural and functional analysis of lexical bundles in medical research articles: a corpus-based study. *International Journal of Information Science and Management*, *13*(1), 51-69.

Johnston, K. M. (2017). *Lexical bundles in applied linguistics and literature writing: A comparison of intermediate English learners and professionals* (Unpublished master's thesis). Portland State University, Portland, OR.

Kwary, D. A., Ratri, D., & Artha, A. F. (2017). Lexical bundles in journal articles across academic disciplines. *Indonesian Journal of Applied Linguistics*, *7*(1), 131-140.

Lin, M. H., & Kuo, C. H. (2014). The rhetorical functions of lexical bundles in computer science research article introductions. *Asian ESP Journal*, *1*(2), 244-275.

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, *4*(3), 221-250.

Maher, J. (1986). English for medical purposes. *Language Teaching*, *19*(2), 112-145.

Marco, M. J. L.(2000). Collocational frameworks in medical research papers: A genre-based study. *English for Specific Purposes*, *19*(1), 63-86.

Mbodj-Diop, N. B. (2016). *Lexical bundles in medical research articles: Structures and functions* (Unpublished master thesis). Michigan State University, East Lansing, MI.

Mutiara, R. (2018). Lexical bundles and keywords in psychology research articles, *Asian EFL Journal*. *20* (7), 135-142.

Neely, E., & Cortes, V. (2009). A little bit about: Analyzing and teaching lexical bundles in academic lectures. *Language Value*, *1*(1), 17-38.

North, S. (2005). Different values, different skills? A comparison of essay writing by students from arts and science backgrounds. *Studies in higher education*, *30*(5), 517-533.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge, England: Cambridge University Press.

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in telecommunications research journals. *Journal of English for Academic Purposes*, *21*, 60-71.

Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *Asian EFL Journal*, *47*(1), 10-11.

Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W.

Schmidt (Eds.), *Language and communication* (pp. 191-226). Harlow, England: Longman

Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes*, *14*, 84-94.

Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, *31*, 205-225.

Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing, and use*. Amsterdam, The Netherlands: John Benjamins Publishing.

Shi, H. (2010). *Lexical bundles and moves in agricultural science research articles* (Master's thesis from an institutional database). Retrieved from http://sutir.sut.ac.th:8080/sutir/

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, *12*(3), 214-225.

Stubbs, M. (2005). Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature*, *14*(1), 5-24.

Supanatsetakul, N. (2013). An overview of English language instruction in undergraduate medical program and ways to improve medical student's foreign language skills. *Thammasat Medical Journal*, *13*(4), 561-569.

Supanatsetakul, N. (2014). Perceived readiness and needs of medical students to improve English language skills in medical practice under ASEAN Economic Community (AEC) liberalization in the year 2015. *Thammasat Medical Journal*, *14*(3), 378-386.

Suwanrot, K., Sausukpaiboon, K., Ketdao, N., Suebsunthorn, K., Prathombut, A., Thepwongsa, I., & Boonjaraspinyo, S. (2017). Learning needs assessment for English language of medical students and residents in Srinagarind hospital, faculty of medicine, Khon Kaen university. *Srinagarind Medical Journal*, *32*(5), 454-460.

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC journal*, *42*(2), 155-166.

Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching*, *32*(4), 213-231.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, *21*(4), 463-489.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, *20*(1), 1-28

Table 1

Structural patterns of lexical bundles (Biber et al., 2004)

Type of structural patterns	Sub-types	Examples
1. Verb phrase fragments	1. 1st / 2nd person pronoun +verb phrase fragment	*You don't have to* *I'm not going to*
	2. 3rd person pronoun +verb phrase fragment	*It's going to* *That's one of the*
	3. Discourse maker + pronoun +verb phrase fragment	*I mean you know* *You know it was*
	4. Verb phrase (non passive verb)	*Is going to be* *Is one of the*
	5. Verb phrase (passive verb)	*Is based on* *Can be used to*
	6. Yes-no question fragments	*Are you going to* *Do you want to*
	7. *Wh*- question fragments	*What do you think* *How many of you*
2. Dependent clause fragments	1. 1st/ 2nd person pronoun dependent clause fragments	*I want you to* *You might want to*

Type of structural patterns	Sub-types	Examples
	2. *Wh*-clause fragments	*What I want to* *What's going to happen*
	3. *If*- clause fragment	*If we look at* *If you have a*
	4. *To*-clause fragment	*To be able to* *Want to do this*
	5. *That*-clause fragment	*That this is a* *That I want to*
3. Noun phrase and propositional phrase fragments	1. Noun phrase with *of*-phrase fragment	*One of the things* *The end of the*
	2. Noun phrase with other post-modifier fragment	*The way in which* *Those of you who*
	3. Other noun phrase expression	*A little bit more* *Or something like that*
	4. Prepositional phrase expression	*At the end of the* *Of the thing that*
	5. Comparative expression	*As well as the* *As far as the*

Table 2

Lexical bundles in Thai Med Corpus

No	Lexical bundles	Freq	Range	No	Lexical Bundles	Freq	Range
1	of this study was	139	111	35	the results of the	43	31
2	as shown in table	126	60	36	used in this study	43	33
3	was approved by the	114	112	37	objective of this study	41	39
4	is one of the	106	77	38	the end of the	41	26
5	study was approved by	97	97	39	was no significant difference	41	32
6	in the control group	90	25	40	was the most common	40	29
7	in this study was	90	64	41	were not significantly different	40	27
8	in the present study	75	46	42	patients in this study	39	32
9	shown in table the	75	57	43	the faculty of medicine	39	34
10	there was no significant	73	49	44	the objective of this	39	38
11	between the two groups	71	34	45	there were no significant	39	26
12	in this study we	71	54	46	in our study the	38	33
13	amount of blood loss	64	7	47	most of the patients	38	23

No	Lexical bundles	Freq	Range	No	Lexical Bundles	Freq	Range
14	at the time of	61	25	48	of the patients had	38	31
15	test was used to	59	50	49	of this study were	38	35
16	the mean age of	58	43	50	as shown in fig	37	27
17	was found to be	56	37	51	approved by the ethics	36	36
18	purpose of this study	55	50	52	of the patients in	36	25
19	have been shown in	54	38	53	the present study was	36	27
20	in this study were	53	42	54	and in the control	35	10
21	length of hospital stay	53	14	55	the sample size was	35	32
22	are shown in table	51	34	56	this study aimed to	35	30
23	in the experimental group	51	7	57	by the ethics committee	34	33
24	mean and standard deviation	51	43	58	the majority of the	34	22
25	the purpose of this	49	46	59	in the study the	33	29
26	at the end of	48	32	60	included in this study	33	29
27	we found that the	48	34	61	was used to compare	33	28
28	were excluded from the	47	43	62	were found to be	33	31
29	excluded from the study	44	40	63	at the department of	32	21

No	Lexical bundles	Freq	Range	No	Lexical Bundles	Freq	Range
30	is the most common	44	33	64	had a history of	32	22
31	the study was approved	44	44	65	in the study group	32	10
32	a total of patients	43	33	66	the quality of life	32	13
33	it was found that	43	30	67	was used to determine	32	31
34	on the other hand	43	37				

Table 3

Structural patterns of lexical bundles in Thai Med corpus

Types of Structural Patterns	Number of Bundles	Percentage
Noun phrase and prepositional phrase fragment	36	52.94
Verb phrase fragment	30	44.11
Dependent clause fragment	2	2.94
Total	68	100

Table 4

Functional patterns of lexical bundles in Thai Med corpus

Functional patterns	Number of bundles	Percentage
Research-oriented	57	80.28
Location	24	33.80
Procedure	17	23.94
Qualification	13	18.30
Description	3	4.23
Topic	0	0
Text-oriented	14	19.72
transitional signal	1	1.41
resultative signal	6	8.45
structural signal	7	9.86
framing signal	0	0
Participant-oriented	0	0
stance bundles	0	0
engagement features	0	0
Total	71	100