A Multidisciplinary Corpus-based Comparative Analysis: Lexical Bundles in Language Teaching, Health Sciences, and Business Management Research Articles

Woravit Kitjaroenpaiboon
Faculty of Humanities and Social Sciences, Suan Dusit University

Sresuda Wongwiseskul
Faculty of Nursing, Suan Dusit University

Thanakorn Puksa
School of Management Sciences, Suan Dusit University

Benjamas Khamsakul
Thai Language Curriculum,
Faculty of Humanities and Social Sciences, Suan Dusit University

Received : Aug 18, 2021
Revised : Oct 15, 2021
Accepted : Nov 17, 2021

Abstract

This research analyzes and compares lexical bundles (three- to five-word units) in language teaching, health sciences, and business management research articles, including their communicative functions. The corpus comprises 90 internationally published research articles from the world's top five journals in the three disciplines. Cargil and O'Connor's (2009) Introduction Method Results and Discussion Structure was applied as a framework for compartmentalizing the sections of the research articles and ANTCONC was applied as a key concordancer. In this study, lexical bundles must be found at least five times in every 100,000 words and in five different texts. Conceptual frameworks of corpus scholars were utilized to analyze communicative functions of the lexical bundles. The findings reveal 182 lexical bundles in the language teaching research articles. One-way repeated measure ANOVA details there are no statistically significant differences between the use of them in all four sections (p = .150). One-hundred and eight lexical bundles are found in the health sciences research articles with some statistically significant differences between the use of them in some sections (p = .016). One-hundred and seventy-eight lexical bundles are found in the business management research articles with some statistically significant differences between the use of them in some sections (p < .001). In total, there are 371 lexical bundles in all 90 research articles. There are no statistically significant differences between the use of them in the research articles from the three disciplines (p = .687). Functional analysis reveals the lexical bundles provide 19 different communicative functions.

Keywords: Lexical Bundles, Research Articles, Language Teaching, Health Sciences, Business Management

Introduction

Lexical bundles are sequences of two or more words used with high frequency in discourses. They can represent genre and section of a research article (RA) (Huimin, 2010). Swales (1992) states that lexical bundles in an academic genre differ from ones in general language. For instance, the lexical bundles *in order to avoid* are frequently found in academic language more than in general language. Scholars proposed that only reading academic RAs could not help academics automatically improve their academic writing proficiency as the literary language in RAs is specific and not similar to general English. It has specific stylistic patterns and meanings in each discipline (Kitjaroenpaiboon, 2016). For example, the lexical bundle *cohesion of the*, in physics, *cohesion* means the intermolecular whilst in linguistics, it means how parts of a text are connected together. To write an RA, academics need guidance and to frequently practice writing (Kitjaroenpaiboon & Getkham, 2016a). Our point of view aligns with scholars stating that most academics do not know how to use academic lexical bundles for writing an RA (Cortes, 2013).

Collaborations between academics from different disciplines are necessary for scientific research (Simone et al., 2018). To solve challenges, academics, researchers, and students from different disciplines should work together (Morrison, 2014) and thus should understand how to use lexical bundles within the discipline and in others as well (McLaughlin & Parkinson, 2018). Most academics, researchers, and students take it for granted that, despite utilization in different disciplines, the same English language is similarly used (Joseph et al., 2010). Scholars pointed out that the nature of each discipline results in its specific linguistic characteristics. Natural science scientific paper writers tend to imply that their studies and findings are important whilst the social sciences writers tend to provide persuasive evidences that a need exists for the studies (Boutelier et al., 2011). This aligns with Conrad (1996) who proposed that different disciplines apply discipline-specific languages. By definition, when writing scientific papers, academics should apply a specific language to a specific discipline (Berkenkotter et al., 1991).

Upon reviewing studies focusing on lexical bundles in scientific papers (e.g., Damshevska, 2019; Hyland, 2008; Panthong & Poonpon, 2020; Wongwiwat, 2016), we found that a lexical bundle provides different communicative functions depending on the context it occurs in. For example, when the bundle *found that* occurs in the introduction section of RAs, it functions as referring to other studies (e.g., *numerous scholars found that*). However, when the bundle occurs in the result section, it functions to report findings (e.g., *we found that*).

Having RAs published is important for academics since it signifies academic success (Poggensee, 2016). English RAs help boost the world's academic advancement (Kanoksilpatham, 2005). If researchers or academics want their RAs to be accessible to others, theirs must be written in English (Genc & Bada, 2010). Nevertheless, not all academics can succeed in having their RA published since English is neither their first nor their second language (Kitjaroenpaiboon, 2016). Their English research writing proficiency is somewhat limited (Fadda, 2012). They do not know what lexical bundles are to be used for writing their RAs (Cortes, 2013). With this problem in mind, this paper thus explores and compares lexical bundles and their communicative functions, in each section of internationally published language teaching, health sciences, and business management RAs and between the three disciplines to determine whether the use of lexical bundle in different sections and different disciplines are similar or different. Language teaching, health sciences, and business management have been gaining popularity within the academic field as seen from an increase

of RAs published in the disciplines (Lindstromberg & Eyckmans, 2017). Therefore, RAs from the three disciplines were investigated to analyze and compare which lexical bundles are frequently used.

In this regard, lexical bundle knowledge can help academics comprehend their discipline-matters more accurately and will contribute to their success in professional communication skills (Chirobocea-Tudor, 2018; Cortes, 2004). Providing RA lexical bundle guidance in the three disciplines and between the three disciplines is an underlying reason the researchers conducted this study.

The key objectives of this research are:

1) to study lexical bundles in each section of internationally published language teaching, health sciences, and business management research articles

2) to compare whether the extracted lexical bundles are differently used in each section of each discipline

3) to compare whether the extracted lexical bundles are differently used between the three disciplines

4) to investigate the extracted lexical bundles' communicative functions.

Methodology

This study focuses on analyzing a corpus of internationally published RAs. Thirty internationally published RAs, from each discipline, were collected from the top five highest impact factors international journals.

The corpus of this study consisted of three sub-corpora. With regards to Scimago Journal Rank 2019 (Scimago, 2019), the top five highest impact factors international journals in language teaching are *Journal of Second Language Writing*, *Language Learning*, *Studies in Second Language Acquisition*, *Reading and Writing*, and *English for Specific Purposes*. The top five highest impact factors international journals in health sciences are *World Psychiatry*, *Diabetes Care*, *Stroke*, *American Journal of Clinical Nutrition*, and *Pediatric Obesity*. The top five highest impact factors international journals in business management are *Journal of Finance*, *Journal of Financial Economics*, *Journal of Accounting and Economics*, *Strategic Management Journal*, and *Journal of Accounting Research*. Complying the methodologies of numerous corpus linguists (e.g., Baoya, 2015, Getkham, 2010; Kanoksilpatham, 2005), we randomly selected six RAs from these journals published between 2016 and 2019 to help increase the generalizability of the results.

Subsequently, the four researchers and three assistants analyzed and compartmentalized four key sections (i.e., introduction, methodology, results, and discussion) in the 90 selected RAs by applying Cargil and O'Connor's (2009) Introduction, Methodology, Result, and Discussion (IMRD) Structure of RAs as a framework.

'Lexical bundle' refers to the highest frequency word strings with two or more words in corpora (Biber et al., 1999; Hyland, 2008). Nevertheless, a word string, that can be identified as a lexical bundle, is one which occurs five times upwards per 100,000 words, in sub-corpora of a single register, over a range of five different texts (Biber et al., 2004; Nesi & Basturkmen, 2006). However, two-word bundles are too numerous while six or more-word bundles were too rare to occur or they do not meet the cut-off point criteria (Hyland, 2008).

In this study, the three- to five-word bundles were studied as they have been found to possess more content (Nasrabady et al., 2020). The cut-off frequency adopted a moderately high threshold at five times per 100,000 words and the dispersion threshold was set at occurring over five different texts in each corpus to retrieve the highest frequency and generally used lexical bundles in each discipline. To explore the lexical bundles, ANTCONC (Anthony, 2020) was applied as a concordance program to detect and count lexical bundles' frequencies.

To compare similarities and differences of the lexical bundles in each data set, one-way repeated measure ANOVA in PASW for Windows was applied. However, before statistical comparison, the frequencies of all lexical bundles needed to be normalized and the statistically significant difference value (*p*) was set at .05 (Biber, 1995).

To analyze communicative functions of the extracted lexical bundles, we synthesized conceptual frameworks of numerous corpus scholars (e.g., Baoya, 2015; Biber et al., 1999; Getkham, 2010; Kanoksilpatham, 2005; Kitjaroenpaiboon, 2016; Kitjaroenpaiboon & Getkham, 2016a; 2016b) and found that lexical bundles could provide 31 communicative functions. They are desire, direction, intention, ability, introduction, elaboration, condition, identification, tangible, intangible, time, politeness, request, further communication, offer, expectation, hybrid function, specific reference, action, evaluation, claim, knowledge, purpose, contradiction, ownership, generality, commentary, modified information, references to present research, tentativeness, and reporting results. These 31 functions were applied as a framework to analyze communicative functions of the lexical bundles in this study. During this stage, to provide reliability, group discussions were held. We and two native English professors together studied contexts in which lexical bundles occur to analyze their communicative functions. The researchers agreed that a unanimous view is required to conclude the functional analysis process of each lexical bundle.

Results

In this study, there are three sub-corpora. The corpus of language teaching RAs comprises 228,891 words, the corpus of health sciences RAs consists of 101,967 words, and the corpus of business management RAs contains 302,552 words. The results of the analysis are presented in the particular order of the research objectives.

1) Lexical Bundles in Each Section of Language Teaching, Health Sciences, and Business Management Research Articles

Table 1. Number of lexical bundles in each section of the research articles from the three disciplines

Discipline	***Sections***				***Total***
	Introduction	***Methodology***	***Results***	***Discussion***	*(by excluding the repeatedly occurring bundles)*
Language Teaching	95	50	58	68	182
Health Sciences	14	43	15	53	108
Business Management	119	87	16	29	178

Table 1 details that in the language teaching RAs, 95 lexical bundles are found in the introduction sections, 50 lexical bundles are found in the methodology sections, 58 lexical bundles are found in the results sections, and 68 lexical bundles are found in the discussion sections. In the health sciences RAs, 14 lexical bundles are found in the introduction sections, 43 lexical bundles are found in the methodology sections, 15 lexical bundles are found in the results sections, and 53 lexical bundles are found in the discussion sections. In the business management RAs, 119 lexical bundles are found in the introduction sections, 87 lexical bundles are found in the methodology sections, 16 lexical bundles are found in the results sections, and 29 lexical bundles are found in the discussion sections.

2) Comparison of the Lexical Bundles in Each Section of Language Teaching, Health Sciences, and Business Management Research Articles

A comparative analysis, conducted to determine whether the extracted lexical bundles are similarly or differently used in each section of each discipline, reveals that in the language teaching RAs, a total of 182 lexical bundles were found. Nine lexical bundles are similarly found in all four sections, 14 lexical bundles are similarly found in three sections, 34 lexical bundles are similarly found in two sections, while 125 lexical bundles are found in one section. After receiving the frequency of the 182 lexical bundles in the language teaching RAs, we normalized their frequencies (to 100,000 words). Further, we analyzed them by one-way repeated measure ANOVA in PASW for Windows to determine whether their means are statistically significantly different (as shown in Table 2).

Table 2. One way repeated measure ANOVA for analyzing similarities of lexical bundles in the language teaching research articles

Source		**SS**	**df**	**MS**	***F***	***p***
Lexical bundles in each section of the language teaching RAs	Between groups	448.558	3	149.519	1.779	.150
	Within groups	45633.611	543	84.039		

Remark: $p > .05$

Table 2 details **no** statistically significant differences between the use of the lexical bundles in the four sections of the language teaching RAs ($F = 1.779$ and $p = .150$). It can be said that the lexical bundles in the four sections of the language teaching RAs are **not different**.

In the health sciences RAs, 108 lexical bundles are found. One lexical bundle is similarly found in all four sections, 14 lexical bundles are similarly found in two sections, while 93 lexical bundles are found in one section. Again, we analyzed the normalized frequencies by one-way repeated measure ANOVA (as shown in Table 3).

Table 3. One way repeated measure ANOVA for analyzing similarities of lexical bundles in the health sciences research articles

Source		SS	df	MS	*F*	*p*
Lexical bundles in each section of the health sciences RAs	Between groups	4192.491	3	1397.497	3.501	.016*
	Within groups	128122.467	321	399.125		

***Remark**: *p* < .05

Table 3 details some statistically significant differences between the use of the lexical bundles in the four sections of the health sciences RAs ($F = 3.501$ and $p = .016$). It can be said that lexical bundles in the four sections of the health sciences RAs are **different**. We further studied a pairwise comparison table (as shown in Table 4).

Table 4. Pairwise comparison

(I) Section	(J) Section	Mean Difference (I-J)	Std. Error	*p*
Introduction	Methodology	-4.461	2.866	.736
	Results	.459	2.639	1.00
	Discussion	-6.995*	2.474	.034*
Methodology	Results	4.921	2.615	.375
	Discussion	-2.534	3.025	1.000
Results	Discussion	-7.455*	2.655	.036*

***Remark** *p* < .05

Table 4 details some statistically significant differences between the use of lexical bundles in the introduction and the discussion sections and in the results and the discussion sections of the health sciences RAs ($p < .05$). However, there are **no** statistically significant differences in the introduction and the methodology sections, in the introduction and the results sections, in the methodology and the results sections, and the methodology and the discussion sections ($p > .05$).

In the business management RAs, 178 lexical bundles are found. Three lexical bundles are similarly found in all four sections, 17 lexical bundles are similarly found in three sections, 30 lexical bundles are similarly found in two sections, while 128 lexical bundles are found in one section. We further analyzed the normalized frequencies by one-way repeated measure ANOVA (as shown in Table 5).

Table 5. One way repeated measure ANOVA for analyzing similarities of lexical bundles in the business management research articles

Source		SS	df	MS	*F*	*p*
Lexical bundles in each section of the business management RAs	Between groups	4197.314	3	1399.104	15.403	.00*
	Within groups	48233.678	531	90.835		

Remark *$p < .05$

Table 5 details some statistically significant differences between the use of the lexical bundles in the four sections of the business management RAs ($F = 15.403$ and $p < .001$). It can be said that lexical bundles in the four sections of the business management RAs are **different**. We further studied a pairwise comparison (as shown in Table 6).

Table 6. Pairwise comparison

(I) Section	(J) Section	Mean Difference (I-J)	Std. Error	*p*
Introduction	Methodology	-.410	1.128	1.000
	Results	5.353*	.889	0.00*
	Discussion	3.611*	.880	0.00*
Methodology	Results	5.764*	1.004	0.00*
	Discussion	4.022*	1.073	0.001*
Results	Discussion	-1.742	1.059	.611

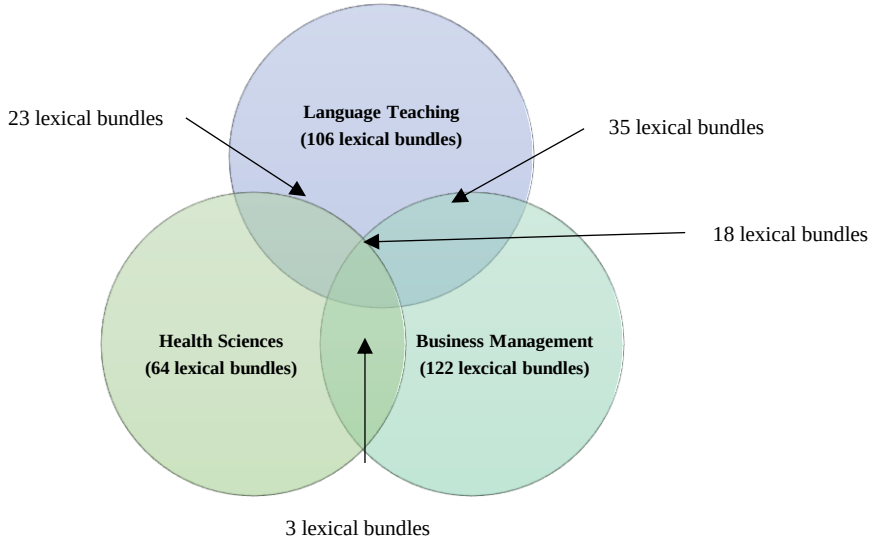
Remark *$p < .05$

Table 6 details some statistically significant differences between the use of lexical bundles in the introduction and the results sections, in the introduction and the discussion sections, in the methodology and the results sections, and the methodology and the discussion sections of the business management RAs ($p < .05$). However, there are **no** statistically significant differences in the introduction and the methodology sections and in the results and the discussion sections ($p > .05$).

3) Comparison of the Lexical Bundles in the Language Teaching, Health Sciences, and Business Management Research Articles

A comparative analysis, again conducted to determine whether the extracted lexical bundles are similarly or differently used between the three disciplines, reveals 371 lexical bundles in the RAs. Eighteen lexical bundles are similarly found in all three disciplines, 23 lexical bundles are similarly found in the language teaching and the health sciences, 35 lexical bundles are similarly found in the language teaching and the business management, and three lexical bundles are similarly found in the health sciences and the business management. One-hundred and six lexical bundles are frequently found in only the language teaching, 64 lexical bundles are frequently found in only the health sciences, and 122 lexical bundles are frequently found in only the business management (as shown in Figure 1).

Figure 1. Numbers of lexical bundles found in the research articles from the three disciplines



We then analyzed normalized frequencies of the 371 lexical bundles (N = 371) by one-way repeated measure ANOVA (as shown in Table 7).

Table 7 One way repeated measure ANOVA for analyzing similarities of lexical bundles in the language teaching, health sciences, and business management research articles

Source		SS	df	MS	*F*	*p*
Lexical bundles in the RAs	Between groups	23.536	2	11.768	.376	.687
from the three disciplines	Within groups	23181.306	740	31.326		

Remark *p > .05*

Table 7 details **no** statistically significant differences between the use of the lexical bundles in the RAs from the three disciplines (F = .376 and *p* = .687). It can be said that lexical bundles used in the language teaching, health sciences, and business management RAs are **not different**.

4) Communicative function of the Lexical Bundles Frequently in the Language Teaching, the Health Sciences, and the Business Management Research Articles

Applying conceptual frameworks of corpus scholars (i.e., Baoya, 2015; Biber et al., 1999; Getkham, 2010; Kanoksilpatham, 2005; Kitjaroenpaiboon, 2016; Kitjaroenpaiboon & Getkham, 2016a; 2016b) to analyze communicative functions of the 371 lexical bundles, we found that the lexical bundles in this study provide 19 communicative functions namely (1) Action, (2) Evaluation, (3) Identification, (4) Reporting Results, (5) Knowledge, (6) Specific Reference, (7) Time, (8) Commentary, (9) Contradiction, (10) Ownership, (11) Tentativeness, (12) Ability, (13) Claim, (14) Direction, (15) Intangible, (16) Tangible, (17) Elaboration, (18) References to Present Research, and (19) Hybrid Function (as shown in Table 8).

Table 8. Lexical bundles and their communicative functions

Communicative Function	**Discipline**		
	Language Teaching RAs (182 lexical bundles)	**Health Sciences RAs (108 lexical bundles)**	**Business Management RAs (178 lexical bundles)**
1) Action (37 lexical bundles)	- *are presented in* - *are shown in* - *by the first* - *is illustrated in* - *measured by the* - *occurred in the* - *participated in the* - *presented in table* - *used in the* - *was used as* - *was used to* - *were asked to* - *were used in* - *were used to*	- *assessed by using* - *analyses were performed* - *by using a* - *by using the* - *calculated as the* - *did not include* - *included in the* - *used to assess* - *used to identify* - *was approved by* - *was defined as* - *was obtained from* - *was used for* - *was used to* - *were classified as* - *were not included* - *were obtained from* - *were used to* - *written Informed consent was obtained*	- *are driven by* - *by showing that* - *can be used* - *examine the effects of* - *paper is organized as follows* - *paper proceeds as follows*
2) Evaluation (7 lexical bundles)	- *considered to be*	- *considered statistically significant* - *shown to be*	- *are robust to* - *prior to the*

		- *significantly associated with* - *significantly higher in*	
3) Identification (14 lexical bundles)	- *each of the* - *for each of* - *of the same* - *of these studies* - *one of the* - *participants in the* - *research on the* - *the participants in* - *this type of*	- *in patients with* - *one of the* - *with type .. diabetes* - *of at least* - *of patients with*	- *each of the* - *of the sample* - *one of the*
4) Reporting Results (17 lexical bundles)	- *and found that* - *results of the* - *results showed that* - *results suggest that* - *showed that the* - *shown in table* - *shows that the* - *table … shows the* - *the findings of* - *there was a* - *there was no*	- *results suggest that* - *findings suggest that* - *no significant differences* - *significant differences in* - *there was a* - *there was no* - *there were no*	- *also found that* - *and found that* - *and show that* - *results suggest that* - *shows that the*
5) Knowledge (13 lexical bundles)	- *according to the* - *few studies have* - *in line with* - *in relation to* - *in the literature*	- *according to the* - *few studies have* - *in accordance with* - *in line with* - *in previous studies*	- *according to the* - *in response to* - *in the previous*

	- refers to the *- studies suggest that*	*- in relation to* *- not associated with* *- studies have shown*	
6) Specific Reference (8 lexical bundles)	*- a second language* *- English as a second language* *- learners of English* *- native English speakers* *- native speakers of English*	*- body mass index* *- children and adolescents*	*- difference in differences*
7) Time (7 lexical bundles)	*- at the same time* *- at the time of*	*- at the time of* *- after adjustment for*	*- a given year* *- at the same time* *- at the time of* *- during the sample period* *- the sample period* *- the time of*
8) Commentary (2 lexical bundles)		*- to our knowledge*	*- for future research*
9) Contradiction (2 lexical bundles)	*- did not differ*	*- did not differ*	*- in contrast to*
10) Ownership (7 lexical bundles)	*- we found that*	*- we found that*	*- we also found that* *- we analyze the* *- we control for* *- we examine the* *- we find that* *- we focus on* *- we found that*
11) Tentativeness (10 lexical bundles)	*- likely to be* *- more likely to*	*- likely to be* *- more likely to*	*- less likely to* *- likely to be*

	- it may be *- it would be* *- may not be* *- more or less* *- seems to be*	*- may have been* *- may not be*	*- more likely to* *- the probability of*
12) Ability (5 lexical bundles)	*- be able to* *- the ability to* *- their ability to* *- were able to*		*- allows us to* *- the ability to* *- their ability to*
13) Claim (8 lexical bundles)	*- this suggests that*	*- is the first* *- the first study to* *- this is the first study*	*- contributes to the literature* *- paper contributes to* *- this leads to* *- to the extent that* *- this suggests that*
14) Direction (3 lexical bundles)	*- it should be*	*- are needed to* *- should be interpreted*	
15) Intangible (28 lexical bundles)	*- analysis of the* *- depending on the* *- in the context of* *- in the discourse* *- in the following* *- information about the* *- knowledge of the* *- of the original* *- on the basis of* *- scores on the* *- the meaning of*	*- an increased risk of* *- on the basis of*	*- a function of* *- a measure of* *- an increase in* *- analysis of the* *- average number of* *- control for the* *- data from the* *- depending on the* *- depends on the* *- information about the* *- in the context of*

	- understanding of the *- with each other* *- with regard to*		*- in the first* *- in the form of* *- in the next* *- in two ways* *- of the firm* *- the distribution of*
16) Tangible (38 lexical bundles)	*- a group of* *- a number of* *- a series of* *- a set of* *- a variety of* *- all of the* *- as part of* *- at least in* *- at least one* *- changes in the* *- compared to the* *- in the case of* *- the age of* *- the case of* *- the following research questions* *- the nature of* *- the next section* *- the number of* *- the proportion of* *- the second author*	*- a number of* *- a total of* *- in a large* *- as part of* *- compared with the* *- the association between* *- the number of* *- version of the* *- the general population* *- with respect to*	*- a large number of* *- a number of* *- a sample of* *- a variety of* *- an indicator variable* *- at least one* *- change in the* *- changes in the* *- data for the* *- descriptive statistics for* *- in the sample* *- the dependent variable is* *- the nature of* *- the number of* *- the proportion of* *- value of the* *- variables used in* *- the internet appendix* *- with respect to*

17) Elaboration (11lexical bundles)	- *as well as* - *as a result* - *because of the* - *can be seen* - *due to the* - *in addition to* - *in other words* - *on the other hand*	- *as well as* - *be due to* - *be explained by* - *because of the* - *in addition to*	- *and in turn* - *as a result* - *as well as* - *due to the* - *in addition to* - *on the other hand*
18) References to Present Research (13 lexical bundles)	- *for this study* - *in a study* - *in the analysis* - *in the present study* - *in the study* - *in these studies* - *in this study* - *the current study* - *the present study*	- *in the current study* - *in the study* - *in this study* - *the current study*	- *in this case* - *in this paper* - *in this section*
19) Hybrid Function (141 lexical bundles)			
19.1) Hybrid Function: Evaluation and Tentativeness (1 lexical bundle)	- *appear to be*		
19.2) Hybrid Function: Action and Knowledge (4 lexical bundles)	- *as compared to* - *as opposed to* - *as shown in*	- *as described previously*	
19.3) Hybrid Function: Action, Tangible, and Intangible	- *followed by a* - *followed by the*		

(2 lexical bundles)			
19.4) Hybrid Function: Evaluation and Knowledge (6 lexical bundles)	*- similar to the* *- related to the*	*- related to the*	*- associated with a* *- associated with the* *- consistent with the* *- relative to the*
19.5) Hybrid Function: Evaluation and Ownership (1 lexical bundle)			*- consistent with our*
19.6) Hybrid Function: Evaluation and Claim (1 lexical bundle)			*- consistent with this*
19.7) Hybrid Function: Evaluation and Reporting Results (1 lexical bundle)	*- found to be*		
19.8) Hybrid Function: Evaluation, Intention, Claim, and Purpose (1 lexical bundle)	*- It is important to*		*- It is important to*
19.9) Hybrid Function: Evaluation and Tentativeness (1 lexical bundle)	*- it is possible*		
19.10 Hybrid Function: Identification and Tangible	*- for the first*		

(1 lexical bundle)			
19.11) Hybrid Function: Identification and References to Present Research (3 lexical bundles)	- *of the present* - *of this study*	- *of the study* - *of this study*	
19.12) Hybrid Function: Reporting Results and Knowledge (3 lexical bundles)	- *have shown that* - *they found that*	- *has been shown* - *have shown that*	- *they found that*
19.13) Hybrid Function: Reporting Results, Tangible, and Intangible (3 lexical bundles)	- *the results for* - *the results of*	- *the results of*	- *a result of* - *the results of*
19.14) Hybrid Function: References to Present Research and Ownership (6 Lexical bundles)	- *in our study*	- *in our study*	- *in our analysis* - *in our data* - *in our sample* - *of our results* - *our second hypothesis*
19.15) Hybrid Function: Knowledge and Generality (2 lexical bundles)	- *there is a* - *there is no*		- *there is a* - *there is no*
19.16) Hybrid Function: Time, References to Present			- *our sample period*

Research and Ownership (1 lexical bundle)			
19.17) Hybrid Function: Intention and Purpose (21 lexical bundles)	- *be used to* - *in order to* - *is needed to* - *study is to* - *study was to* - *to determine the* - *to engage in* - *to ensure that* - *to note that* - *to investigate the*	- *in order to* - *to account for* - *to assess the* - *to examine the* - *to identify the*	- *the decision to* - *to address this* - *to capture the* - *to control for* - *to estimate the* - *to examine the* - *to examine whether* - *to reduce the*
19.18) Hybrid Function: Ownership and Commentary (6 lexical bundles)		- *we showed that* - *we hypothesized that*	- *we assume that* - *we argue that* - *we estimate the* - *we show that*
19.19) Hybrid Function: Ownership and Action (2 lexical bundles)			- *we use a* - *we use the*
19.20) Hybrid Function: Tentativeness, Tangible, and Intangible (1 lexical bundle)			- *the likelihood of*
19.21) Hybrid Function: Tentativeness and Evaluation			- *tend to be*

(1 lexical bundle)			
19.22) Hybrid Function: Ability Tangible, and Intangible (1 lexical bundle)			*- the ability of*
19.23) Hybrid Function: Claim and Reporting results (1 lexical bundle)			*- provide evidence that*
19.24) Hybrid Function: Direction and Intention (1 lexical bundle)	*- need to be*	*- need to be*	
19.25) Hybrid Function: Intangible and Time (2 lexical bundle)	*- at the end of*		*- at the beginning of* *- at the end of*
19.26) Hybrid Function: Tangible and Intangible (65 lexical bundles)	*- a range of* *- between the two* *- difference between the* *- differences between the* *- differences in the* *- in light of* *- in terms of* *- in the same* *- main effect of*	*- an association between* *- differences in the* *- the cross sectional* *- the development of* *- the effect of* *- the effects of* *- the prevalence of* *- the relationship between* *- the risk of* *- the use of*	*- a decrease in* *- a form of* *- a source of* *- at the level* *- difference between the* *- differences between the* *- focus on the* *- half of the* *- in terms of* *- in the same*

	- most of the *- part of a* *- part of the* *- some of the* *- structure of the* *- the absence of* *- the acquisition of* *- the comparison of* *- the degree of* *- the development of* *- the effect of* *- the effects of* *- the existence of* *- the extent to which* *- the lack of* *- the majority of* *- the presence of* *- the process of* *- the question of* *- the relationship between* *- the total number of* *- the use of*		*- leads to a* *- most of the* *- part of the* *- some of the* *- the absence of* *- the change in* *- the coefficient on* *- the cost of* *- the demand for* *- the difference between* *- the effect of* *- the effects of* *- the existence of* *- the extent of* *- the extent to which* *- the fraction of* *- the impact of* *- the interests of* *- the introduction of* *- the level of* *- the literature on* *- the percentage of* *- the presence of* *- the quality of* *- the ratio of* *- the relation between*

			- *the remainder of* - *the rest of* - *the risk of* - *the role of* - *the sample to* - *the sensitivity of* - *the size of* - *the total number of* - *the use of* - *the value of*
19.27) Hybrid Function: Purpose, Tangible, and Intangible) (1 lexical bundle)	- *the purpose of*		
19.28) Hybrid Function: Tangible, Intangible, and Knowledge (1 lexical bundle)	- *based on the*	- *based on the*	- *based on the*
19.29) Hybrid Function: Intangible and Claim (1 lexical bundle)	- *the importance of*	- *the importance of*	- *the importance of*
TOTAL	**371 lexical bundles**		

Table 8 shows the 19 communicative functions of the 371 lexical bundles. All communicative functions derive from the communicative functions of lexical bundles as proposed by corpus scholars (i.e., Baoya, 2015; Biber et al., 2004; Getkham, 2010; Kanoksilpatham, 2003; Kitjaroenpaiboon, 2016; Kitjaroenpaiboon & Getkham, 2016a; 2016b). The functional analysis reveals that some lexical bundles provide one communicative function, while some provide two or more functions (known in this study as Hybrid Function) depending on the context in which they are found. For instance, *'the importance of'* has two communicative functions. They are Intangible and Claim.

Conclusion and Discussion

A total of 182 lexical bundles are found in language teaching RAs. The findings differ from Hyland (2008) who found 20 lexical bundles in applied linguistic RAs. The difference might be possibly due to the number of words contained in the lexical bundles. Hyland (2008) investigated four-word lexical bundles in his corpus while this study analyzes three- to five-word bundles.

A total of 108 lexical bundles are revealed in the health sciences RAs. It is also discrepant from Panthong and Poonpon's (2020) finding of 67 lexical bundles in medical RAs, Kwary et al.'s (2017) finding of 62 lexical bundles in health sciences RAs, and Cortes' (2004) finding of four lexical bundles in biology RAs. A possible reason affecting the differences might be that those studies also focused on exploring four-word bundles.

Additionally, a total of 178 lexical bundles are found in the business management RAs. The findings are discrepant from Damshevska (2019) who found 40 lexical bundles in business economic RAs. The difference might be because Damshevska (2019) focuses only on investigating four-word lexical bundles in his corpus.

It could be said that the number of words contained in a lexical bundle results in the findings. In other words, the fewer the numbers of words a lexical bundle comprises, the more frequently a lexical bundle is found and vice versa (Cardinali, 2015; Neely & Cortes, 2011; Nesi & Basturkmen, 2006).

Our observation is that 40 lexical bundles (i.e., *as well as, based on the, in order to, in terms of, in this study, one of the, the number of, the relationship between, the use of, a number of, in other words, in relation to, it is important to, most of the, on the other hand, related to the, showed that the, some of the, there is a, of this study, in addition to, the present study, as a result, be able to, compared to the, differences in the, found to be, in light of, in the context of, it may be, likely to be, may not be, more likely to, need to be, the ability to, the development of, the effects of, the importance of, the lack of, the nature of*) are similarly found in the introduction and in the discussion sections of the language teaching RAs. Similar to Bal (2010), lexical bundles found in the introduction sections are similarly found in the discussion sections. This might be because a discussion section plays the role of a mirror reflecting the content provided in an introduction section. RA writers always mention other previous studies and compare their findings with others in these sections (Lim, 2005). **No** statistically significant differences between the use of the 182 lexical bundles in all four sections of the language teaching RAs are found ($p = .150$). This implies that lexical bundles in all four sections of the language teaching RAs are quite **similar**. The findings are in line with Kitjaroenpaiboon and Getkham (2016b) who reported that linguistic characteristics in all four sections of language teaching RAs are quite identical.

Another observation is that three lexical bundles (i.e., *as well as, in order to, one of the)* are similarly found in the introduction and the method sections of the health sciences RAs. Three lexical bundles (i.e., *as well as, a total of, included in the*) are similarly found in the method and the results sections. Three lexical bundles (i.e., *as well as, compared with the, not associated with*) are similarly found in the results and the discussion sections. This might be because stylistic patterns in introduction are similar to ones in the method sections, ones in the method are similar to ones in the results, and ones in the results are similar to ones in the discussion. Writers provide an overview of a research methodology in the introduction and

detail the methodology again in the method sections, explain populations in the study in the method and the results sections, and provide conclusions of the results in the discussion sections (Misak et al., 2005). Some statistically significant differences between the use of the 108 lexical bundles in all four sections of the health sciences RAs are found (p = .016). This implies that the use of lexical bundles in some sections of the health sciences RAs are **different**. The findings differ from Bineta (2016) and Kitjaroenpaiboon and Getkham (2016a) who reported that lexical bundles and linguistic structures are identically used in all four sections of medical and nursing RAs. The findings of this paper are discrepant from those two might be presumably due to the different disciplines analyzed. Kitjaroenpaiboon and Getkham (2016a) state that each discipline has its own specific stylistic pattern despite being in the same science.

The other observation is that 36 lexical bundles (i.e., *as well as, a number of, the effect of, the number of, based on the, changes in the, due to the, in terms of, in the same, in this case, the level of, the likelihood of, the probability of, the value of, consistent with the, relative to the, the change in, the impact of, a variety of, an increase in, as a result, at the time of, focus on the, in addition to, less likely to, most of the, one of the, prior to the, tend to be, the cost of, the effects of, the extent to which, the relation between, the time of, we focus on, with respect to*) are similarly found in the introduction and the methodology sections of the business management RAs. This might plausibly be because the content of an introduction section and a methodology section similarly presents an overview of research and discusses other previous studies (Weissberg & Buker, 1990). Some statistically significant differences between the use of the 178 lexical bundles in all four sections of the business management RAs are found (p < .001). This implies that the use of lexical bundles in some sections of the business management RAs are **different**. However, the findings differ from Betul (2019) who found that similar lexical bundles are applied through all four sections of the economic RAs.

A total of 371 lexical bundles are found in the language teaching, health sciences, and business management RAs. **No** statistically significant differences between the use of the 371 lexical bundles in all RAs are found (p = .687). This implies that the use of lexical bundles in the three disciplines are quite **similar**. The findings differ from Kwary et al. (2017) who reported that lexical bundles found in health science RAs differ from lexical bundles found in social sciences. However, they are in line with Betul (2019) in that similar lexical bundles are found through the economic, the educational, the history, the medical, the psychological sciences, and the sociology RAs.

We also found that the lexical bundles with functions such as desire, introduction, elaboration, condition, politeness, request, further communication, offer, and expectation of Biber et al. (2004) are not found in this study. This lack of functions might be because Biber et al. (2004) investigated spoken discourse; however, RAs is in written academic discourse of nature is formal and conventional (Bailey et al., 2004). Subsequently, lexical bundles with the above communicative functions are not found in the analysis.

In summary, the lexical bundles in each section of the RAs are both similar and different. Plausibly, a factor affecting similarities of the lexical bundle uses is that the language applied for writing in all sections of RAs is an academic language which is rather conventional and formal and differs from general language (Ranney, 2012). Presumably, a factor causing differences is the underlying communicative purposes of each section which result in different uses of lexical bundles in each section (Rao, 2018). The introduction section provides an overview of related works and the importance of a study. The methodology section elaborates a research design. The results section presents the findings. Meanwhile, the discussion section

presents interpretations and comparisons of the findings (Pho, 2008). For example, in this study *few studies have* is frequently found in the introduction sections to mention previous studies. *Participants in the* is frequently found in the methodology sections to refer to the research populations. *Are shown in* is frequently found in the results sections to present an informational table. Also, *we found that* is frequently found in the discussion sections to conclude and compare findings with others.

That the 371 lexical bundles in these three disciplines are quite similar has been confirmed by **no** statistically significant difference (p = .687). Hyland (2012) says that academic language is always used for writing RAs in all disciplines and shares some stylistics in common among the disciplines. Academic language is always formal, conventional, and applied some similar lexical bundles (Hyland, 2007). For instance, *according to the*, *based on the*, and *results suggest that* are generally found in academic language. However, each discipline has its use of some discipline-specific lexical bundles (Ranney, 2012). For example, in this study, *English as a second language* is found explicitly in the language teaching RAs. In comparison, *in patients with* and *children and adolescents* are specifically found in the health sciences RAs. In contrast, *descriptive statistics for* and *difference in differences* are only found in the business management RAs. The researchers view that these discipline-specific lexical bundles are not seen as frequently as general academic lexical bundles. Subsequently, **no** statistically significant differences between the use of the total of 371 lexical bundles in the RAs from these three disciplines were revealed.

We view that studying lexical bundles and their communicative functions is essential for learning academic language in each discipline and they provide non-native English, novice, and inexperienced researchers RA writing guidance. Therefore, before academics or researchers write a discipline-specific RA or even a multidisciplinary RA, a lexical bundle analysis can be used to determine how lexical bundles are needed to write a text and prepare the bundle lists accordingly. Thus, as noted by Chirobocea-Tudor (2018), understanding lexical bundles with their communicative functions is a helpful though daunting task to enhance comprehension and utilization of lexical bundles in a discipline-specific context in a particular field of study or even multidisciplinary context.

The relationships between lexical bundles and sections of RAs and between lexical bundles and disciplinary variations have been discussed in numerous studies. In the academic genre, for example, the existing research studies showed some different usages of lexical bundles across conventional sections as well as across disciplines in the written form (Damshevska, 2019; Hyland, 2008; Panthong & Poonpon, 2020; Wongwiwat, 2016). As stated by scholars (Huimin, 2010; Hyland, 2012), the distribution of bundles not only characterizes particular genres, but also is a section and disciplinary marker. This study adds considerable empirical evidences in viewing lexical bundles as an intradisciplinary and interdisciplinary marker. Within the RA genre, despite the use of lexical bundles being similarly distributed in a wide range when examining the disciplines with interest, some mild variations could be treated as distinction marks.

This study contributes to the research on intradisciplinary and interdisciplinary study with the examples from language teaching, health sciences, and business management RAs. The intradisciplinary and interdisciplinary relationships, as revealed in the results, sees rather significant convergence compared with the divergent usages.

Limitations and Recommendations for Further Studies

This study helps shed light on lexical bundles in academic writing in the three specific disciplines (i.e., language teaching, health sciences, and business management). All data sets were retrieved from the disciplines. Accordingly, generalizing of the results is restricted to those specific corpora. For further studies, future research might be conducted to determine how the data-driven approach can best be facilitated in English for Academic Purposes or English for Specific Purposes instruction. This can contribute to teaching academic writing.

Acknowledgements

Academically, we have always been supported by Associate Professor Dr. Kanyarat Getkham, a specialist in corpus linguistics as well as applied linguistics, and are grateful for her valuable, helpful, and constructive comments on our research study. We are also deeply thankful to Miss Sunee Steyn for her assistance in academicizing our language.

The Authors

Woravit Kitjaroenpaiboon graduated a Doctor of Philosophy (Language and Communication) from Graduate School of Language and Communication, National Institute of Development, Thailand. Currently, He is a teacher at English Language Curriculum, Faculty of Humanities and Social Sciences, Suan Dusit University. His e-mail address is woravit_kit@dusit.ac.th.

Sresuda Wongwiseskul graduated a Doctoral Degree of Education (Higher Education) from Faculty of Education, Chulalongkorn University, Thailand. Currently, she is dean of Faculty of Nursing, Suan Dusit University. Her e-mail address is sresuda_won@dusit.ac.th.

Thanakorn Puksa graduated a Doctoral Degree of Management (Business Management) from School of management Sciences, Suan Dusit University, Thailand. Currently, he is a deputy dean of School of Management Sciences, Suan Dusit University, Thailand. His e-mail address is thanakorn_puk@dusit.ac.th.

Benjamas Khamsakul graduated a Doctor of Philosophy (Thai Language) from Faculty of Humanities, Kasetsart University, Thailand. Currently, she is a teacher at Thai Language Curriculum, Faculty of Humanities and Social Sciences, Suan Dusit University. Her e-mail address is benjamas_kha@dusit.ac.th.

References

Anthony, L. (2020). *AntConc (Version 3.5.9)* [Computer Software]. Waseda University. Available from https://www.laurenceanthony.net/software.

Bailey, A. L., Butler, F. A., LaFramenta, C., & Ong, C. (2004). *Towards the characterization of academic language in upper elementary science classrooms.* Center for the Study of Evaluation National Center for Research on Evaluation, Standards, and Student Testing Graduate School of Education & Information Studies, University of California, Los Angeles.

Bal, B. (2010). *Analysis of four-word lexical bundles in published research articles written by Turkish scholars*. [Unpublished master's thesis, Georgia State University].

Baoya, Z. (2015). *Moves and inter-move linguistic variation in education research articles.* [Unpublished doctoral dissertation, Suranaree University of Technology].

Berkenkotter, C., Huckin, T., & Ackerman, J. (1991). *Social context and socially constructed texts: The initiation of a graduate student into a writing research community*. University of Wisconsin Press.

Betul, B. (2019). Lexical bundles in published research articles: A corpus-based study. *Journal of Language and Linguistic Studies, 15(*2), 520-534.

Biber, D. (1995). *Variation across speech and writing.* Cambridge University Press.

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., & Quirk, R. (1999). *Longman grammar of spoken and written English.* Longman.

Bineta, M. (2016). *Lexical bundles in medical research articles: Structures and functions.* [Unpublished master's thesis, Michigan State University].

Boutellier, R., Gassmann, O., & Raeder, S. (2011). *What is the difference between social and natural sciences?.* University of St. Gallen.

Cardinali, R. (2015). Lexical bundles in biology research articles: Structure and function across corpora and sections. *Indonesian Journal of Applied Linguistics, 7*(3), 1-19.

Cargil, M., & O'Connor, P. (2009). *Writing scientific research articles: Strategies and steps.* Wiley-Blackwell.

Chirobocea-Tudor, O. (2018). *Vocabulary acquisition in ESP: Perspectives, strategies and resources*. Retrieved July 18, 2019, from https://www.researchgate.net/publication/330324952_ Vocabulary_ Acquisition_in_ESP_ Perspectives_Strategies_and_Resources

Conrad, S. (1996). Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education*, *8*(3), 299-326.

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. (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.

Damshevska, N. (2019) Method sections of management research articles: A pedagogically motivated qualitative study. *Article in English for Specific Purposes*, *25*(3), 282-309.

Fadda, H. (2012). Difficulties in academic writing: From the perspective of King Saud University postgraduate students. *English Language Teaching*, *5*(3), 123-130.

Genc, B., & Bada, E. (2010). English as a world language in academic writing. *The Reading Matrix*, *10*(2), 142-151.

Getkham, K. (2010). *A corpus-based study of applied linguistic research articles: A multidimensional analysis.* National Institute of Development Administration.

Huimin, S. (2010). *Lexical bundles and moves in agricultural science research articles.* [Unpublished master's thesis, Suranaree University of Technology].

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, *16*(3), 148-164.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(2), 4-21.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, *32*(2), 150-169. http://dx.doi.org/10.1017/S0267190512000037

Joseph, D., Jones, J., Tennant, F., & Hook, S. A. (2010). Academic literacy: The importance and impact of writing across the curriculum – a case study. *Journal of the Scholarship of Teaching and Learning*, *10*(2), 34-47.

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, *24*(3), 269-292.

Kitjaroenpaiboon, W. (2016). *A comparative corpus-based study of language teaching and nursing research articles: A multidimensional analysis.* [Unpublished doctoral dissertation, National Institute of Development Administration].

Kitjaroenpaiboon, W., & Getkham, K. (2016a). Patterns of linguistic features and their communicative functions in nursing research articles. *International Journal of Management and Applied Science*, *3*(3), 98-103.

Kitjaroenpaiboon, W., & Getkham, K. (2016b). Stylistic patterns in language teaching research articles: A multidimensional analysis. *PASAA: Journal of Language Teaching and Learning*, *52*(1), 169-208.

Kwary, D.A., Ratri, D., & Artha, A. (2017). Lexical bundles in journal articles across academic disciplines. *Indonesian Journal of Applied Linguistics*, *7*(1), 132-140.

Lim, J. M. H. (2005). *Reporting results in management journal articles.* Universiti Putra Malaysia Press.

Lindstromberg, S., & Eyckmans, J. (2017). The power of sound in L2 vocabulary learning. *Language Teaching Research*, *21*(3). 341-361.

McLaughlin, E., & Parkinson, J. (2018). We learn as we go: How acquisition of technical vocabulary is supported during vocational training. *English for Specific Purposes*, *50*(4), 14-27.

Misak, A., Marusic, M., & Marusic, A. (2005). Manuscript editing as a way of teaching academic writing: Experience from a small scientific journal. *Journal of Second Language Writing*, *14*(3), 122-131.

Morrison, J. (2014). Incoming NSF director faces challenges in Congress. *Nature*, *507*(1), 285.

Nasrabady, P., Shirvan, M. E., & Golparvar, S. E. (2020). Exploring lexical bundles in recent published papers in the field of applied linguistics. *Journal of World Languages*, *6*(3), 175-197. https://doi.org/10.1080/21698252.2020.1797992

Neely, E., & Cortes, V. (2011). A little bit about: Analyzing and teaching lexical bundles in academic lectures. *Language Value*, *1*(1), 17-38.

Nesi, H., & Basturkmen, H. (2006). Lexical bundles and discourse signalling in academic lectures. *International Journal of Corpus Linguistics*, *11*(3), 283-304.

Panthong, P., & Poonpon, K. (2020). Lexical bundles in Thai medical research articles. *Journal of Studies in the English Language*, *15*(1), 59-106.

Pho, P. D. (2008). Research article abstracts in applied linguistics and educational technology: A study of linguistic realizations of rhetorical structure and authorial stance. *Discourse Study*, 3(2), 78-95.

Poggensee, A. (2016). *The effects of globalization on English language learning: Perspectives from Senegal and the United States.* Western Michigan University.

Ranney, S. (2012). Defining and teaching academic language: Developments in K-12 ESL. *Language and Linguistics Compass*, *6*(9), 560-574.

Rao, C. (2018). The use of English language in research. *Journal for Research Scholars and Professionals of English language Teaching*, *8*(2), 101-135.

Scimagojr.com. (2019). *Homepage on the internet*. New York: Scimago Institution Ranking [cited 2019 Oct 3]. Retrieved May 8, 2019 from https://www.scimagojr.com/

Simone, V. G., Misha, V., Jacob A. P., Watts, S., Janey S. P., Preeti, R., & Jeremy, M. D. (2018). The Importance of interdisciplinary research training and community dissemination. *Interdisciplinary Research Education*, *8*(5), 611-614. DOI: 10.1111/cts.12330.

Swales, J. M. (1992). *Aspects of article introductions*. University of Michigan Press.

Weissberg, R., & Buker, S. (1990). *Writing up research: Experimental research report writing for students of English.* Prentice Hall.

Wongwiwat, T. (2016). Move analysis and lexical bundles analysis of conference abstract: A case study of Thailand TESL international conferences. [Unpublished doctoral dissertation, Thammasat University].