



## **The Effects of Using Pedagogic Corpus as Target Formulas for Direct Instruction in an Academic Writing Class**

**Nor Ashikin Ab Manan<sup>1</sup>, Paramaswari Jaganathan<sup>2</sup> and Ambigapathy Pandian<sup>2</sup>**

<sup>1</sup>Academy of Language Studies, Universiti Teknologi MARA (Perak), 32610 Bandar Baru Seri Iskandar, Perak, Malaysia

<sup>2</sup>School of Languages, Literacies and Translation, Universiti Sains Malaysia, 11800 Pulau Pinang, Malaysia

Corresponding author: Nor Ashikin Ab Manan, Academy of Language Studies, Universiti Teknologi MARA (Perak), 32610 Bandar Baru Seri Iskandar, Perak, Malaysia

---

**Abstract.** The purpose of this study is to investigate the effects of academic formula instruction using target formulas chosen from pedagogic corpus, on the use of the target formulas in an academic essay writing test, and the academic essay writing test scores. Two intact groups of Diploma in Computer Science students assigned as the experimental and control groups participated in the study. Each group consists of forty mixed ability ESL learners who were enrolled in an academic writing course. The study addressed two research questions. (1) What are the effects of formula instruction on the students' use of the target formulas? (2) What are the effects of formula instruction on the students' academic writing performance?

**Keywords:** Academic formulas, Formula instruction, Pedagogic corpus, Writing Performance

---

## 1. INTRODUCTION

Many researchers in the field of second language acquisition (SLA) have acknowledged the importance of formula acquisition for native-like competence and fluency (refer to Boers, Eykmans, Kappel, Stengers, Demecheleer, 2006; Boers & Lindstromberg, 2012; Roever, 2012; Wood, 2010; Wray, 2002;), and there is evidence which suggests that L2 learners can gain a lot of benefits from mastering formulaic language since knowledge of formulas correlates significantly with L2 proficiency level (Dai & Ding, 2010; Kennedy & Thrope, 2007; Keshavarz & Salimi, 2007; Ohlrogge, 2009). However, there has been no conclusive agreement on which of these formulaic expressions should be directly taught in the second language classrooms, and the most suitable approach to formula instructions.

### *1.2 Objectives of the Study and Research Questions*

The purpose of this study is to investigate whether direct instruction of academic formulas chosen from “pedagogic corpus” (Willis, 2003, p.163) is beneficial in improving ESL learners’ academic writing ability. The objectives of the study are to determine the effects of direct teaching of academic formulas chosen from pedagogic corpus on the use of the target academic formulas (TAF) in the academic essay writing test and the students’ academic essay writing test scores.

The study aims at addressing two research questions. (1) What are the effects of formula instruction on the students’ use of the TAF? (2) What are the effects of formula instruction on the students’ academic writing performance?

## 2. LITERATURE REVIEW

### *2.1 The Importance of Academic Writing to Tertiary Level Learners*

Academic writing skill is important for tertiary level learners since students’ academic performance is evaluated mostly based on written works (Kelley, 2008; O’Ferrell, 2005), and academic writing is a literacy practice which connects the students’ admission into their disciplinary communities and the acquisition of the formal conventions associated with them (Leibowitz, Goodman, Hannon & Parkerson, 1997). The main characteristics of written academic English are its formal style of expression and precise word choice (Coffin et al., 2003) and one of the

defining features of academic prose is academic vocabulary. As stressed by Schoonen, Van Gelderen, Stoel, Hulstijn, De Glopper (2001, p. 33) academic writing draws heavily on linguistic resources a writer has and “a large vocabulary and a rich and flexible repertoires of sentence frames” will assist the writer to be clear and concise in his writing.

However, it has been reported that many Malaysian undergraduates lack both receptive and productive vocabulary knowledge expected for tertiary level studies (Jamian, Sidhu & Muzafar, 2008; Mathai, Jamian & Nair, 2004; Mokhtar, 2010) which in turn contributes to their poor academic writing performance. Findings from research have established the need for academic vocabulary instruction to develop undergraduates’ proficiency in academic writing. Hinkel (2004) has proposed that in addition to grammar, academic vocabulary should also be explicitly taught in an academic writing class for ESL learners. Nevertheless, due to limited time allocated to developing academic writing at tertiary level, the selection of academic vocabulary to be explicitly taught has to be narrowed down in order for the proposal of direct instruction to be practical.

To address this matter the study has turned to second language acquisition (SLA) research which has accrued evidence on the highly formulaic nature of language based on research conducted in the fields of corpus linguistics and psycholinguistics (Biber, Conrad & Cortes, 2004; Biber & Barbieri, 2007; Conrad, 2008; Ellis, 1996; Erman & Warren, 2000; Foster, 2001; Howarth, 1998; Rayson, 2008; Sinclair, 1991; Wray, 2002). Since multiword lexis or formulas, “fulfil the same functions as single words” (Boers & Lindstromberg, 2012, p. 84), similar to vocabulary knowledge which has been found to be a strong predictor of general proficiency (Lewis, 2002; Schmitt, Jiang & Grabe, 2011; Singleton, 2000), L2 learners’ knowledge of multiword lexis has been found to correlate highly with proficiency level (Keshavarz & Salimi, 2007; Al-Zahrani, 1998; Zhang, 1993).

### *2.3 Vocabulary of Academic Prose*

Vocabulary used in academia is often made up of multiword combinations (Biber & Barbieri, 2007; Cortes, 2002, 2004, 2006; Coxhead & Byrd, 2007; Schmitt,

2004). Corpus driven research has been conducted by many researchers (Biber, 2006; Simpson-Vlach & Ellis, 2010; Hyland, 2012) in the field of language teaching to identify the most commonly used word combination or formulas in academic discourse. According to Hyland (2012, p.150)“...these sequences...are simply extended collocations that appear more frequently than expected by chance, helping to shape meanings in specific contexts and contributing to our sense of coherence in a text.” There are many different definitions and concepts of multiword unit and chunks but the term ‘formula’ in this study is adopted from the definition of formula proposed by Wray (2002, p. 9) who defines a formula as a sequence whether continuous or discontinuous of words or other elements, “...which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar”.

#### *2.4 The importance of Formula Instruction*

According to Coxhead and Byrd (2007) academic formulas are important to writers and speakers because their repetition offers users (particularly students) ready-made sets of words to work with. Meunier (2012) stresses that formula instruction seems relevant for three reasons: (a) formulaicity is ubiquitous in language (Sinclair, 1991; Rayson, 2008), (b) formulaic use has been shown to be a marker of proficiency in L2 (Cortes, 2004; Hyland 2008), and (c) studies have demonstrated that L2 language learners find formulaicity challenging (Ang, Abdul Rahim, Tan & Salehuddin, 2011; Naderishahab & Tahririan, 2013), thus teaching the formulas would help learners improve their knowledge and use of formulas which in turn would improve their proficiency level.

#### *2.5 Pedagogical Approach to Academic Formulas*

One of the constraints faced by academic writing teachers is limited time allocated for academic writing classes. Thus, to ensure practicality of direct instruction of academic formulas, the number of formulas to be directly taught has to be narrowed down to a manageable size. Among the major problems faced by academic writing teachers are on deciding ‘*which formulas to teach?*’ and ‘*how to*

*teach them?* Willis (2003) states that, “Pedagogically the main problem with phrases is that there are so many of them...” (166). Thus, this study has adopted the proposal by Sinclair and Renouf’s (1988) to focus on the common uses of the common words and Willis’s (2003, p. 163) suggestion of using “pedagogic corpus”, corpus made up of texts used in the classroom as the resource for formula instruction. This is advantageous as the lexical phrases or formulas chosen for teaching are extracted from texts that learners have already processed for meaning. In addition, selecting the formulas to teach in this manner would ensure better contextualization, increase relevance and promote higher level of motivation among learners.

The next concern is *‘how to teach’* these formulas. In the absence of a well-defined methodological framework, Granger (2011) suggested that lexical approach could be integrated progressively via “mini-action programmes” as proposed by Lewis (2000: 153). For example, ESL teachers could conduct local experiments which are integrated into the teachers’ preferred or imposed teaching curriculum. Researcher such as Kozłowski and Seymour (2003) have turned to current language learning theories which suggest that second language learners could store language in chunks and language patterns need to be heard, written, spoken and read repeatedly so that they would be imprinted in the learners long-term memory. Similarly, Wood (2002) also stressed that repeated exposure to written language that deals with specific content and in particular genre would facilitate comfort with written expression. Additionally, for most ESL learners the teaching materials used in the classroom provide the only context for ‘priming’ (Hoey, 2005). Thus, it is essential that ESL teachers provide ‘helpful’ priming in the classroom by utilizing materials that provide essential shortcuts to priming. Some of the examples given include usage notes, drilling exercises, texts or tapes with repeated instances of word sequence, collocational observations and illustrations.



instrument for the study was an Academic Essay Writing (AEW) test which was used to measure the academic writing performance. The AEW test requires the subjects to write an academic essay based on their understanding of the issues discussed in two articles of the same theme which were provided. The test was adapted from the Academic Writing course past years' final examination test paper for the Academic Writing course to ensure the subjects' familiarity with the test format. The data was collected by administering pre and post Academic Essay Writing (AEW) test at the beginning and the end of the study period respectively.

The independent variable for the experimental group was the Direct Instruction of Academic Formula (DIAF) which was the *treatment*, plus academic writing instruction employing process writing approach. The independent variable for the control group was the academic writing instruction employing process writing approach without the treatment. The dependent variables for the study were; (a) the scores of the pre and post Academic Essay Writing (AEW) test, and (b) the number of target academic formula used in the post AEW test.

### *3.2 Research Population and Samples*

The study involved two groups of Diploma in Computer Science students from a public Malaysian University assigned as the experimental and control groups. Each group consists of forty mixed ability ESL learners who were enrolled in an academic writing course (N=80). The Academic Writing course is the university's requirement for all diploma students in semester three. The pre requisites for this course are Preparatory English and Intermediate English which are offered in semester one and two respectively. Thus, the subjects for the study who were in their third semester had already passed both the pre requisite courses. They were from similar age group, ranging from 19 to 21 years old. They were homogeneous with regard to their mother tongue (Malay), cultural and educational background as well as the length of exposure to formal English as a second language (ESL) instruction.

Table 3.1: The Subjects' Demography

Group	Gender		Proficiency Level			Total (N)
	Male	Female	Advanced	Intermediate	Beginner	
Experimental	14	26	3	24	13	40
Control	15	25	4	23	13	40

Table 3.1 shows the subjects' demography. Their proficiency level was determined based on the results of the Intermediate English course which they attended in the second semester.

### 3.3 Selection of Target Academic Formula (TAF)

Thirty TAFs were selected from the Academic Formula List (AFL) by Simpson-Vlach and Ellis (2010). Table 3.2 shows the selected formulas identified from the top 200 Written AFL list as target formulas to be used in the study. The target formulas were selected based on Willis's (2003, p.163) proposal that formula instruction should include "pedagogic corpus". The criteria for formula selection for this study were as follows: the formula should be used at least once a) in the Academic Writing course's prescribed textbook entitled EAP Crossing Borders (Michael et al., 2010) and / or b) in the supplementary material for the course. Table 3.2 shows the target academic formulas.

Table 3.2: The Target Academic Formulas

in relation to [1]	can be/ is/ are affected by[11]	due to the fact that[21]
in response to [2]	give rise to[12]	as a consequence[22]
(from)( the) point of view (of) [3]	as well as[13]	as a result of[23]
to distinguish between[4]	more/less likely to[14]	due to the[24]
the relationship between[5]	there are (three/a few/many) [15]	can be achieved[25]
in conjunction with[6]	there are several[6]	appears to be/ does not appear to be[26]
according to the[7]	there is/are no[17]	there has been/there have been[27]
can be considered[8]	on the basis of[18]	a large number of[28]
a variety of[9]	in terms of (the) [19]	the number of[29]
with regard to[10]	in accordance with[20]	(there) are a number (of) [30]



Subjects from the experimental groups were exposed to the formulas through direct instruction while the subjects from the control groups were exposed to the formulas indirectly when using the course's prescribed textbook and the supplementary materials. Another consideration in TAF selection is that the formulas have to be recognizable to the subjects and are commonly found in teaching materials not only for the academic writing course but other courses in their academic discipline (i.e. lab report, term paper and written assignments).

### *3.5 Instructional Procedure*

Since the study was conducted among students who were attending an academic writing course, the experimental variable, DIAF, was designed to fit into the course's current syllabus with minor adjustment to the course's schedule and scheme of work.

Table 3.3: DIAF Exercise Schedule

Week	Time	Type of Activity
Three	2 hours	DIAF (Theme 1: Social Sciences and Humanities) Highlighting TAF in the RC passages Fill-in-the-blank/matching (sentence level) Sentence construction using the formulas
Five	2 hours	Practice for Theme 1 Develop Thesis statement & Topic sentences using TAF. Write introductory & body paragraph of a cause & effect essay.
Eight	2 hours	DIAF (Theme 2: Science and Technology) Highlighting TAF in the RC passages Fill-in-the-blank/matching (paragraph level) Sentence construction using the target formulas
Nine	2 hours	Practice for Theme 2 Develop Thesis statement & Topic sentences using TAF Write a draft of problem & solution essay
Ten	2 hours	Write full Essay based on Theme 1 (Writing a cause & effect essay)
Eleven	2 hours	Write full Essay based on Theme 2 (Writing a problem & solution essay)

Both the experimental and control group's lecturers were furnished with lesson plans prepared for fourteen weeks. The control group utilised the course's existing lesson plans while the experimental group's lesson plans for week 3, 5, 8, 9, 10 and 11 were designed to accommodate DIAF. Since minor adjustments were made to the scheme of work, the lesson plans were also adjusted accordingly. DIAF involves three types of activities which were conducted during the experimental period. Table 3.3 shows the activities conducted during the experimental period.

### *3.6 Data Collection*

Pre AEW test was conducted during the second week while post AEW test was conducted during the fourteenth week of the study period. The answer scripts for both pre and post AEW test were scored by two independent scorers who were not involved in the study. The marks awarded by both independent scorers were tabulated and averaged. The average scores were taken as the subjects' final pre and post AEW test scores. Marks allocation for the AEW test are as follows: (a) eight marks for content, (b) six marks for language, and (c) six marks for organization. The total scores were then divided by twenty and converted to 100%. After the second scorer had finished marking the post AEW test papers, the answer scripts were passed to the *third scorer*. The duty of the third scorer was to manually record the number of target academic formula used by the subjects.

## **4. DATA ANALYSIS AND FINDINGS**

### *4.1 Addressing the First Research Question*

To address the first research question, the frequency of TAF used in the post AEW test for the experimental and the control groups was compared. Figure 4.1 shows the frequency of TAF used in the AEW test for the experimental and the control group based on a scale. The use of 0 to 4 TAF is considered low (L), the use of 5 to 8 TAF is considered as moderate (M) while the use of more than 8 TAF is considered as high (H). Figure 4.1 shows the frequency of TAF used in the AEW test for the experimental and the control group based on a scale. Based on Figure 4.1 it can be seen that the experimental group used more TAF compared to the control group. Sixteen subjects from the experimental group are considered low (L) users of

the TAF, twenty subjects fall into the category of moderate user (M) while four subjects are considered as high users (H). On the other hand, thirty one subjects of the control group are low users (L) while nine subjects are moderate (M) users.

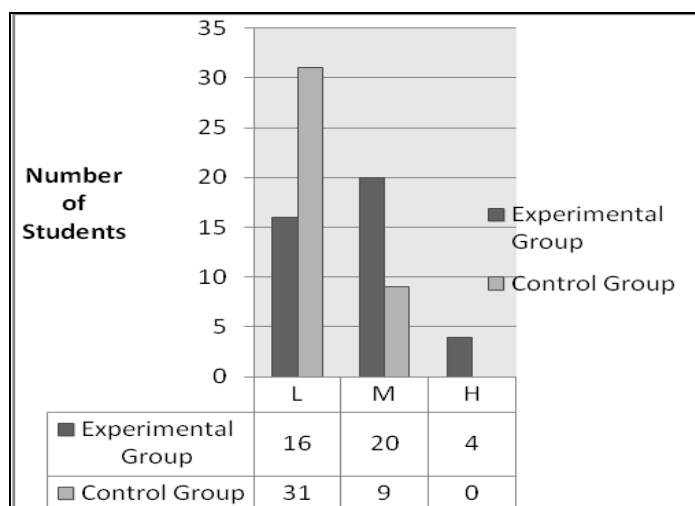


Figure 4.1: The Frequency of TAF Used in the Post AEW Test

The use of 0 to 4 TAF is considered low (L), the use of 5 to 8 TAF is considered as moderate (M) while the use of more than 8 TAF is considered as high (H). Based on Figure 4.1 it can be seen that the experimental group used more TAF compared to the control group. Sixteen subjects from the experimental group are considered low (L) users of the TAF, twenty subjects fall into the category of moderate user (M) while four subjects are considered as high users (H). On the other hand, thirty one subjects of the control group are low users (L) while nine subjects are moderate (M) users.

Table 4.2 shows the list of TAFs and how frequent each of them was used during the post AEW test by the experimental and control groups. It can be seen from Table 4.2 that the experimental group used more and a wider range of TAFs compared to the control group. The experimental groups used TAFs 209 times in total while the control group used them 106 times. It can be concluded that DIAF encourages the students to use TAFs in their writing.

Table 4.2: Frequency of TAF used in the Post AEW Test

		Experimental Group	Control Group	
	<b>TARGET ACADEMIC FORMULA (TAF)</b>	No. of Use	No. of Use	Total
1	in relation to	3	0	3
2	in response to	0	0	0
3	from the point of view	2	0	2
4	to distinguish between	0	0	0
5	the relationship between	1	0	1
6	in conjunction with	1	0	1
7	according to the	29	55	84
8	can be considered	2	0	2
9	a variety of	25	10	35
10	with regard to	1	0	1
11	is/ are / can be affected by	2	0	2
12	give rise to	5	0	5
13	as well as	6	0	6
14	more/less likely to	3	1	4
15	(there) are a number (of)	5	2	7
16	a large number of	5	8	13
17	there are (three a/few/many)	17	6	23
18	there are several	23	9	32
19	the number of	16	8	24
20	there have been/ there has been	10	4	14
21	there is no	5	2	7
22	appears to be/ does not appear to be	5	0	5
23	on the basis of	1	0	1
24	in terms of	9	0	9
25	in accordance with	0	0	0
26	due to the fact that	8	0	8
27	as a consequence	3	0	3
28	as a result of	11	1	12
29	due to	10	0	10
30	can be achieved	1	0	1
	<b>TOTAL</b>	<b>209</b>	<b>106</b>	<b>315</b>

#### 4.2 Addressing the Second Research Question

To answer the second research question, one way ANCOVA on the overall mean score of the post AEW test with the overall mean score of pre AEW test assigned as covariates was conducted to determine whether the difference between the mean scores of the experimental and the control groups in the post AEW test was significant if their previous knowledge is statistically controlled. Table 4.3 shows the results of one-way ANCOVA for the overall scores.

Table 4.3: Results of One-Way ANCOVA for the Overall Scores.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9284.582 <sup>a</sup>	2	4642.291	57.399	.000
Intercept	3445.063	1	3445.063	42.596	.000
PRE_AEW	6753.332	1	6753.332	83.500	.000
METHOD	2474.103	1	2474.103	30.591	.000
Total	296950.000	80			
Corrected Total	15512.188	79			
Dependent Variable: POSTAEW_SCORES					
a. R Squared = .599 (Adjusted R Squared = .588)					
b. Computed using alpha = .05					

Based on the table, the probability value obtained for 'PRE\_AEW' is 0.000 which is smaller than the predetermined alpha value of 0.05. This indicates that there is significant difference in the mean scores between the experimental and control groups when students' previous knowledge is statistically controlled. The probability value obtained for 'METHOD' (Experimental group: *Process writing class with DIAF*; Control group: *Process writing class without DIAF*) is also 0.000 which is smaller than the predetermined alpha value of 0.05. There also exists adequate evidence to show that there is significant difference in the mean scores between the experimental and control groups according to methods of instruction. The value of adjusted  $R^2$  is 0.588 which means that the independent variable

(METHOD) can account for 58.8 % of the variance in the dependent variable (post AEW test scores).

Next, to examine whether there is significant difference between the means of the experimental and the control groups for the three separate components of the post AEW test when the students' previous knowledge is statistically controlled, one way ANCOVA was conducted on the scores of the 'Content', 'Language' and 'Organization' components with the mean scores of each component in the pre AEW test used as their covariates.

Table 4.4: The results of one-way ANCOVA for AEW Test Components

Source	Type III Sum of Square	df	F	Sig	R <sup>2</sup>	Adjusted R <sup>2</sup>
Pre-Content	480.442	1	38.940	0.000	0.369	0.353
Method	92.579	1	7.504	0.008		
Pre-Language	1030.253	1	49.381	0.000	0.511	0.498
Method	681.176	1	32.650	0.000		
Pre-Organization	359.598	1	23.865	0.000	0.359	0.343
Method	237.227	1	15.744	0.000		

Based on Table 4.4, the probability value obtained for all the three components; 'Pre-Content', 'Pre-Language' and 'Pre-Organization' is 0.000. This suggests that there is significant difference in the mean scores between the experimental and control groups when students' previous knowledge of each component is statistically controlled. The probability values obtained for 'METHOD' for 'content', is 0.008 while the values for both 'language' and 'organization' are 0.000 respectively. These values are also smaller than the predetermined alpha value of 0.05 which means that there is significant difference in the mean score between the experimental and control groups according to methods of instruction. The adjusted R<sup>2</sup> values for 'content', 'language' and 'organization' components are 0.353, 0.498 and 0.343 respectively. These results indicate that the independent variable (DIAF) can account for 35.3% of the variance in the 'content' component,

49.8% of the variance in the 'language' component and 34.3% of the variance in the 'organization' component.

It can be inferred from the ANCOVA results that the experimental group performed significantly better than the control group in the overall scores as well as the scores for all the three writing components after going through the treatment (DIAF). DIAF can account for more of the variances in 'language' component compared to 'organization' and 'content' components.

Finally, to quantify the strength of the difference between the means of the experimental and the control groups, the effect size of DIAF on the overall academic writing performance was calculated.

Table 4.5: Effect Size of DIAF on the Academic Writing Performance

AEW TEST COMPONENTS	EFFECT SIZE	INTERPRETATION
Content	0.45	small
Language	0.98	large
Organization	0.86	large
Overall	0.87	large

Table 4.5 shows the effect size for the overall AEW test performance and the three different components of the AEW test. Based on Table 4.5, it can be seen that effect size of DIAF on academic writing 'content' component is small but the effect size is large for 'language' and 'organization' components.

## 5. CONCLUSIONS AND SUGGESTIONS

Drawing from the findings of the study, it can be concluded that DIAF has positive effects on the subjects' academic writing performance since the experimental group outperformed the control group in the post AEW test for all the three writing components ('content', 'language' and 'organization') and the effect size of DIAF on the 'content' component of the AEW test is small but the effect size for 'language' and 'organization' components is large. It can be concluded that the direct instruction of academic formulas (DIAF), using target formulas chosen from

pedagogic corpus, incorporated into an academic writing class is beneficial at enhancing the students' receptive knowledge of the target formulas as well as improving their academic writing skills. Although the time allocated for DIAF in the study was limited, the effect size of DIAF as a whole is large. This suggests the importance of formula instruction and the flexibility of DIAF that it could be successfully implemented in an academic writing class with an institutionally imposed syllabus. Thus, it is recommended that DIAF be implemented in academic writing class and the target formula selection should be based on pedagogic corpus.

### References

- [1] Ang, L.H., Abdul Rahim, H., Tan, K.H., & Salehuddin, K. (2011). *Collocations in Malaysian English learners' writing: A corpus-based error analysis*. *3L: Language, Linguistics, Literature*, 17(Special issue):31-44.
- [2] Al-Zahrani, M. S. (1998). *Knowledge of English lexical collocations among male Saudi college students majoring in English at a Saudi university* (Doctoral dissertation). ProQuest Dissertations and Theses database. (UMI No. 9825581).
- [3] Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263-286.
- [4] Biber, D., Conrad, S., & Cortes, V. (2004). Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- [5] Biber, D., (2006). *University Language: A corpus-based study of spoken and written registers*. Amsterdam, the Netherlands: John Benjamins.
- [6] Boers, F. & Lindstromberg, S. (2012). Experimental and Intervention Studies on Formulaic Sequences in a Second Language. *Annual Review of Applied Linguistics*, 32, 83-110.
- [7] Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, H. (2006). Formulaic sequences and perceived oral proficiency: putting a lexical approach to the test. *Language Teaching Research*, 10, 245-261.
- [8] Coffin, C., Curry, M. J., Goodman, S., Hewings, A., Lilis, T. M., & Swann, J. (2003). *Teaching Academic Writing : A toolkit for higher education*. London: Routledge.
- [9] Conrad, S. (2008). Writing Myth 6: Corpus-based research is too complicated to be useful for writing teachers. In J. Reid (Ed.), *Myths about teaching writing* (pp.115-139). Ann Arbor: University of Michigan Press.



- [10] Cortes, V. (2002). Lexical bundles in Freshman composition. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 131–145). Amsterdam: John Benjamins Publishing Company.
- [11] Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397–423.
- [12] Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17, 391–406.
- [13] Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of the academic prose. *Journal of Second Language Writing*, 16, 129–147.
- [14] Dai, Z. & Ding, Y. (2010). Effectiveness of text memorization EFL learning of Chinese students. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 71–87). New York, NY: Continuum.
- [15] Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18, 91–126.
- [16] Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20, 29–62.
- [17] Foster, P. (2001). Rules and routines: a consideration of their role in task-based language production of native and non-native speakers. In Bygate, M., Skehan, P. and Swain, M., (Eds.), *Researching pedagogic tasks: second language learning, teaching, and testing* (pp. 75–93). Harlow, UK: Longman
- [18] Granger, S. (2011) From phraseology to pedagogy: Challenges and prospects. In T. Herbst, P. Uhrig and S. Schüller (eds), *The Phraseological View of Language. A tribute to John Sinclair*. (pp. 123–146). Berlin & New York: Mouton de Gruyter.
- [19] Hinkel, E. (2004). *Teaching Academic ESL Writing: Practical Techniques in Vocabulary and Grammar*. Mahwah, New Jersey: Lawrence Erlbaum.
- [20] Hoey, M. (2005). *Lexical Priming: A new theory of words and language*. New York: Routledge.
- [21] Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19, 22–44.
- [22] Hyland, K. (2012). Bundles in Academic Discourse. *Annual Review of Applied Linguistics*, 32, 150–169.
- [23] Hyland, K. (2008). 'As can be seen: Lexical bundles and disciplinary variation.' *English for Specific Purposes*, 27, 4–21.
- [24] Jamian, L. S., Sidhu, G. K., & Muzafar, M. (2008). Assessing UiTM TESL Students' Knowledge of Vocabulary. *Asian Journal of University Education*, 4(2), 79–100.
- [25] Kelley, M. J. (2008). The impact of weblogs on the Affective states and academic writing of 12 undergraduates. Unpublished PhD. Thesis. The Faculty of the Curry School of Education University of Virginia, U.S.A.

- [26] Kennedy, C, & Thorpe, D. (2007). A corpus-based investigation of linguistic responses to an IELTS academic writing task. In L. Taylor & P. Falvey (Eds.), *Studies in Language Testing* (Vol. 19 IELTS Collected Papers, pp. 316-377).
- [27] Keshavarz, M. H., & Salimi, H. (2007). Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics*, 17, 81-92.
- [28] Kozlowski, A., & Seymour, L. (2003). The importance of Collocation in English Language Teaching. *Contact*, 29(3), 41-49.
- [29] Lewis, M. (2000). *Teaching Collocation. Further Development in the Lexical Approach*. Boston: Thomson Heinle.
- [30] Lewis, M. (2002) *Implementing the Lexical Approach*. Boston: Thomson Heinle.
- [31] Leibowitz, B., Goodman, K., Hannon, P. & Parkerson, A. (1997). The role of a writing centre in increasing access to academic discourse in a multilingual university. *Teaching in Higher Education*. 2 (1), 5-19.
- [32] Mathai, J., Jamian, L.S.& Nair, S. (2004) Assessing Malaysian University Students' English Vocabulary Knowledge. In Wilaiwan Khanittanan & Paul Sidwell, eds. *SEALSXIV: papers from the 14th meeting of the Southeast Asian Linguistics Society (2004), Volume 1*. Canberra, Pacific Linguistics, 2008, pp.219-237.
- [33] Meunier, F. (2012). Formulaic Language and Language Teaching. *Annual Review of Applied Linguistics*, 32, 111-129.
- [34] Micheal, G. R. D., Dhillon, P. K., Jaafar, H. H., Umadevi, S., Askandar, R., Ghafar, N., et al. (2010). *EAP Crossing Borders*. Petaling Jaya, Malaysia: Pearson Longman.
- [35] Mokhtar, A. A. (2010). Achieving Native-like English Lexical Knowledge: The Non-native Story. *Journal of Language Teaching and Research*, 1(4), 343-352.
- [36] Naderishahab, M. & Tahririan, M. H. (2013). Foreign Language Learners' Collocational Competence and its Relationship to their Productive Skills. *The Iranian EFL Journal*, 9(3), 24-45.
- [37] O'Farrell, C. (2005). The write approach: Integrating writing activities into your teaching. In G. O'Neill, S. Moore & B. McMullin (Eds.), *Emerging Issues in the Practice of University Learning and Teaching*. Retrieved from: <http://www.aishe.org/readings/2005-1>.
- [38] Ohlrogge, A. (2009). Formulaic expressions in intermediate EFL writing assessment. In R. Corrigan, A. Moravcsik, H. Ouali, & K. M. Wheatley (Eds.) *Formulaic Language; Vol 2. Acquisition, loss, psychological reality , and functional explanations* (pp/ 387-404). Amsterdam, the Netherlands: John Benjamins.
- [39] Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4), 519-549.
- [40] Roever, C. (2012). What learners get for free: learning of routine formulae in ESL and EFL environments. *ELT Journal*, 66, 10–21.

- [41] Schmitt, N. (Ed.). (2004). *Formulaic sequences*. Amsterdam, the Netherlands: John Benjamins.
- [42] Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*, 95, 26-43.
- [43] Schoonen, R., Van Gelderen, A., Stoel, R., Hulstijn, J., De Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary-school students. *Language Learning*, 6 (1), 31-79.
- [44] Simpson-Vlach, R., & Ellis, N. C. (2010). An Academic Formula List: New Methods in Phraseology Research. *Applied Linguistics*, 31(4), 487-512. Retrieved from <http://www.applij.oxfordjournals.org>
- [45] Sinclair, J. & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter & M. McCarthy. *Vocabulary and language teaching* (pp.140-158). London: Longman.
- [46] Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.
- [47] Singleton, D. M. (2000). *Language and the lexicon: An introduction*. London: Oxford University Press.
- [48] Willis, D. (2003). *Rules, Pattern and Words. Grammar and Lexis in English Language Teaching*. Cambridge: Cambridge University Press.
- [49] Wood, D. (2002). Formulaic language in acquisition and production: implications for teaching. *TESL Canada Journal*, 20(1), 1-15.
- Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and applications*. London: Continuum.
- [50] Wray, A. (2002). *Formulaic Language and Lexicon*. Cambridge, UK: Cambridge University Press.
- [51] Zhang, X. (1993). *English Collocations and Their Effect on the Writing of Native and Non-native College Freshmen*. Unpublished PhD Dissertation. Indiana university of Pennsylvania: Pennsylvania.