

ISSC 2016 : International Soft Science Conference

A Corpus-Based Analysis Of Lexical Verbs In MUET Argumentative Essays Of Pre-University Students In Malaysia

Anita Kanestion^{a*}, Manvender Kaur Sarjit Singh^a, SarimahShamsudin^b,
Hishamuddin Isam^a, Naginder Kaur^c, Gurmit Singh Pertap Singh^a

* Corresponding author: Anita Kanestion, anitakanestion@yahoo.com

^aAwang Had Salleh Graduate School, Universiti Utara Malaysia, Sintok, Malaysia, anitakanestion@yahoo.com

^bLanguage Academy, Universiti Teknologi Malaysia, Kuala Lumpur

^cUniversiti Teknologi Mara, Arau, Perlis, Malaysia

Abstract

<http://dx.doi.org/10.15405/epsbs.2016.08.56>

Argumentative writing has been regarded as an essential form of written discourse in Malaysia, specifically in Malaysian University English Test (MUET). However, limited research has been conducted on the MUET argumentative essay writing component for pre-university students. The purpose of this preliminary study is to investigate a compiled representative learner corpus to conduct an analysis of the linguistic features; lexical verbs used in Band 3 and Band 5 argumentative essays produced by students of a pre-university college in Malaysia. Subsequently, the lexical verbs utilized in the respective essays were compared. As a descriptive analysis and corpus-based study, the use of lexical verb forms was examined, reflecting the learners' knowledge of applying this linguistic constituent into their written essays. The method applied in the analysis is a computer-based learner corpus analysis. Results revealed that Band 5 essays contained more '-ing' participle (VVG) form, whereas Band 3 essays revealed the use of more past participle (VVN) form. Surprisingly, both groups avoided the base form of lexical verb (VVO). The pedagogical implications based on the results of the study were proposed and the ideas for future research were discussed.

© 2016 Published by Future Academy www.FutureAcademy.org.uk

Keywords: Corpus-Based Analysis; Lexical Verbs; Malaysian University English Test (MUET); Argumentative Essay; POS Tagging.

1. Introduction

One of the most significant current discussions in the area of Second Language Acquisition (SLA) is learner corpus. Corpus is denoted as a body of written or spoken text which is naturally produced and compiled by using computer and directly provides an empirical evidence of the language structures and



its use (Biber, Conrad and Reppen, 1998; Leech, 1998). According to Kennedy (1998) a learner corpus is defined as computerized textual database produced authentically by foreign language learners. It is believed that learner corpus can be an effective tool to investigate language learning processes and to improve the present language learning strategies since it serves as a basis for linguistic analysis and description (Zeba and Asim, 2014; Manvender and Sarimah, 2011). To access and evaluate the linguistic forms in learners' language, a corpus-based analysis is needed and the present study is aimed to analyse the use of lexical verb forms found in a compiled representative learner corpus of MUET argumentative essays written by pre-university students.

Hyland and Milton (1997) have argued that most non-native writers are unable to use lexical verbs in the argumentative essays as it involves critical lexical, tense and voice choices. This is due to the fact that lexical verbs are used to communicate the writers' stance towards statements and audience. In addition, Salazar and Verdager (2009) add that these verbs are employed by native speakers to express modality or in other words to express their voice and attitude towards issues raised. Expressing modality is one of the criteria needed to write argumentative essays as the readers are required to be convinced by the writers' ideas and justifications. In contrast, Guo (2006) has found out that non-native learners prefer to choose verbs for the writing purpose compared to native speakers who prefer to choose nouns. He has also revealed that the non-native learners overuse the base form of lexical verbs as it is influenced by the learners' first language and underuse other forms of lexical verbs, VVD, VVN, VVG and VVZ (refer to Table 2).

In another study, Abdullah and Noor (2013) have investigated the use of lexical verbs and verb-noun collocations between native speakers and non-native speakers in two learner corpora; LOCNESS (Louvain Corpus of Native English Essays) and WECMEL (Written English Corpus of Malay ESL Learners). The study reveals that both learners have utilized a common verb form of lexical verb and verb-noun collocation that is verb infinitive (VVI) in their argumentative essays. Interestingly, Malay learners are found to overuse past tense (VVD) whereas the native learners underuse past participle (VVN). Albeit, this study has highlighted the significant contribution of how lexical verbs and verb-noun collocation are used among Malay students at tertiary level who scored a distinction in English or known as competent writers, yet it only gives the opportunity to study how competent writers in Malaysian setting use the linguistic features.

According to recent statistic, Malaysian students who sit for Malaysian University English Test (hereafter MUET), only scored Band 2 and Band 3 (80%) in their writing component. Thus, there is a need to tackle this issue as MUET is compatible to IELTS and it is the nation's aspiration to produce competent user of English in order to achieve a developed country. It can be deduced that, most of the studies have attracted much attention to analyse linguistic features by employing contrastive analysis (CA) which involves comparing native speakers' learner corpora to non-native learner corpora and errors committed by the students at primary, secondary and tertiary level, yet the study on the language use by pre-university students, in Malaysia, who are the majority to be enrolled in public universities is still lacking. Pre-university students' writing skills are essential as it is one of the components tested in MUET and contributes as the second highest weighting (30%). In MUET writing, the students are expected to write one focused report writing and an argumentative essay. Besides, researches based on

MUET, especially writing are still insufficient. Therefore, the present study aims to provide an empirical evidence on the use of lexical verbs by competent writers (Band 5) and modest writers (Band 3) in a compiled learner corpus contributed by pre-university students.

2. Research Questions

This study is aimed to answer the following research questions:

1. What are the types of lexical verbs used in the compiled Band 5 and Band 3 argumentative essays by the pre-university students?
2. What are the frequencies of lexical verbs used in the compiled Band 5 and Band 3 argumentative essays by the pre-university students?

3. Methodology

3.1. Corpus

To determine the linguistic features (lexical verbs) of written texts of the MUET academic writing, a genre-specific corpus was compiled using total of twelve argumentative essays of pre-university students in a college in Malaysia. The MUET past year question was used for the purpose of the study and the essays were rated by two raters who had the experience of teaching and marking for more than 10 years. The twelve texts in the corpus were selected by using a purposive sampling method to fit the objectives of the study. All the texts, argumentative essays (coded as AE), were coded individually, according to Band 5 and Band 3 (six texts respectively) and the number of words (refer to Table 1).

3.2. Method

For the present analysis, the twelve texts were tagged for POS using the latest online trial version of the tagging software CLAWS (the Constituent Likelihood Automatic Word-tagging System) tagger (the CLAWS7 Tagset). The CLAWS tagger was developed by the UCREL at the University of Lancaster in the early of 1980s and has since been continuously improved. The tagger has consistently achieved 96-97% accuracy and having an error rate of only 1.5%. The POS tagging was supported by a list of coding for CLAWS7 tagger. Nevertheless, for the purpose of the study only lexical verbs are used and the list of coding is represented in Table 2.

For this study, the corpus was tagged horizontally as the nature of it is to give a general descriptive view of the linguistic constituents of the individual text as well as facilitate the frequency computation using concordance software. Therefore, the horizontally tagged version of the corpus was coded individually and loaded for the frequency analysis. A concordance tool (AntConc3.4.3w) was then used to compute the frequencies of the related linguistic constitutions (lexical verbs) which is available free of charge and can be downloaded from the internet.

Table 1. Argumentative Essay Coding and Number of Words

No.	Band 5 Essays	No. of words	No.	Band 3 Essays	No. of words
1	5AE1	886	1	3AE1	509
2	5AE2	763	2	3AE2	776
3	5AE3	553	3	3AE3	477
4	5AE4	779	4	3AE4	534
5	5AE5	721	5	3AE5	348
6	5AE6	570	6	3AE6	451

Table 2. List of Coding

Codes	CLAWS7 tagger
VV0	base form of lexical verb (e.g. give, work)
VVD	past tense of lexical verb (e.g. gave, worked)
VVG	-ing participle of lexical verb (e.g. giving, working)
VVN	past participle of lexical verb (e.g. given, worked)
VVZ	-s form of lexical verb (e.g. gives, works)

4. Results and Discussion

The concordance tool was used to compute the frequency of lexical verbs in the respective tagged corpus. The frequency analysis of lexical verbs were shown in Table 3 and Table 4 based on the bands. The total number of lexical verbs used were calculated manually as the software was unable to do so.

Table 3. Frequency of Lexical Verbs and Examples in Band 5 essays

Corpus		Codes					Total
		VV0	VVD	VVG	VVN	VVZ	
Band	5AE1	-	1	16	9	11	37
	5AE2	-	3	9	12	15	39
	5AE3	-	3	6	5	15	29
	5AE4	-	1	22	8	14	45
	5AE5	-	3	14	12	14	43
	5AE6	-	-	5	3	18	26
	Total	-	11	72	49	87	219
Examples		-	Saw, provided, shouldered, brought	Summing, sending, controlling	Provided, leveled, determined	Contributes, possesses, develops, realizes, lies	

Table 3 illustrates the frequency of lexical verbs and examples of use in the tagged corpus of Band 5 essays. The total number of lexical verbs used in the tagged essays is 219. As highlighted in the table, the most recurring part-of-speech is the 's' form of lexical verbs (VVZ) or also known as third person singular form with 87 occurrences altogether followed by -ing participle VVG. The abundant use VVZ seems to point to the native speaker's proficiency in using the present tense to talk about literary works, authors and characters (Guo, 2006). Furthermore, lexical verbs are utilized to express their attitude and in this context the writers choose VVZ and VVG so that their voice could be heard by the readers. This is probably due to the nature of the argumentative essays as it requires them to provide opinions based on the issues raised. Moreover, the analysis also shows that the students have avoided the use of the base form of lexical verb, which is coded as VV0 and utilized least number of past tense

of lexical verbs (VVD) with only 11 occurrences. Particularly, the tagged sub-corpus coded as 5AE6 has portrayed zero use of both VV0 and VVD form. Generally, past participle (VVN) is used by the students in their essays with 49 occurrences, however the four sub-corpus (5AE1, 5AE3, 5AE4 and 5AE6) have utilized the verb in a limited way as the occurrences are in between 3 and 10.

Table 4. Frequency of Lexical Verbs and Examples in Band 3 essays

Corpus	Codes					Total
	VV0	VVD	VVG	VVN	VVZ	
Band 3						
3AE1	-	-	3	3	8	14
3AE2	-	3	8	21	12	44
3AE3	-	-	-	-	4	4
3AE4	-	11	6	5	7	29
3AE5	-	-	2	3	-	5
3AE6	-	6	6	6	2	20
Total	-	20	25	38	33	116
Examples	-	Planned, respected, gained	Thinking, dealing, causing,	Defined, aided, measured	Contributes needs, admires	

Table 4 denotes the frequency of lexical verbs and examples of use in the tagged corpus of Band 3 essays. Compared to the number of lexical verbs exists in Band 5 essays, Band 3 essays only hit 116. This displays that Band 3 writers are not less likely to express their opinions to convince the readers. Evidently, the frequency analysis shows that the most frequently used part-of-speech is past participle (VVN) with 38 occurrences altogether and sub-corpus coded as 3AE2 recorded the highest frequency with 21 occurrences. This contradicts with Abdullah and Noor's (2013) findings that non-native learners underuse VVN in their essays compared to native learners. Nevertheless, 3AE3 has avoided from using this part-of-speech in the essay. Even though, there are some avoidance identified in the use of past tense (VVD) in the respective sub-corpus, 3AE1, 3AE3 and 3AE6, the findings reveal that Band 3 writers have overused the verb form compared to Band 5. Similar to Band 5 written essays, the use of the base form of lexical verb, which is coded as VV0, has been avoided. The findings of the research rejects Guo's (2006) findings that non native learners overuse the base form of lexical verbs. Nevertheless, the researcher agrees with his assumption that the native language of the learners could possibly influence the lexical verbs choice. In this context, the students are influenced by the national language, Bahasa Malaysia, as it is an inflected language. The analysis also reveals that the 'ing' form of the participle tense of lexical verbs (VVG) are used in a limited way as the frequency is recorded below 10 occurrences. Unexpectedly, the findings also has shown that 2 sub-corpus, 3AE3 and 3AE5 have only used the least number of lexical words, the first has used only VVZ in the essay with 4 occurrences whereas the latter has used two, VVD and VVG.

5. Conclusion

The major concern of this study was to analyse the linguistics constituent; the LEXICAL VERBS by using a corpus-based analysis. As stated earlier, 12 texts were compiled to form a representative corpus in relation to MUET band to achieve the objectives. Practically, both Band 5 and Band 3 essays used a variety of lexical verb forms in their argumentative essays. There were four types of lexical

verbs commonly used in the both Band 5 and Band 3 written essays, which were past tense (VVD), -'ing' form (VVG), past participle (VVN) and -s form (VVZ). Surprisingly, the most apparent conclusion was both Band 5 and Band 3 essay writers had avoided the base form (VVO) in their argumentative essays. Contrary to the avoidance of base form, Band 3 writers findings also showed limited response in employing -'ing' form (VVG) and -s form (VVZ) in the essays. Moreover, the findings of this study also implied that both Band 5 and Band 3 essays underused past tense (VVD) and rejected Abdullah and Noor's (2013) findings.

This study has a few significant pedagogical implications for the teachers and students both in theoretical and practical aspects. Albeit, the lexical verbs are analysed descriptively, the lexical verbs identified in Band 5 essays could be employed to improve the Band 3 essays. This can offer a better guideline for teachers, especially pre-university colleges in producing competent students who are being able to communicate effectively as argumentative writing is essential at tertiary level as well as work place. Furthermore, it can also provide a room for the teachers as well as the learners to understand their weaknesses and strengths of writing an argumentative essay and find ways to improvise their written product with an effective intervene.

The findings have shed some useful insights into the use of lexical verb forms by pre-university students in their argumentative essays. The analysis has highlighted the significant difference between Band 5 and Band 3 writers. Nevertheless, it had been focused only on a small sample of a learner corpus that had been compiled especially for this analysis. Therefore, the findings probably were not suitable for generalization on a larger corpus. Future researches can be conducted with more learners and with larger corpus as it would benefit the teachers and students. In addition, studies on distributional patterns of lexical verbs as well as error analysis of lexical verbs in the MUET essays in relation to bands can be carried out in the future.

Acknowledgements

The authors acknowledge the support of RACE Grant.

References

- Abdullah, S., & Noor, N. M. (2013). Contrastive Analysis of the Use of Lexical Verbs and Verb-noun Collocations in Two Learner Corpora: WECMEL vs. LOCNESS. (Published PhD Thesis). Kobe University.
- Biber, D., Conrad, S., & Reppen, R (1998). *Corpus Linguistics Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Guo, X. (2006). Verbs in the Written English of Chinese Learners: A Corpus-based Comparison between Non-native Speakers and Native Speakers. (Published PhD Thesis). University of Birmingham (2006).
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6, 183-205.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. London: Longman
- Leech, G. (1998). Learner corpora: What they are and what can be done with them. In S. Granger (Ed.), *Learner English on computer* (pp. xiv-xx). London: Addison Wesley Longman.
- Manvender, K., & Sarimah, S. (2011). Extracting noun forms: A lesson learnt. *International Journal of Language Studies* 5(4): 19-32.
- Salazar, D. and Verdaguer, I. (2009). Polysemous Verbs and Modality in Native and Non-Native Argumentative Writing: A Corpus-Based Study. *International Journal of English Studies*, 9(3), pp. 209-219.
- Zeba, I., & Asim, M. (2014). Genre Analysis of Argumentative Essays of Pakistani EFL Learners. *Journal of Education and Practice*, 5 (30), 95-100.