A CORPUS-BASED STUDY OF *BE*-COPULA IN NATIVE SPEAKER AND NON-NATIVE SPEAKER LEARNERS' ARGUMENTATIVE ESSAYS

[*1]Roslina Abdul Aziz & [2]Zuraidah Mohd Don

[1] Akademi Pengajian Bahasa, Universiti Teknologi MARA Cawangan Pahang,

Lintasan Semarak, 26400 Bandar Jengka, Pahang, Malaysia.

[2] Akademi Bahasa, Universiti Teknologi Malaysia,

81310 Johor Bahru, Johor, Malaysia.

*Corresponding author: leenaziz@uitm.edu.my

ABSTRACT

Background and Purpose: There is little research on the use of *BE*-copula in learner compositions in terms of how *BE* is used, and what functions it performs at specific stages of composition. This study aims to investigate the functions of *BE*-copula in argumentative essays written by non-native speaker (NS) and native speaker (NNS) learners and determine the similarities and differences in the way the learners use *BE*-copula in their essays.

Methodology: This descriptive study employed a corpus-based methodology with the corpus data obtained from the Malaysian Corpus of Learner English (MACLE) and Louvain Corpus of Native English Essays (LOCNESS). Only data from L1-Malay learners, which were extracted from MACLE, and which form a sub-corpus of L1-Malay learner argumentative essays, were analyzed in the study. Adopting Hyland's three-stage structure of an argumentative essay as the framework for analysis, data were analyzed for the distribution of *BE*-copula and the functions it performed in the learner essays.

Findings: Results from the quantitative analysis revealed a similar pattern in the use of *BE*-copula between NNS and NS learners, with both groups showing an inclination towards the present forms (*is, are*) and *BE*-copula construction. A *t*-test analysis revealed a more significant use of *BE*-copula in terms of both forms and functions by the NS learners. The qualitative analysis revealed that even though the

NNS learners exhibited almost similar composing elements as the NS counterparts, their texts were stylistically simple, more constrained, less fluent, and effective due to limited syntactic variety.

Contributions: The empirical findings from the in depth quantitative and qualitative analyses have enabled more insightful conclusions to be drawn about the NNS learners' use of *BE*-copula in their writing. The present study has direct pedagogical implications for the teaching of academic writing in the ESL context.

Keywords: Argumentative essay, academic writing, *be*-copula, corpus-based, learner corpus research.

Cite as: Abdul Aziz, R., & Mohd Don, Z. (2022). A corpus-based study of be-copula in native speaker and non-native speaker learners' argumentative essays. *Journal of Nusantara Studies, 7*(2), 21-43. http://dx.doi.org/10.24200/jonus.vol7iss2pp21-43

1.0 INTRODUCTION

There is little research on the use of *BE*-copula in learner compositions in terms of how *BE* is used, and what functions it performs at specific stages of composition (Staples & Reppen, 2016). Previous studies on learner writings have focused mainly on lexical usage (e.g. Aijmer, 2002), syntactic development (e.g. Biber & Reppen, 1998; Hinkel, 2002, 2003), syntactic complexity (Friginal, Li, & Weigle, 2014; Staples & Reppen, 2016), grammatical features (Imani & Habil, 2012), and style and discoursal differences (Bloch, 2010; Flowerdew, 2003), with most focusing on the use of stative *BE* as part of the analysis for the syntactic complexities and grammatical features of learner compositions.

Clauses with *BE* as the main verb are regarded as syntactically simple (Hinkel, 2002, 2003), and are often associated with structures used in conversation. Most often, the constructions involve the use of copula *BE* in simple propositions, consisting mostly of a subject and an adjectival predicate (e.g. *The money making **is** important of course*). More recent studies on learner writing (Friginal et al., 2014; Staples & Reppen, 2016) acknowledge the importance of *BE* clauses in academic writing, for instance in expressing stance (Staples & Reppen, 2016). Staples and Reppen (2016) in their analysis of the choice of verbs used by L1 learner writers in expressing stance in long arguments found that they notably used more *BE* controlling complement clauses. This has enabled the L1 writers to express stance less overtly, which is considered more effective than using conversational verbs such as *think*. The finding led the researchers to suggest adding the ability to use *BE* controlling complement clauses to

the framework of more advanced writers, which was proposed by Biber, Gray, and Poonpon (2011).

This study is part of a larger investigation into the use of *BE* by a group of non-native speaker (NNS) learners of English. It investigates the use of copula *BE* by the NNS learners at different stages of argumentative essays, and compares the patterns of use to that of a group of native-speaker (NS) learners of comparable age and educational level. The aim is to ascertain the functions of *BE*-copula in argumentative essays and determine the similarities and differences in the way the NNS and NS learners use *BE*-copula in their essays. Drawing from Hyland's (1990) three-stage structure of argumentative essays, which are Thesis, Argument and Conclusion, the study addresses the following research questions:

1. What are the distributional patterns of copula *BE* in the argumentative essays written by NNS and NS learners?
2. How is copula *BE* used at different stages of the argumentative essays?

2.0 LITERATURE REVIEW

2.1 *BE* in L2 Writing

Studies on the use of *BE* among older L2 learners tend to focus primarily on the position of *BE* in interlanguage grammar and the role of L1 interference in the acquisition of *BE* (e.g., Ang, Tan, & Lye, 2020; Dissanayake & Dissanayake, 2019; Smirkou & Smirkou, 2018). Ang et al. (2020), who utilized corpus-based techniques to identify and analyze errors associated with subject verb agreement (SVA) and copula *BE* in Malaysian L2 learners' writing reported various types of errors in the writings including mis-selection, omission, blends, and overinclusion of *BE*. Smirkou and Smirkou (2018) in their study found that copula omission elicited in the speech among Moroccan learners of English occurred more frequently in situations where Arabic does not require the use of the copula, namely the present tense, progressive aspect, and passivization. The study reported the main source of omission is the negative transfer from Moroccan Arabic into English. Dissanayake and Dissanayake (2019) also reported syntactic errors involving auxiliary and copula *BE*, where *BE* was inserted in non-obligatory context in the writings of undergraduates in Sri Lanka, whose mother tongue is Sinhala. The researchers attributed the errors to overgeneralization of the target language rules.

These studies have undoubtedly revealed various aspects of L2 acquisition process and development, especially the acquisition of verbal morphology and factors affecting acquisition. Very little, however, is known about how *BE* is actually used in learner compositions, and what

functions it performs at specific stages of the compositions. According to Hinkel (2003), heavy reliance on *BE*-copula constructions among NNS learners renders their writings simplistic and conversational, and so lacking elegance, as clauses with *BE* as the main verb are simpler than those containing verbs with higher semantic and lexical content (Biber, Johansson, Leech, Conrad, & Finegan, 1999). The use of *BE* (or the lack of it) has been studied in relation to L2 proficiency. Imani and Habil (2012) found that the *BE* structure was commonly used in the writing of NNS postgraduate students irrespective of their level of proficiency, not because of its simple constructions, but rather owing to the lack of awareness of the preferred structures in academic writing.

Friginal et al. (2014) studied the relationship between linguistic distributions and the quality of writing in highly rated NS and NNS learner compositions, and found that even though these compositions would not generally have high frequencies of *BE*, the verb is considered one of their very important linguistic features. A high positive score for stative *BE* in two of the six clusters, used to categorize highly rated learner essays, was also reported by Jarvis (2007). The clusters were also characterized by high positive mean scores for text length, demonstratives, nominalizations and adverbial subordination (Cluster 3), and high positive mean scores for text length, diversity of vocabulary, and passives (Cluster 5), all of which are essential features of academic writing.

Clauses with *BE,* even though considered simpler than that with higher lexical and semantic verbs, form an indispensable and important part of academic prose. Staples and Reppen (2016) in their analysis of the choice of verbs used by learner writers in expressing stance in long arguments found that L1 English learners notably used more *BE* controlling complement clauses. This has allowed the L1 writers to express stance less overtly, which is considered more effective than using conversational verbs such as *think*. The finding led the researchers to suggest the ability to use *BE* controlling complement clauses to be added to the framework of more advanced writers, which was drawn by Biber et al. (2011) and reported in the Biber, Gray, and Staples (2016).

2.2 Hyland's (1990) Argumentative Essay Analysis Framework

Based on a detailed study of top 10% essay scripts (N=68) written for the high school matriculation in English, Hyland (1990) proposed a model for the elements of the argumentative essay, which is characterized by three stages, which are Thesis, Argument and Conclusion. Each stage is expressed by structural moves, some of which are obligatory, while others optional. The thesis stage has five moves, which are Gambit, Information, Proposition,

Evaluation and Marker, but only Proposition is obligatory. On the other hand, the Argument stage is made up of four moves, which are Marker, Restatement, Claim and Support, and only Restatement is optional. The conclusion includes four moves, which are Marker, Consolidation, Affirmation and Close, with only Consolidation obligatory.

The framework has since been adopted to analyze the rhetorical structures of argumentative essays of learners of different L1 backgrounds. Among the studies include Imtiaz and Mahmood (2014) and Kanestion, Singh, Kaur, and Shamsudin (2017). Imtiaz and Mahmood (2014) explored the move structures of 33 argumentative essays by Pakistani college and university students extracted from ICLE (The International Corpus of Learner English), while Kanestion et al. (2017) investigated the linguistic features used to signal the moves in 24 argumentative essays by the pre-university students in a pre-university institution in Malaysia. Besides application to the entire essay, the framework has also been applied to investigate the specific parts of learner essays. Liu's (2015) comparative study that focused on the moves and wrap-up sentences in the concluding paragraphs of English expository essays produced by first year, second year and third year Chinese students, and Kanestion and Singh's (2019) examination of the introductory paragraphs of 14 argumentative essays of NNS learners in Malaysia are examples of studies taking this route. Hyland's (1990) model has mainly been applied to explore the rhetorical structures of essays, and to the knowledge of the researchers, the present study will be first to adopt the framework in investigating the realization of a specific linguistic feature (i.e., *BE*-copula) in the move structures of learner argumentative writings. Hyland's (1990) analytical framework is deemed most suitable for the purpose of this study as it contains very detailed move structures, which enable the researcher to investigate the functions of copula *BE* in depth.

3.0 RESEARCH DESIGN

3.1 Corpus Data

The data for this study consist of two learner corpora. The first is a subset of the Malaysian Corpus of Learner English (MACLE) (Knowles et al., 2006), which is a mono-generic corpus consisting of argumentative essays written by second to fourth year undergraduates at the University of Malaya between 2004 and 2005. It amounts to 800,000 words at the time the study was conducted. It consists of essays written by learners from three major ethnic groups in Malaysia: Malay, Chinese and Indian. This study is a part of a larger investigation into the use of *BE* by L1-Malay learners, and for that reason, only scripts written by L1-Malay learners were analyzed for this study. The subset consists of 366 scripts and approximately 198,262

words. Since MACLE is modeled after the International Corpus of Learner English (ICLE) (Granger, Dagneaux, & Meunier, 2002), the topics for the essays were adopted from ICLE. Table 1 below summarizes the topics together with the number of essays written for each topic.

Table 1: The distribution of essays according to topics in the NNS corpus

	Essay topic	Number
1	Crime does not pay.	40
2	The prison system is outdated. No civilized society should punish its criminals: It should rehabilitate them.	22
3	Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.	88
4	The role of censorship in society.	30
5	All armies should consist entirely of professional soldiers: There is no value in a system of military service.	1
6	The Gulf War has shown us that it is still a great thing to fight for one's country.	6
7	Feminists have done more harm to the cause of women than good.	7
8	In the words of the old song: 'Money is the root of all evil'.	115
9	Some people say that in our modern world, dominated by science, technology and industrialization, there is no longer a place for dreaming and imagination. What is your opinion?	57
	Total	366

The NS learner corpus is a subset of the Louvain Corpus of Native English Essays (LOCNESS), which is a native speaker component of the International Corpus of Learner English (ICLE) (Granger, 1996). ICLE consists of 436 argumentative and literary essays written by native speakers of British English (from the UK) and American English (from the United States) for their assignments and examinations. In ensuring the homogeneity of the NS data and the comparability of the NS and NNS data, only argumentative essays from the American component were used for this study, which means that the literary essays were excluded from the analysis. The exclusion has left the sub-corpus with approximately 149,573 words as shown in Table 2.

The American learner sub-corpus is considered more comparable to the L1-Malay sub-corpus in terms of size, level and length of learners' education and the writing tasks. The use of the American sub-corpus also eliminated the potential effects of English variety (American vs. British variety) in the NS data. Both the American learner sub-corpus and MACLE consist of timed and untimed essay scripts on argumentative topics concerned with contemporary

social issues. The general equivalence in genre, text types and topics in both corpora suggests considerable similarities in the language requirements of the tasks, so that they are highly comparable in the use of *BE*. Table 2 below summarizes the composition of argumentative essays in the American learner sub-corpus.

Table 2: The distribution of essays according to universities in the American learner sub-corpus

	No of essays	No of words
1. Marquette University	46	54,285
2. Indiana University at Indianapolis	28	13,454
3. Presbyterian College, South Carolina	6	12,447
4. University of South Carolina	53	52,885
5. University of Michigan	43	16,502
Total	**176**	**149,573**

3.2 Data Annotation and Analysis

CLAWS (Constituent Likelihood Automatic Word-tagging System) developed by UCREL at the University of Lancaster was used to tag the NS learner sub-corpus. CLAWS is a hybrid part-of-speech tagger using both the probabilistic and rule-based elements. We did not, however, use CLAWS to tag MACLE because it was not able to handle the errors in the ESL learner corpus. Findings from the study conducted by Roslina and Zuraidah (2019) showed that CLAWS achieved accuracy of 93.6% for NNS texts, which is less than for the BNC (96-97%). For this reason, the L1-Malay learner sub-corpus was manually annotated using specially developed tagsets (Roslina, 2018).

WordSmith Tools Version 5 was employed for data analysis as it enabled frequency analysis of the *BE* to be administered automatically using the Concord and WordList tools. The software also allows for analysis outputs to be saved in various formats (txt., XML text, excel spreadsheet or print to file) to be analyzed later. WordSmith Tools Version 5 consisted of three major tools; Concord, WordList, and Keywords, was used to analyze both the NS and NNS learner sub-corpora. The major part of the data analysis made use of the Concord Tool, which identified the occurrences of *BE* in the learner sub-corpora. The overall number of occurrences was counted according to the forms and functions of *BE*. The Concord Tool was also utilized to generate the concordances of copula *BE* used at the different stages of the essays, namely Thesis, Argument and Conclusion. The contextual analysis of *BE* was administered by clicking on the *BE* node selected, which enabled the verb to be examined in the specific context of

occurrence. Only 10% of texts from each sub-corpus were analyzed for this purpose and the analysis for the NNS learner texts involved only the well-formed *BE* clauses.

Meanwhile, the Statistical Package for Social Sciences (SPSS) version 24.0 was employed to administer descriptive statistics on the datasets to obtain the mean and standard deviation for each finite *BE* form, and a *t*-test was administered to discern the possible difference between the means of the two groups in the use of *BE*-copula.

3.3 Analytical Frameworks

The scope of the study involves finite *BE* (i.e., *am*, *is, are, was, were*) functioning as a main verb or as a copula. Copula *BE* is mainly used to link a subject NP to its predicate, which can take the form of a phrase (noun phrase or adjective phrase), or a clause (*that* clause or infinitive *to* clause). Copula *BE* in this study was analyzed using Hyland's (1990) three-stage structure for argumentative essays, namely Thesis, Argument and Conclusion. These stages represent the organizing principles of argumentative essays (Hyland, 1990). Each structure is expressed in terms of moves as summarized in Table 3.

Table 3: Elements of structure of the Argumentative Essay

STAGE	MOVE
. *Thesis.* Introduces the proposition to be argued	**(Gambit)** Attention Grabber – controversial statement or dramatic illustration. **(Information)** Presents background material for topic contextualization. **Proposition** Furnishes a specific statement of position. **(Evaluation)** Positive gloss- brief support of proposition **(Marker)** Introduces and/or identifies a list.
2. *Argument* Discusses grounds for thesis (four move argument can be repeated indefinitely)	**Marker** Signals the introduction of a claim related it to the text **(Restatement)** Rephrasing or repetition of proposition **Claim** States reason for acceptance of the proposition **Support** States the grounds which underpin the claim.
3. *Conclusion*	**(Marker)** Signals conclusion boundary. **Consolidation** Presents the significance of the argument stage to the proposition. **(Affirmation)** Restates propositions. **(Close)** Widens context or perspective of proposition.

4.0 ANALYSIS AND DISCUSSION

4.1 Distribution of *BE* in NS and NNS Learner Essays

The SPPS version 24.0 was used to calculate the mean and standard deviation for each finite *BE* form. The overall distribution is shown in Table 4. The analysis revealed remarkable similarities in the patterns of use of all the finite *BE* forms, with both corpora employing more present forms than the past forms. In total, 6952 tokens and 5248 tokens of contracted and

uncontracted finite *BE* were used by the NNS and NS learners respectively. Both the NS and NNS learners used mostly *is* and *are* in their essays. As shown in Table 4, *is* and *are* have higher means than *was* and *were* in both corpora. The NNS learner corpus recorded the mean of 10.78 and 6.17 for *is* and *are* respectively compared to only 1.09 and .62 for *was* and *were.* On the other hand, the NS learner corpus means were 15.14 and 7.82 for *is* and *are* respectively, compared to only 3.52 and 2.13 for *was* and *were*.

Table 4: Distribution of *BE* in the NS and NNS learner data

	Mean		Std. Deviation	
	NNS	**NS**	**NNS**	**NS**
Am	0.11	0.17	.56	.62
Is	10.78	15.14	7.35	11.67
Are	6.17	7.82	4.95	6.90
Was	1.09	3.52	2.13	4.56
Were	0.62	2.13	1.29	3.31

According to Swales (1990), the present tense is conventionalized in academic writing. It is used in references to current knowledge, and when used in this context, the reported information is close to the writer's beliefs and opinions. The higher distribution of present *BE* forms is also determined by the writing genre. Past tenses are mostly associated with narratives, case studies or descriptions of historical events, while the present tense is associated with formal academic prose dealing with generalizations, observations, or descriptions (e.g., Paltridge, 2001; Swales & Feak, 2012). The essays compiled for MACLE and LOCNESS are argumentative, and they deal with contemporary issues requiring the learners to argue for or against an issue using here-and-now arguments or what McCarthy (1994) refer to as 'now-relevance' (p. 102). Hinkel (2004) added that writers are able to construct more effective academic text when they use the simple present tense.

In terms of the mean use of *BE*, Table 4 reveals that NS learners surprisingly recorded higher mean scores for almost all forms, especially the form *is*, which recorded the mean of 15.14 in the NS learner data, compared to 10.78 in the NNS learner data. The NS learners also scored slightly higher means for other forms (*are*, *was, were*). This is contrary to past studies (Hinkel, 2002, 2003) that recorded higher usage of *BE* among the NNS learners.

A *t*-test was also administered to gauge the difference in the use of *BE* between NS and NNS learners. Table 5 below summarizes the findings.

Table 5: Independent sample *t*-test of *BE* forms

| | | | t-test for Equality of Means | | |
| | | | df | Sig. (2-tailed) | 95% Confidence Interval of the Difference | |
					Lower	Upper
am	Equal variances assumed	1.23	538	0.219	0.17	0.039
	Equal variances not assumed	1.19	310.83	0.236	0.18	0.043
is	Equal variances assumed	5.27	538	0	5.98	2.73
	Equal variances not assumed	4.51	240.32	0	6.25	2.45
are	Equal variances assumed	3.16	538	0.002	2.66	0.62
	Equal variances not assumed	2.81	260.37	0.005	2.79	0.49
was	Equal variances assumed	8.46	538	0	2.99	1.87
	Equal variances not assumed	6.7	209.6	0	3.15	1.72
were	Equal variances assumed	7.6	538	0	1.89	1.12
	Equal variances not assumed	5.81	198.38	0	2.02	0.99

As shown in Table 5, there are significant differences in the use of *is, are, was, were* (p-value < 0.05). The results indicate that the NS learners use *BE* more in their writing than the NNS learners. This finding provides further support for Imani and Habil's (2012) finding, that the use of *BE* is not determined by learner proficiency. In accordance with this finding, the NS learners tend to use higher means of *BE* in their writing than the NNS learners, suggesting very clearly that proficiency has very little influence on the decision learners made to use *BE* clauses. Instead, the use could be determined by the need to fulfill the requirements of academic writing conventions, which will be dealt with further in the following sub-section.

4.2 Distribution of *BE* According to Functions

In order to determine what *BE* is mainly used for in the sub-corpora, we analyzed the three major functions it performs: copula, auxiliary passive and auxiliary progressive. The analysis involves only the uncontracted finite *BE* forms, namely *is, are, was* and *were*. Table 6 below summarizes the findings.

Table 6: Distribution of *BE* according to functions

	NS		NNS	
	Token	**%**	**Token**	**%**
Copula	3893	54.3	5252	63.6
Aux-Passive	799	11.2	1023	12.4
Aux-Progressive	364	5.1	497	6

The NNS and the NS learners show an almost similar trend in the overall use of *BE*. Across the sub-corpora, copula *BE* constructions occur the most, followed by auxiliary *BE* in the passive voice and auxiliary *BE* in the progressive aspect. The NNS learners have a slightly higher percentage of copula *BE* constructions (63.6%) than the NS learners (54.3%). There is generally a heavier reliance on copula *BE* constructions, compared to auxiliary *BE* in both the NNS and NS learner essays.

Milton (2001) reported similar findings. He found that auxiliary *BE* is relatively rare in the interlanguage grammar of the Hong Kong ESL learners and claimed that this was the result of learners' avoidance of the passive voice and aspectual verb forms (Milton, 2001). An avoidance strategy could not be used to explain the lower frequency of auxiliary *BE* in both the NNS and NS corpora in this study, and the register and the writing genre are more likely explanations. The progressive aspect is mainly used to refer to ongoing events or a temporary state (Quirk, Greenbaum, Leech, & Svartvik, 1985), and is more common in speech (Biber et al., 1999). There may be less opportunity to use this aspect in argumentative essays than in other genres such as narratives.

Table 7 below summarizes the findings of an independent *t*-test on the data to analyze significant differences between NNS and NS learners in the use of each function.

Table **7**: Independent sample *t*-test of *BE* functions

			95% Confidence Interval of the Difference	
	Mean Difference	**T test statistics**	**Lower**	**Upper**
Copula	8.01	6.90*	5.72	10.30
AuxProgressive	1.05	4.85*	0.62	1.48
AuxPassive	1.70	4.77*	1.00	2.40

*p-value < 0.05

As shown in Table 7, all three functions recorded a p-value < 0.05, indicating significant differences in the use of the functions in the learner data. The finding indicates that NS learners use significantly higher means of all the *BE* functions.

4.3 How *BE* is Used at Different Stages of Argumentative Essays

In finding out how *BE* is used at different stages of the argumentative essay, the verb was analyzed in its context. The findings are presented and discussed according to the three stages of argumentative essays - thesis, argument, conclusion.

The thesis stage consists of potentially 5 moves - gambit, information, proposition, evaluation and marker. This stage introduces the topic and advances the writer's proposition. *BE*-copula constructions are mainly used by the NNS and NS learners in this stage to provide the background information of the topic and to move forward their propositions.

Information moves are considered almost universal features of argumentative writing. They often include a restricted class of illocutions namely, definition, classification, descriptions, critiques or 'straw man' arguments (Hyland, 1990, p. 70). The information moves that are expressed using *BE*-copula construction in the NNS and NS learner essays are restricted to only definition and classification. According to Biber et al. (1999), one of the main functions of *BE*-copula in academic writing is to describe and classify a subject NP. It is common for a topic to be introduced using the definition of its key concept(s) and usually this is done using a simple *NP + BE + NP* structure. As can be seen in the following extracts, terms are defined by both the NS and NNS learners in an almost similar fashion. The simple *BE* constructions are often expanded with attendant elements (e.g., relative clause or infinitive clause) used to complement the subject as in the samples:

1. **Steroids** *are* chemical cousins of the hormone testosterone. They promote muscle growth and increase aggression, spurring athletes to more - strenuous workouts. (NS)
2. **Censor** *is* an act to remove parts which are considered offensive, politically unacceptable or even a threat to security depending on the values, culture or beliefs of a certain society. (NNS)

It is also common for the information moves to consist of a description of a term/concept/idea/organization. These descriptions provide the background information on the topic, and have become an almost conventionalized feature of the thesis stage. Both the NNS

and NS learner essays contain this feature in the thesis stage, and the construction is normally realized in *NP+BE+NP* structure as shown below:

3. **The Exxon corporation** *is* one of the largest corporations in the world and owns many smaller corporations and businesses as well. (NS)
4. **Robbery, rape, criminal breach of trust, house breaking, and murder** *are* the crimes that criminals specialize in order to get what they want. (NNS)

Another common usage of *BE*-copula in the information move is to classify. In the NNS learner essays, existential *there* clauses are often used for classifying. Biber (1988) comments that existential *there* is predominantly more common conversational genres (Biber et al., 1999; Brazil, 1995; Quirk et al., 1985). A vast majority occur with *BE* as the main verb, which characterizes it as syntactically and lexically simple (Biber, 1988). According to Jiang and Hyland (2020), existential *there* is very common in academic writing. It is not restricted to only classifying but performs useful persuasive purposes across sections of the research article and is an important rhetorical tool, "allowing writers to build interaction with readers and claim credit for their ideas" (p. 28). As shown in Extract (5), existential *there* constructions produced by the NNS learner do possess the qualities characterized by Biber (1998) as being syntactically and lexically simple and were used mainly to classify. However, the NS learners exhibited more advanced use of the existential *there*, whereby it functions as a rhetorical tool in building interaction with the reader as exemplified by Extract 6 below:

5. *There are* different types of censorship; among them are internet, music, television, film, movie and radio censorship. Although *there are* a wide variety of censorships, they are carried out for almost the same reasons and objectives. (NNS)
6. *There are* two sides to this issue. **On one side** *are* chemists armed with state-of-the-art equipment for detecting the use of such substances. **On the other** *are* many athletes determined to use any means available to enhance their performance and get away with it. (NS)

Perhaps, the most prominent use of *BE*-copula constructions in the essays written by NS learners is in the realization of the proposition move. *It*-clefts tend to be used in a proposition statement, which is commonly realized in *IT + BE + ADJ* pattern such as *it is obvious* or *it is clear* as shown in Extracts (7) below. Most often, the proposition statements precede two

opposing arguments. The *BE*-copula clause justifies the writer's stance, and at the same time adjective intensification is used to influence the reader into accepting the reasonableness of the proposition. This finding is in line with Matesic and Memisevic (2016), who found that the use of adjective intensifications in conveying stance was relatively more common in papers written by native speaker authors than Croatian authors. They reported that English linguistic texts contain 3 times more evaluative adjectives per 1,000 words than the ones in Croatian.

Generally, *it*-cleft is rarely used by the NNS learners in the proposition move, but the construction can be found in the other parts of the learner essays, for instance, in the support move (refer to Extract 16). This finding suggests that NNS learners are perhaps unaware and unfamiliar of the use of *it*-cleft in expressing propositions. As suggested by Imani and Habil (2012), L2 learners' lack of awareness of the convention of academic writing might influence the grammatical features or syntactic structures they decide to use (or not use) in their writing. After considering these articles that cover both sides of the issue, **it is obvious** that prayer does not belong in the public-school classroom, as the articles that oppose prayer in public schools refute and weaken considerably the arguments for the reintroduction of prayer in public schools as a way to cure modern social ills. (NS)

BE-copula is also used in propositions that involve the writer making a declaration of opinion by using attributive adjectives as exemplified in (8) to (9) below. Other than to express specific evaluation (Biber et al., 1999), Hyland (1998) explained that in academic texts, predicative adjectives function as lexical and strategic hedges. They can be used to either express a predicated state of affairs or a degree of variation relative to an expected norm. The declaration aims to give maximum effects to the proposition, with minimal regard for the opposing view. This technique is often used to realize the claim move in the argument stage (Hyland, 1990), but in this study, it is also used by the NNS and NS learners to convey their stance.

7. Surrogate motherhood ***is not* an ethical way** for infertile couples to have a child because of the emotional strains and ties present among the surrogate mother and the other two parents. (NS)
8. It (Prison system) ***is a waste** of taxpayers' money and an outdated, cruel, and barbaric way that should be abolished long ago. (NNS)

The argument stage consists of a four-move cycle, which can be repeated indefinitely in a specific order. The moves include marker, restatement of proposition, claim, and support. The use of *BE* in the argument stage centers mainly on the claim and support moves.

In both the NS and NNS learner essays, *BE*-copula can be used to state the claims. A claim constructed with *BE* often follows a specific structure; it is commonly framed by a marker that is phrased together with the subject NP, which is then linked to the subject predicate using *BE*. The use of *BE* controlling complement as exemplified in Extract (10) and Extract (11) are also common in the long arguments of highly rated NS learner essays analyzed by Staples and Reppen (2016). They argue that the structure is an effective method to establish a writer's claim, as it enables writers to achieve subtleness in their claim statements.

1. **The final problem** which arises *is* the modern tendency of moral relativism, due in part to individualism. (NS)
2. **One additional argument** *is* that if the nation aspires to be a global player and a centre of excellence that attracts students from the four corners of the world, our universities must have great teaching staff. (NNS)

The support move is an indispensable part of the claim, is generally tied to the claim as a pair, and it reinforces the claim by providing the relevant evidence and reasons to support it (Hyland, 1990). The most common use of *BE*-copula in the support move is to state facts or what the writer generalizes as the 'truth'. In this context, *BE*-copula indicates certainty, which is required in statements of fact. Most often the statements are not even supported by factual details, as they are perceived by the society as the truth, for instance, the idea that a woman is physically not strong enough to fight a war (Extract 12).

3. Women just can't fight in a war very well. They *are not* **strong** enough and if captured could be raped. (NS) In Malaysia, the provision under the Article 8(1) of the Federal Constitutions state that all person men and women *are* **equal** before the law and entitle to the equal protection of the law. (NNS)

Another common usage of *BE*-copula is to provide an explanation for an argument which is realized using the *This + BE* construction. The demonstrative pronoun *this* in the construction is used to draw attention to a referent, or can be used by writers to reinforce text cohesion (Hinkel, 2001; Swales, 2005; Wulff, Römer, & Swales, 2012). It can be used as a determiner

of a noun phrase (*This problem is* X), or can also be used as a free-standing pronoun (*This is because* X) (Wulff et al., 2012). According to Gray (2010), *BE* in academic writing is often followed by either an adjective (*BE+ADJ*), a word indicating an explanation (*BE+because*) or a noun or noun phrase (*BE+NP*). Corpus-based studies by Swales (2005) and Wulff et al. (2012) reported about 20-30% use of pronominal *this/these* by expert writers, and Gray (2010) reported that 40% of pronominal instances are followed by copula *BE*, indicating that pronominal *this/these* followed by *BE* is a common occurrence in academic writing. Galti (2016) also reported the common use of demonstrative pronouns in research articles by Nigerian authors, which according to the researchers, was used to provide clear and unambiguous presentations of facts and ideas in the writings.

The *This + BE* pattern is mostly used by the NNS learners as a pronoun, and often the construction provides an explanation for the cause(s) of observed phenomena, where *BE* is preceded by either a *to*-clause or a conjunction that signals a cause-and-effect relationship like *because, due to*. When compared to their NS counterparts, the NNS learners tend to have a very restricted use of *This + BE* constructions, and they also tend to be syntactically simpler. In the NS learner essays, other than providing explanation (*BE+because*), pronominal *This + BE* pattern is also used to reinforce an argument already presented, which is usually achieved through the *BE + NP* pattern or to emphasise the importance of a previously mentioned argument or to indicate evaluation/stance, which is realized using the *BE + ADJ* pattern. This finding is in line with that of Gray (2010), who also reported more varied use of *this* cluster in the NS expert writers' research articles, in which about 60% of *This + BE* constructions are followed by a noun or noun phrase, 26% by an attributive adjective, and only about 13% by *because*.

The NS learners' *This + BE* constructions are also syntactically more complex. As exemplified in Extract 15, copula *BE* is predicated by a noun which is complemented by a *that*-clause then further extended by two independent clauses coordinated by conjunctions (i.e., *and, because*). The findings suggest that while both the NNS and NS learners have appropriately and successfully used *This + BE* to create cohesion in their compositions, the NNS learners tend to produce syntactically simpler *this* clusters, which perform relatively fewer functions than those produced by the NS learners.

4. Many instances children becoming violent from what they read or see, like killing and shooting others. **This** *is* because, children cannot distinguish between the real world and the world of fiction. (NNS)

5. In the 90s, however, a person who chooses the tough "homemaker" position is made to feel a bit ashamed that he or she has taken the option of a non-marketplace, non-public and non-financially rewarded job. Perhaps **this** *is* due to the fact that women have successfully entered the once male-dominated work world, and thus the stay-at-home Mom feels torn or ambivalent because the media makes her feel behind the times. (NS)

It-clefts are also common in the support move. It normally takes the *IT + BE + ADJ* structure. Unlike pronoun *it*, the *it* used in the *it*-cleft clauses does not refer to anything. The focus of this construction is the post-predicate complement clause. According to Biber et al. (1999), *that*-clauses and *to*-clauses are the most common types of complement clause in *it*-extraposed position, and they can complement both adjectives and verbs. In both the NNS and NS learner essays, post-predicate complements are used mainly to complement adjectives. Attributive adjectives like *important*, *difficult* or *easy* are used to relay the writer's evaluation or judgment of an issue being discussed. Hasselgård (2014) in her investigation of *it*-clefts in five corpora representing Norwegian learners of English, novice L1 writers of English, and specialist L1 academic writing found that L2 learners underused clefts and the types of clefted constituent and the choice of subordinator in the cleft clause are more restricted in L2 writing than the L1 counterparts. Similar to this study, even though *it*-cleft structure is employed by both the NS and NNS learners, NNS learners tend to produce structurally shorter *it*-clefts with comparatively fewer attendant elements.

6. When one reads their literature, **it** *is* **not surprising** that 90% of our families have been destroyed or damaged by divorce because of relationships and adultery. (NNS)
7. **It** *is* **difficult** to imagine, however, a society where workers are encouraged to take employment seriously - to say nothing of finding satisfaction from employment - and where the workforce is filled with a diversity of willing workers, without considering the fundamental value of childcare providers. (NS)

There are four possible move sequences for this stage and they include marker, consolidation, affirmation and close. Only the consolidation move is considered indispensable in the conclusion stage, while the others are optional.

BE-copula constructions in the conclusion occur mostly in the affirmation move. The function is mainly to reaffirm the writer's stance on the topic. This is achieved either by restating the evaluation made in the proposition or establishing a final evaluation on the topic. Normally,

the evaluation is realized using the *BE + ADJ* or *BE + NP* sequences. The affirmation helps to reinforce the writer's stance on the topic, and the use of attributive adjectives attempts to influence the reader into adopting the same opinion.

1. To wholly implement rehabilitation on all sorts of criminals *is* **foolish**. It *is* **logical** to rehabilitate an addict (although it does not always work), but not all criminals *are* **suitable** for rehabilitation. (NNS)

2. It *is* **obvious** that crime pays. This is less of a choice than it *is* **an act of taking advantage of an opportunity**. There are few alternatives for low-income youths, and fewer obstacles in the quest for cash. (NS)

5.0 CONCLUSION

In conclusion, there is almost a similar pattern of the use of *BE* between NNS and NS learners, with both groups showing an inclination towards the present forms (*is, are*) and *BE*-copula construction. *T*-tests revealed more significant use of *BE* in terms of both forms and functions by the NS learners, indicating more prominence use of *BE* in the NS learner writings. Nonetheless, the in-depth qualitative analysis of the use of *BE*-copula found even though the NNS learners exhibit almost similar composing elements as the NS counterparts, their texts were stylistically simple, more constrained, less fluent and effective as the result of limited syntactic variety. The study recommends the use of corpora as one way to address some of the problems NNS speakers are facing with *BE*-copula. The NNS learners should be exposed to a large sample of lexical and syntactic varieties and be consciously taught about these varieties and the incorporation of corpora in their EAP writing class is an excellent way to achieve these objectives.

REFERENCES

Aijmer, K. (2002). Modality in advance Swedish learners' written interlanguage. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and language teaching* (pp. 57-76). John Benjamins.

Ang, L. H., Tan, K. H., & Lye, G. Y. (2020). Error types in Malaysian lower secondary school student writing: A Corpus informed analysis of subject-verb agreement and copula be. *3L: The Southeast Asian Journal of English Language Studies, 26*(4), 127-140.

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.

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

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use the characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5–35.

Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics, 37*(1), 639-668.

Biber, D., & Reppen, R. (1998). Comparing native and learner perspectives on English grammar: A study of complement clauses. In S. Granger (Ed.), *Learner English on computer* (pp. 145–158). Addison Wesley Longman.

Bloch, J. (2010). A concordance-based study of the use of reporting verbs as rhetorical devices in academic papers. *Journal of Writing Research, 2*(2), 219–244.

Brazil, D. (1995). *A grammar of speech*. Oxford University Press.

Dissanayake, S. D., & Dissanayake, C. B. (2019). Common syntactic errors made by the undergraduates in writing English as a second language. *International Journal of Liberal Arts and Social Science, 7*(11), 1-8.

Flowerdew, L. (2003). A combined corpus and systemic-functional analysis of the problem–solution pattern in a student and professional corpus of technical writing. *TESOL Quarterly, 37*(3), 489–511.

Friginal, E., Li, M., & Weigle, S. C. (2014). Revisiting multiple profiles of learner compositions: A comparison of highly rated NS and NNS essays. *Journal of Second Language Writing, 23*(1), 1–16.

Galti, A. M. (2016). Awareness of students on the use of affective strategy and their level of speaking anxiety. *International Journal of Multidisciplinary Research and Development, 3*(1), 319-322.

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast: Paper from a symposium on text-based cross-linguistic studies. Lund studies in English Vol. 88* (pp. 37–51). Lund University Press.

Granger, S., Dagneaux, E., & Meunier, F. (2002). *The international corpus of learner English handbook and CD-ROM.* Presses Universitaires de Louvain.

Gray, B. (2010). On the use of demonstrative pronouns and determinersas cohesive devices: A focus on sentence-initial this/these in academic prose. *Journal of English for Academic Purposes, 9*(1), 167-183.

Hasselgård, H. (2014). It-clefts in English L1 and L2 academic writing. The case of Norwegian learners. In K. Davidse, C. Gentens, L. Ghesquière, & L. Vandelanotte (Eds.), *Corpus Interrogation and Grammatical Patterns* (pp. 295–319). John Benjamins.

Hyland, K. (1990). A genre description of the argumentative essay. *RELC Journal*, *21*(1), 66-78.

Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.

Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied Language Learning, 12*(1), 111–132.

Hinkel, E. (2002). *Second language writers' text. Linguistic and rhetorical features.* Lawrence Erlbaum Associates.

Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly, 37*(2), 275-301.

Hinkel, E. (2004). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar.* Lawrence Erlbaum.

Imani, A., & Habil, H. (2012). NNS postgraduate students' academic writing: Problem-solving strategies and grammatical features. *Procedia-Social and Behavioural Sciences, 66*(1), 440-471.

Imtiaz, Z., & Mahmood, M. A. (2014). Genre analysis of argumentative essays of Pakistani EFL Learners. *Journal of Education and Practice, 5*(30), 95-100.

Jarvis, S. (2007). Theoretical and methodological issues in the investigation of conceptual transfer. *Vigo International Journal of Applied Linguistics, 4*(1), 43–71.

Jiang, F. K., & Hyland, K. (2020). "There are significant differences…": The secret life of existential there in academic writing. *Lingua, 233*(1), 1-31.

Kanestion, A., & Singh, M. K. S. (2019). A corpus-based genre analysis moves in introductory paragraph of argumentative writing. *International Journal of Academic Research in Business and Social Sciences, 9*(7), 821–831.

Kanestion, A., Singh, S., Kaur, M., & Shamsudin, S. (2017). Developing a framework for writing skill: A corpus-based analysis of the written argumentative essays. *Sains Humanika, 9*(4-2), 39–47.

Knowles, G., Zuraidah, M. D., Jariah, M. J., Rajeswary, S., Janet, Y., & Sathiadevi. (2006). The Malaysian corpus of learner English: A bridge from linguistics to ELT. In H. Azirah & H. Norizah (Eds.), *Varieties of English in Southeast Asia and beyond*. University of Malaya Press.

Liu, D. (2015). Moves and wrap-up sentences in Chinese students' essay conclusions. *SAGE Open, 1*(1), 1-9.

Matesic, M., & Memisevic, A. (2016). Pragmatics of adjectives in academic discourse: From qualification to intensification. *Jezikoslovlje, 17*(1), 179-206.

McCarthy, M. (1994). It, this, and that. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 266-275). Routledge.

Milton, J. (2001). Elements of a written interlanguage: A computational and corpus-based study of institutional influences on the acquisition of English by Hong Kong Chinese Students. In G. James (Ed.), *Research report Vol 2*. Language Centre, the Hong Kong University of Science and Technology.

Paltridge, B. (2001). *Genre and the language learning classroom.* The University of Michigan Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language.* Longman.

Roslina, A. A., & Zuraidah, M. D. (2019). Tagging L2 writing: Learner errors and the performance of an automated part-of-speech tagger. *GEMA Online® Journal of Language Studies, 19*(3), 140-155.

Roslina, A. A. (2018). *A corpus-based study of the use of "BE" in Malay ESL learner essays* (Unpublished doctoral dissertation). University of Malaya.

Smirkou, A., & Smirkou, M. (2018). The development of the copula among Moroccan learners of English: The lightness of Be. *IOSR Journal of Humanities and Social Science, 23*(5), 67-78.

Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing, 32*(1), 17–35.

Swales, J. M. (1990). Nonnative speaker graduate engineering students and their introductions: Global coherence and local management. In U. Connor & A. Johns (Eds.), *Coherence in writing* (pp. 189-207). TESOL.

Swales, J. M. (2005). Attended and unattended "this" in academic writing: A long and unfinished story. *ESP Malaysia, 11*(1), 1-15.

Swales, J. M., & Feak, C. (2012). *Commentary for academic writing for graduate students.* University of Michigan Press.

Wulff, S., Römer, U., & Swales, J. (2012). Attended/unattended *this* in academic student writing: Quantitative and qualitative perspectives. *Corpus Linguistics and Linguistic Theory, 8*(1), 129-157.