Error Types in Malaysian Lower Secondary School Student Writing: A Corpus-Informed Analysis of Subject-Verb Agreement and Copula *be*

ANG LENG HONG
School of Humanities,
Universiti Sains Malaysia, Malaysia
lenghong@usm.my

TAN KIM HUA
Centre for Research in Language and Linguistics,
Universiti Kebangsaan Malaysia, Malaysia

LYE GUANG YANG
School of Humanities,
Universiti Sains Malaysia, Malaysia

ABSTRACT

The issue of English language proficiency among Malaysian students has always been a subject of ongoing discussion among educators. While English is taught as a second language in primary and secondary schools, many students, especially those from rural areas, have problems using the language correctly. Also, it has been said that students' English language proficiency in most rural contexts takes on the quality of a foreign language rather than a second language. The current study addresses this problem by examining the use of Subject-Verb Agreement and copula **be** *in essays written by 32 Malay lower secondary school students from a rural school in Kedah state, Malaysia. Using corpus-linguistic approach, this study compiled 128 essays and tagged them with online part-of-speech (POS) UCREL CLAWS tagger. This is a novel attempt in integrating Error Analysis (EA) framework with a corpus tagging approach to examine error patterning in learner writing. Using corpus-based techniques, this study identifies and analyses errors associated with SVA and copula* **be** *in Malaysian lower secondary school student writing. The findings showed that student writings were riddled with various types of errors, namely misselection, omission, blends, and overinclusion. Based on these findings, this study offers pedagogical suggestions for improving the teaching and learning of the language.*

Keywords: learner writing; Error Analysis; Subject-Verb Agreement; copula **be***; part-of-speech tagging*

INTRODUCTION

English language is considered as the lingua franca of the commercial sphere with more trades and businesses going global. The importance placed on English in the workplace is highlighted in many local studies that focus on the issue of graduate employability (Nadzrah et al., 2011; Rohani Salleh et al., 2015; Cheong et al., 2016; Lim et al., 2016; Ting et al., 2017; Ilhaamie Abdul Ghani Azmi et al., 2018; Siti Zaidah Zainuddin et al., 2019). These studies show evidence that English language proficiency is one of the most critical soft skills expected by most employers. While there is an emphasis on the importance of English in the workplace, the issue of the low level of English proficiency among Malaysians has always been highlighted by stakeholders in the private sector (Ang et al., 2017; Ilhaamie Abdul Ghani Azmi et al., 2018; Ang & Tan, 2019; Siti Zaidah Zainuddin et al., 2019).

English language development in Malaysia is closely related to the various policies implemented since the country gained its independence. The introduction of the New Education Policy in 1970 marked the beginning of the English language as a Second Language in Malaysia. Two types of public primary schools, namely the national schools and the national-type schools have been established together with the New Education Policy. In national schools,

the medium of instruction was Malay language and English was taught as a subject from Standard One, while in national-type schools, ethnic languages such as Tamil and Mandarin were used as mediums of instruction with English taught as a subject only since Standard Three. Consequently, students' access to English language has suffered from a constant decrease (Hazita Azman, 2016). There is a significant drop in English language proficiency among many Malaysian students, particularly those from the rural areas. Realising the importance of English in both local and global contexts, the Malaysian Ministry of Education has initiated several policies and measures to improve the deteriorating English language standard among Malaysian students. For instance, there were decisions to lower the starting age to introduce English in the primary schools and increase the number of lessons taught in the English language. However, these measures have generally been unsuccessful. It has been reported that fewer than 50% of students who have completed six years of elementary school were literate in English (EPU 2016, pp. 10-2). Although the English language is taught as a second language in the formal schooling system starting from preschool, most students' English proficiency, especially from rural areas, remains unsatisfactory. Past studies have shown that Malaysian learners still face difficulties in English writing (Siti Hamin & Mohd Mustafa, 2010; Ang et al., 2011; Wong, 2012; Manokaran, Ramalingam & Adriana, 2013; Hazita Azman, 2016; Ang & Tan, 2018).

Concerning the problems faced by language learners, numerous studies have used Error Analysis (EA) framework proposed by scholars such as Corder (1981), Ellis (1994), and James (2013) to explore learner language. Errors committed by learners have been the focus in many studies as a good understanding of errors is necessary for improving language pedagogy (Siti Hamin & Mohd Mustafa, 2010; Ang et al., 2011; Satake. 2020), particularly in the context related to the construct of language accuracy or writing precision in learner language. While there are numerous research conducted using EA framework, many of these previous studies rely on the conventional way of identifying errors, i.e., manually inspect the data for possible language inaccuracies. This may result in overlooking certain errors, especially in large amount of writing. A solution to this issue is to extend the analysis to include parts-of-speech (POS). The corpus data is tagged for extracting the relevant grammatical patterns, for instance, verb constructions automatically. This allows researchers to examine the relevant language patterns more comprehensively for language accuracies or writing precision purposes.

LITERATURE REVIEW

Errors are considered signs of underlying problems or result from complex language interactions (Corder, 1981). An error occurs when a deviation arises due to a learner's lack of knowledge (Ellis, 1994). Learners make errors because they have formed a false hypothesis about the target language. This shows their lack of knowledge of the language system. Errors committed by learners are crucial in improving second language teaching and learning (Corder, 1981; Ellis, 1994; James, 2013; Xie, 2019) as errors are regarded as valuable insight for the learners to learn, make hypotheses about the language, and lastly gain improvement in language learning. Learners' errors show that every learner has an interim, still incomplete language system, known as *Interlanguage* (Selinker, 1972). The term Interlanguage was coined by Selinker (1972) to refer to the linguistic system produced when a second language learner attempts to convey meaning in the target language. This linguistic system is a system that is different from the mother tongue or native language of the learner and the target language that the learner is learning (Tarone, 1988).

The purpose of EA is to determine what the learners know and do not know in terms of the target language (Corder, 1974). When the known and unknown are identified, appropriate

guidance and solution can be provided to the learners to improve on the language. By conducting EA, the information and data about the errors and the target forms enable learners to gain sufficient knowledge about the concept of a rule in the target language (Corder, 1981). James (2013) elaborated that EA is the process whereby the error analyst determined and identified the occurrences, characteristics, causes, and consequences of unsuccessful language. He provided a more thorough explanation of EA by describing the learners' interlanguage and the target language system in a more objective way. When the learners' interlanguage and target language systems are elucidated, the researcher can analyse and compare the two linguistic systems to locate mismatches (James, 2013). EA is a fundamental tool in language teaching to re-organise a teacher's perceptions and re-adjust his or her teaching methodology to help learners fix the gaps or weaknesses in their language use. EA has always had a more practical application in second language teaching to resolve the perennial problem of accuracy and is therefore, important in facilitating language teachers in improving second or foreign language teaching.

While EA remains an effective approach in exploring learner language for pedagogical purposes, findings generated can only be as accurate and valid as the theoretical context and system of linguistic explanations applied to them. This indicates that EA framework is limited in describing grammatical aspects on the surface level (Hamilton, 2015; Mcdowell, 2020). The EA framework by Corder (1974, 1981) is therefore not able to facilitate the analysis of errors from the functional perspective. In response to this limitation, researchers (e.g., Kim, 2010; Hamilton, 2015; Mcdowell, 2020) elaborated the conventional EA framework with a Systemic Functional Linguistics (SFL) approach that enables functional analyses and that it offers insights into the system-structure relationships of grammar (Matthiessen & Halliday, 2009; Martin, 2013, Mcdowell, 2020). Undoubtedly, the elaborated EA framework is useful in defining errors across different levels of expressions. Nevertheless, it is worth pointing out that this SFL elaborated EA is more useful in the context where researchers aim at analysing the metafunctions of the language aspects to see how errors disrupt the flow of information (Mcdowell, 2020). It is therefore not surprising that some EA studies still adopt the traditional EA framework since their research aim was to examine errors on the surface structures that are pedagogically more useful and simpler to be discussed in the classroom settings (Hamilton, 2015). This is important as it motivates language teachers to consider using corpora in language teaching to address the phenomenon in which the use of corpora in classroom teaching has not been popular (Flowerdew, 2010; Tribble, 2015).

In analysing error patterns, Corder (1974, 1981) proposed four different ways in which the errors occur: addition of an unnecessary element; omission of an obligatory element; misordering in terms of the sentence position; and misselection. This traditional EA classification has been elaborated over time due to its practical application in describing errors following the theory of Interlanguage (e.g., Dulay, Burt, & Krashen, 1982; James, 2013). In traditional EA framework, error types are identified by following the traditional school grammar fashion: i.e. errors on general word classes. According to James (2013), the most common error types found in learner writing included tenses, prepositions, Subject Verb Agreement (SVA), and weak vocabulary.

EA studies have been conducted in different contexts, involving learners with different first language (L1) backgrounds. For example, focusing on Malaysian context, Marlyna Maros, Tan Kim Hua, and Khazriyati (2007) performed an EA of L1 Malay learner writing in selected rural secondary schools. They found that the most frequent errors made by the rural Malay learners included article, SVA, and copula **be** errors, and these learners had the tendencies to omit the copula **be**. Also, Siti Hamin and Mohd Mustafa (2010) conducted an EA study to investigate SVA errors committed by 20 postgraduate students from a teachers' training institute. The study identified five types of SVA errors: the agreement of person, number,

indefinite expression of amount, coordinated subject, and proximity. The researchers revealed that although the students were in their tertiary level, they still committed many SVA errors in their writing. Using collocational method, Ang et al. (2011) investigated collocation errors in L1 Malay learner writing, finding that the most prominent error pattern was pattern associated with preposition. Kong (2013) conducted a study that examined the verb errors committed by L1 Chinese learners in a secondary school in Malaysia. Errors associated with SVA, simple past and simple present tense were identified. The findings revealed that the error types were attributed to selection, addition, and omission processes. Furthermore, the researcher found that the learners' L1 background, learning environment, and exposure to the target language had affected the way the learners learnt the language. In the foreign context, Dahlmeier, Ng, and Wu (2013) examined errors made by Singaporean learners and found that Singaporean learners frequently made preposition and article errors. Recently, Satake (2020) conducted a study on the errors made by 55 undergraduate students from a university in Tokyo and discovered that L1 Japanese students frequently made lexical errors through omission. Revision tasks were performed by the students as an effort of error correction.

While many studies have been conducted to look into learner writing errors, these past studies employed the conventional method in identifying the errors, i.e., manually inspect the corpus, except Satake (2020). This study integrates EA with the corpus tagging approach in identifying the specific grammatical constructions comprehensively to minimise the chances of overlooking certain SVA and copula **be** errors in learner writing. Specifically, it aims to examine the English essays written by selected lower secondary school learners from a rural area school in Kedah state, Malaysia. Apart from students' low proficiency in the language, this research also considers the debate on rural school students' language ability compared to that of urban school students. Also, this study is a novel attempt in integrating EA with a corpus tagging approach to examine error patterning in learner writing. Using corpus-based techniques, this study identifies and analyses errors associated with SVA and copula **be** in lower secondary school student writing. The research question of this study is:

1) What are the types of Subject-Verb Agreement (SVA) and copula **be** errors that are found in the Malaysian lower secondary student writing?

METHODOLOGY

This corpus-based study aims at identifying grammatical errors, namely SVA and copula **be** errors in a self-compiled learner corpus. The traditional EA framework was adopted in conducting the data collection and analysis procedures. A total of 32 Form Two Malay students from a rural secondary school in Kedah state, Malaysia, were asked to write English essays in classroom settings. Form Two students were chosen because they were the product of the ongoing new curricula (KSSR and KSSM), whereby grammar is more prominent in the teaching and learning process. The participants possessed intermediate proficiency level in the English language. The selection was based on the results of their final examination when the participants were in Form One. In this study, four writing tasks were used to build the learner corpus to investigate the types of SVA and copula **be** errors. The writing tasks were conducted in four sessions whereby the participants were required to write an essay in each session. They were given 45 minutes to complete the essay of not less than 150 words. The requirements were similar to those of the current Form Three school-based assessment, *Pentaksiran Tingkatan 3* (PT3), where the students were asked to write descriptive and narrative essays, respectively, in four sessions. The researcher obtained permission from the principal to conduct the writing sessions.

A total of 128 essays (24,037 words) were collected and digitised manually into the plain text format. The essays were then annotated to show the part-of-speech (POS) of each word using online part-of-speech (POS) UCREL CLAWS tagger. *AntConc* (Version 3.5.8) software was used to generate and analyse the corpus. POS tagging is useful in analysing learners' errors (Roslina Abdul Aziz & Zuraidah Mohd Don, 2019). The data collection procedures adapted the steps proposed by Corder (1974), namely collecting data, identifying errors, classifying errors, analysing errors, and explaining errors. Errors were classified into different categories: misselection, omission, blends, and overinclusion errors (James, 2013). The classification of the error types serves to answer the research question: What types of errors are found in the use of SVA and copular *be* in the student writing?

With regard to the error categories, misselection errors occur when the wrong form of a verb is used in a sentence. The learners' written production would show the incorrect selection of copula *be* and other wrongly selected verb form. Omission errors take place when a word or a linguistic element is omitted in a sentence. The omission errors can be identified in the corpus whereby the verb (for example, copula *be* and the main verb) is omitted from the sentence. Blend errors occur when the learner is indecisive about determining which linguistic forms to select when using SVA constructions. This kind of errors are the blending of two well-defined forms which resulted in erroneous form. Overinclusion errors occur when a redundant word or linguistic element is added to the verb phrase. Table 1 shows the types of errors and their sub-categories.

TABLE 1. Types of errors (adopted from James, 2013)

Type of Error	Sub-category
1) Misselection	a) The use of a singular *be* verb for a plural subject
	b) The use of a plural verb for a singular subject
	c) The use of a singular verb for a plural subject
	d) The use of a plural *be* verb for a singular subject
	e) The use of an improper subject for the *be* verb
2) Omission	a) Omission of the singular copula *be*
	b) Omission of the plural copula *be*
	c) Omission of the main verb
3) Blends	a) Standard blend
	b) Deletion
	c) Addition
4) Overinclusion	a) Double marking
	b) Regularisation

DATA ANALYSIS AND RESULTS

The learner writing was analysed and the SVA and copula *be* errors were identified. The errors were categorised into four types, namely misselection, omission, blend, and overinclusion. Misordering, another type of error, was excluded because there was no error found in this category. The distribution for each type of error is shown in Figure 1.
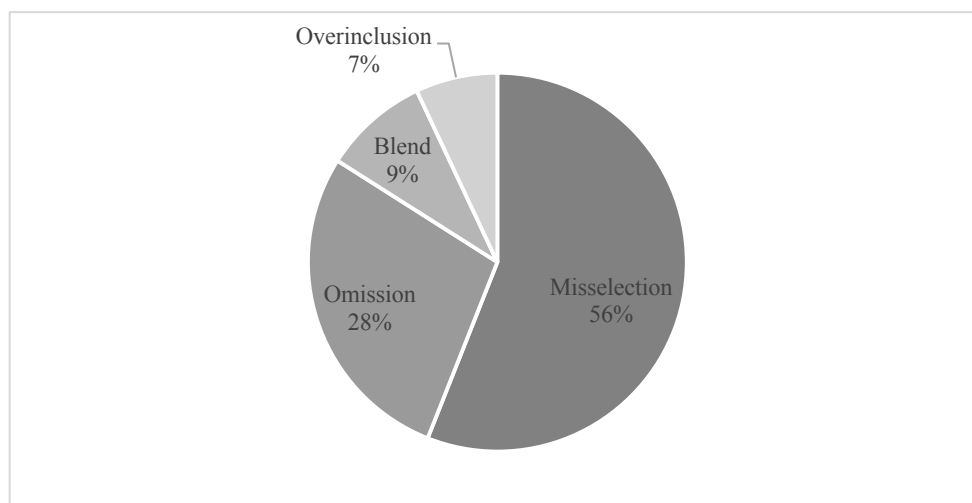
FIGURE 1. Types of errors

Based on the findings, a total of 475 errors were identified. Misselection error is the most frequently committed, constituting 266 errors or 56.00%, followed by omission errors (132 errors or 28%), blend errors (42 errors or 9%), and overinclusion errors (35 errors or 7%).

MISSELECTION ERRORS

The errors under the category of misselection were further classified into different sub-categories, as shown in Figure 2.
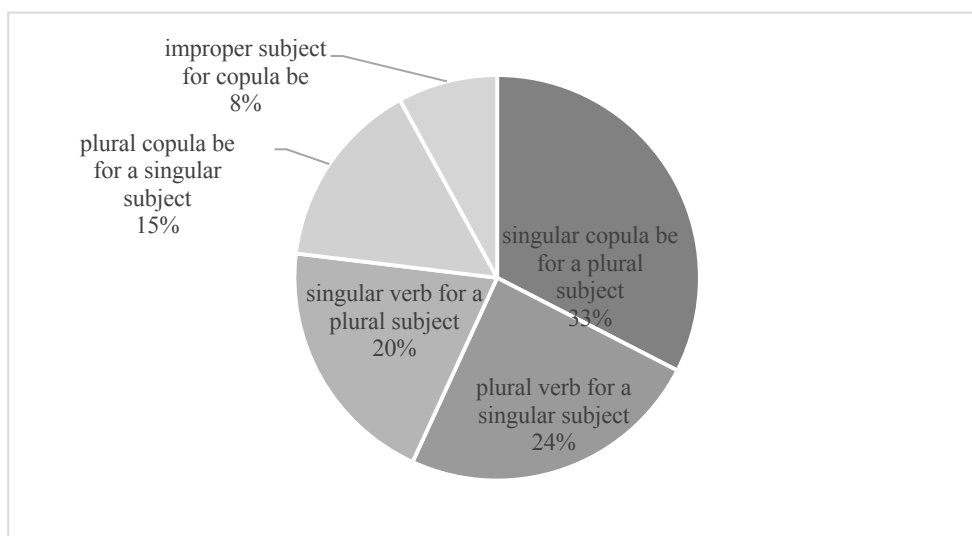

FIGURE 2. Sub-categories of misselection errors

The examples were extracted from the Concordance window in *AntConc* software. Some instances of misselection errors are as follows:
a) The use of singular *be* verb for a plural subject

1. Among the games **is** coconut fruit bowling ….
2. This activities **is** for teachers and ….
3. There **is** many teachers in ….
4. morning, there **was** colourful ribbons ….

In the first and second examples, the singular copula *be* verb in bold was used for the plural subjects **games** and **activities** at the beginning of the examples. The correct selection of copula *be* should be **are** or the past tense form **were** so that the verb agrees with the plural subject. For example 3 and 4, the learner used singular copula *be* **is** and **was** for the plural subjects **many teachers** and **colourful ribbons**. The existential **There** in the two examples carried no emphasis whether the copula *be* should be in singular or plural form. The plural subjects that followed the copula *be* should be identified. Thus, plural copula *be* **are** and **were** should be used.

b) The use of a plural verb for a singular subject

 1. he **appreciate** it ….
 2. the principal **give** the opening speech ….

In the above examples, the learner used a plural verb (in bold) for the singular subject. The suffix **-s** should be placed after the stem word to indicate it as a singular form. The word **appreciate** should be **appreciates** to agree on the singular subject **he** while the word **give** should be **gives** to agree on the singular subject **principal**.

c) The use of a singular verb for a plural subject

1. We **eats** with joy as ….
2. Many teachers **feels** very excited to try ….

The above examples show the use of singular verbs for plural subjects. A singular subject should have a singular verb and vice versa. In these two examples, the subjects were in plural form. However, the verbs **eats** and **feels** were in singular form with the suffix **-s**. The correct selection of verbs should be in plural form as the subject **We** and **Many teachers** were plural subjects. Thus, the verbs should be in the simple present form **eat** and **feel** without the suffix -s which will contribute to the singularity of the verb.

d) The use of a plural *be* verb for a singular subject

1. The competition **were** into ten stations.
2. This activity **were** fun when looked ….

Example 1 and 2 had a plural copula *be* as the verb. However, the subjects **The competition** and **This activity** were singular subjects. Therefore, the verb should be in singular form as well. The correct form would be the singular copula *be* **was** in both examples.

e) The use of an improper subject for the *be* verb

1. **The competition** were bowling using ….
2. **All of the teacher** were required to participate ….
3. **One of the historical site** in Melaka is ….

In example 1 and 2, the copula *be* which was selected was the plural form *were*. The verb was chosen correctly but the wrong selection occurred in the selection of the subjects. For sentence 1, the subject should be in plural form, **The competitions** because the following noun phrase indicated a few examples of different competitions that took place during the event. As

for sentence 2, the indefinite pronoun **All** was used before the antecedent, the object of the preposition (the noun **teacher**). Based on the SVA rule, an indefinite pronoun relies on its antecedent in order to determine whether the pronoun is in singular or plural form. The antecedent simply means the object of the preposition phrase (in this case, the word **teacher**). In context, the countable noun **teacher** should be in plural form **teachers** to indicate the plurality of the indefinite pronoun **All**. Thus, the correct phrase should be **All the teachers were**. In example 3, the subject **One of the historical site** was erroneous because the phrase **One of** should follow a plural noun, in this case **the historical sites**. The selection of copula *be* **is** was appropriate as the phrase indicated singularity.

<div align="center">OMISSION ERRORS</div>

The errors under the category of omission were further categorised into three sub-categories, as presented in Figure 3.
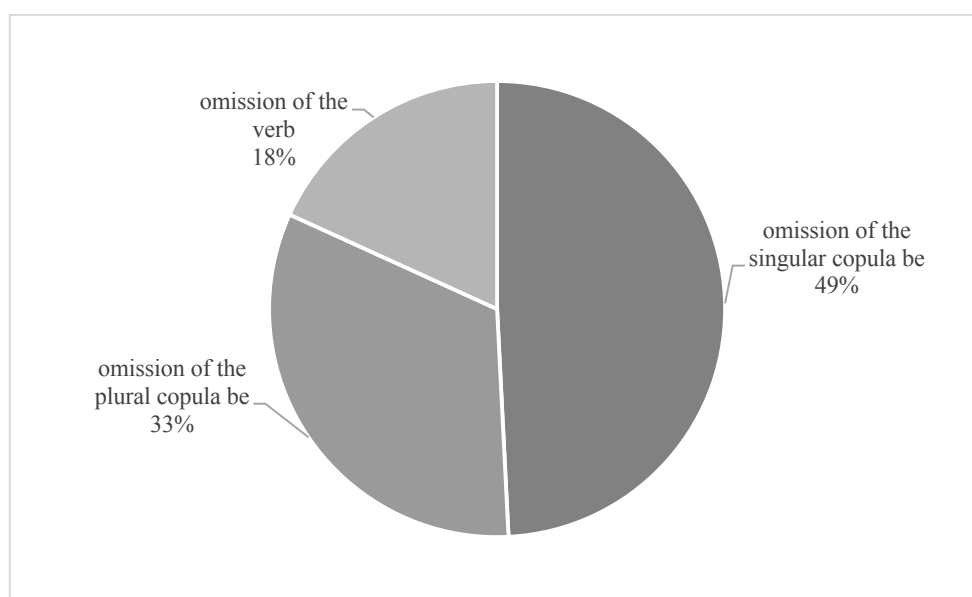

<div align="center">FIGURE 3. Sub-categories of omission errors</div>

Some instances of omission errors are as follows:
a) Omission of the singular copula *be*

1. It Ø memorable day ….
2. It Ø conducted by ….

In example 1 and 2, the copula *be* was missing, making them ungrammatical. The subject was a singular subject **It**. Thus, based on the context, the learners should include a past tense form of singular copula *be* after the subject to talk about the past experience. The correct form should be **was**.

b) Omission of the plural copula *be*

1. The teachers and students Ø tired but happy.
2. security guards Ø very professional ….

In the first example, the learner omitted the plural copula *be*. The correct sentence should include the past form plural copula *be* **were** after the plural subject of **teachers** and

students. Example two also shows the similar situation whereby the learner omitted the plural copula ***be***. The present form plural copula ***be*** should be placed in the position of the Ø.

c) Omission of the main verb

1. students and teachers Ø in the open hall ….
2. Neighbour Ø together to help ….

The learners omitted the main verb in the examples above, making them incomprehensible. Based on the context, the omitted main verb in example 1 could probably be action verbs in the plural form such as **gather/assemble** or in the past tense form **gathered/assembled** to show the action of the subjects **students and teachers** coming together into the hall. On the other hand, in example 2, the possible verb to replace the Ø is the verb **work**, which is singular to agree to the singular subject **Neighbour**.

<div align="center">BLEND ERRORS</div>

The blend errors were categorised into three sub-categories, as shown in Figure 4 below. However, only one sub-category of blend errors were found in the learner writing, namely standard blend. Standard blend implies that the errors consist of part of each target sentence structure and those structures are inhibited and/or partly is used.
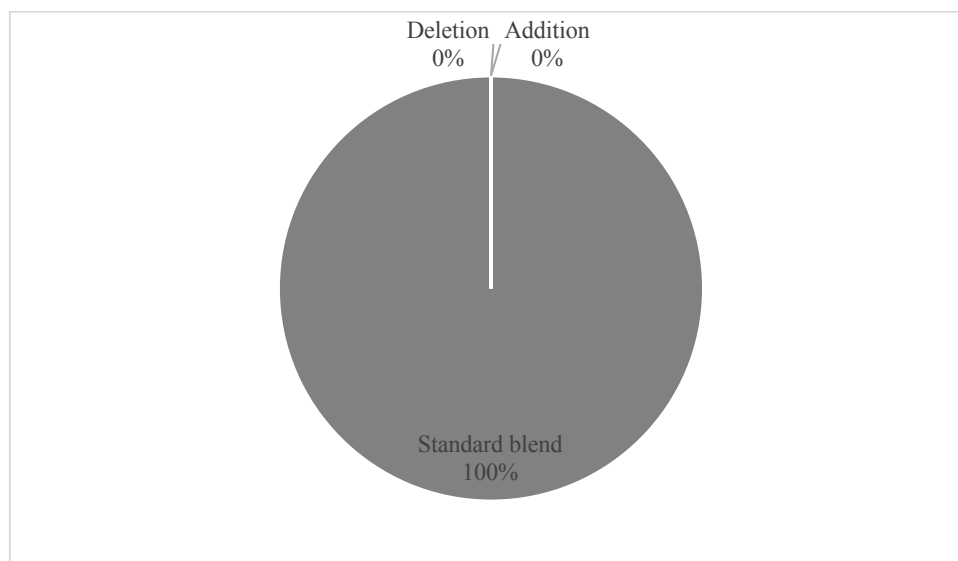

FIGURE 4. Sub-categories of blend errors

Some instances of blend errors are as follows:
Standard blends

1. teachers who …   **was divides** us into small ….
2. **Our has** 400 students ….

In example 1, there was a blend of two grammatical structures in the target language. For instance, in example one, the subject was a plural subject, and thus the misselection of the auxiliary **was** occurred. Based on the context, the learner was trying to construct a past continuous tense sentence beginning with an adverbial clause **At that time**. Hence, the selection of the verb **divides** was also erroneous. The correct form should be **were dividing** whereby the auxiliary verb agrees with the plural subject followed by a present participle verb

form to indicate continuous action. As in example 2, the learner committed misselection of the subject as well as the verb **has**. The pronoun in the example 2 should be a subject pronoun **We** instead of possessive determiner **Our**. Since the subject was a plural subject, the verb should agree with the subject by having a plural form **have**.

OVERINCLUSION ERRORS

The overinclusion errors were further categorised into double marking and regularisation, as shown in Figure 5. This type of error is the least common one in the selected lower secondary student writing.
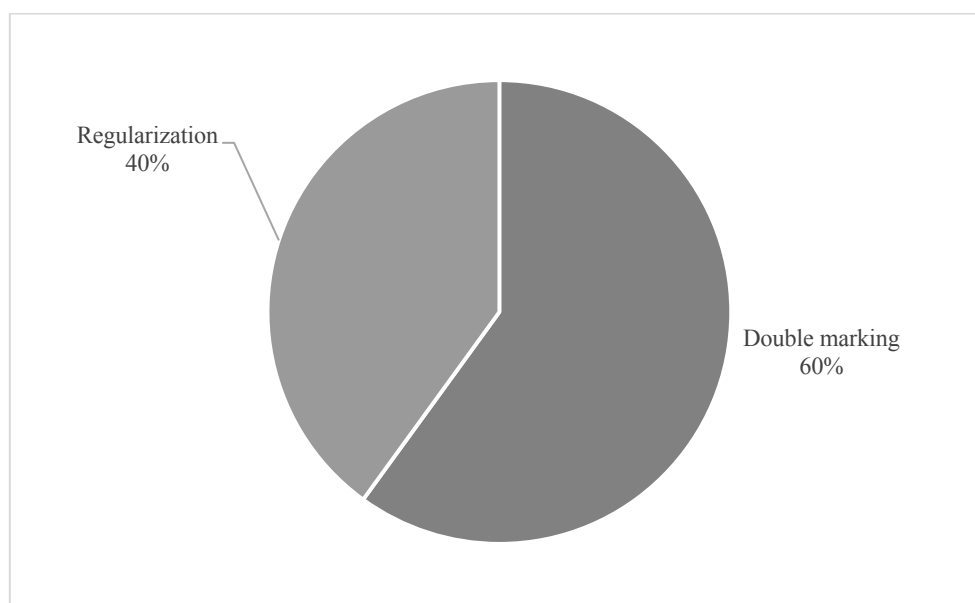

FIGURE 5. Sub-categories of overinclusion errors

Some instances of overinclusion errors are as follows:
a) Double marking

1. My teacher **did divides** the tasks ….
2. it **can attracts** tourist ….

Double marking in overinclusion error means the failure to eliminate certain linguistic items which are not required in other linguistic constructions. Example 1 had an overinclusion of the suffix **-s** to indicate a singular verb. However, the suffix **-s** was not required after the auxiliary verb **do**. Thus, the word **did** which already indicated past tense did not require the verb that followed to be further modified. In example 2, the modal verb **can** required the verb's base form and did not require the overinclusion of any suffixes. The addition of the suffix **-s** by the learner shows double marking.

b) Regularisation

1. The first group **leds** by Farisha ….
2.   Third group **leds** by Sajirah and ….

The verb **led** is the past tense form of **lead**. Regularization occurred when the learner added the suffix **-s** to the past tense form **led**. It involved the ignorance of the exceptions of

adding the suffix **-s** to the past tense form verb to indicate singularity that agrees with the singular subject. The correct form for the two examples is **led**.

DISCUSSION

The findings indicate that the lower secondary school learners faced difficulties in choosing the correct form of verb and subject to form grammatically correct sentences. This shows that some Malaysian learners, particularly those from rural areas were still unable to use the verb (main verb and copula *be*) correctly. Looking at the previous literature, Kong (2013) found out that in terms of SVA, the selection errors (similar to the misselection errors in the present study) constituted the highest frequency (82 errors or 70.7%), followed by omission errors (34 errors or 29.3%). There were no addition and misordering error for the SVA category. Although the participants in his study consisted of Chinese learners, the results were similar in that the error types that constituted the highest frequency was misselection errors. This indicates that learners were aware of the SVA rules, but they were unable to use the correct form of them. Besides, the finding indicates that Malay learners face difficulties in using SVA correctly in school and in institution of higher learning, as indicated in Siti Hamin and Mohd Mustafa's (2010) study. This trend is worrying as SVA has been taught to learners since they were in the primary school. Their failure in mastering the SVA rules should be given more research and pedagogical attention, particularly in the primary school settings.

In terms of misselection errors, the learners of the present study incorrectly selected singular *be* verb for the plural subject. They were confused about the singularity and plurality of the subject as well as the verb. They intended to use words such as **is** and **was** for most subjects without differentiating between the singular or plural noun. Besides, the learners also committed several misselection errors whereby they used a plural verb for a singular subject and vice versa. This indicates that some learners were unaware or unsure about the use of suffix -s that plays different roles in a noun and a verb. The learners were confused and they assumed that the suffix -s would show plurality when added to the verb. Thus, this led to the selection of the erroneous verb form. With regard to omission errors, they were the second most prevalent error type emerging from the data. The finding is similar to that of Marlyna Maros, Tan Kim Hua, and Khazriyati (2007), that Malay learners had the tendencies to omit the copula *be*.

In this present study, there were also blend errors and overinclusion errors identified in the learner writing. Blend errors involved the combination of two grammatical forms which resulted in erroneous structure. On the other hand, overinclusion errors involved the double marking and regularisation errors that deal with redundant linguistic items in the formation of a verb. These errors indicate that the learners were incompetent in using English language in general, although they were exposed to the new curricula (KSSR and KSSM) that emphasise the grammar teaching and learning.

Above all, the findings from the learner corpus analysis are able to answer the research question of the study: What types of errors are found in the use of Subject-Verb Agreement (SVA) and copula *be* in the student writing? It can be concluded that the types of errors in the use of SVA and copula *be* in student writing are misselection, omission, overinclusion, and blend errors, with misselection as being the most prevalent error type.

POTENTIAL APPLICATIONS IN TEACHING AND LEARNING OF ENGLISH WRITING

From the findings discussed, constructive actions should be taken to address Malaysian student writing's weaknesses. Particular attention needs to be given to students from rural areas as they

show overt weaknesses in using correct grammar in their writing. Students should be made aware of the importance of learning and acquiring knowledge of grammar. According to Siti Hamin and Mohd Mustafa (2010), language teachers should inform students explicitly about the importance of the subject and verb since these two linguistic elements are the two most essential components in constructing correct and complete sentences in the English language. Another pedagogical implication is aptly described by Candling (2001) that second language learners' grammatical errors should be seen as a potentially crucial element for understanding the process of language learning. Thus, the findings on EA can determine what learners still need to be taught in the English language classroom. The refined error list with error descriptions and examples can be used in a versatile way to inform the English language teaching and writing. Revision tasks can be given to students to correct the errors they make, for instance by using reference corpus in correcting the errors (Satake, 2020). The findings of this study demonstrate the significance of EA as it provides evidence of the types of errors committed by the learners. After identifying the types of errors, teachers will have the advantage of focusing on the learners' grammar teaching. Specifically, the errors identified in student writing can help language teachers provide constructive feedback to the learners regarding the use of specific grammar rules. By letting the learners know about the erroneous form and the proper form, they can carry out their correction to improve their proficiency in the language.

LIMITATIONS AND FURTHER RESEARCH DIRECTIONS

In further endeavours, researchers may want to overcome or avoid the following limitations of this study. Firstly, the present research attempted to better understand lower secondary Malay learners' grammatical problems. Nonetheless, it provides only a refined explanation of the problems without seeking to analyse their underlying causes. A logical next step is to investigate the possible causes of these grammatical errors. Future research may seek to provide an explanation of the causes of errors by comparing language learners' first language system and that of the target language. Secondly, while the analysis presented here enables the pinpointing of SVA and copula *be* errors, it did not look at errors involving other grammatical aspects. While a detailed analysis of every grammatical inaccuracy is beyond the scope of this research, future work with this corpus-based EA framework may include the analysis of other grammatical aspects in learner language.

CONCLUSION

The present research is a novel attempt in integrating EA with a corpus tagging approach to examine grammatical error patterning in Malay learner writing. The present study adopted the traditional EA framework as it is deemed more suitable and simpler for language teachers to identify and understand common errors made by learners. This study has shed some light on how Form Two Malay learners in rural context used SVA and copula *be* in their writing. While the EA framework has been used in numerous second language research, it is still useful and relevant to today's context as Malaysian learners still have problems using correct grammar after years of learning the language. An important step to be taken by language teachers is to view learners' errors positively. EA has excellent value in classroom research as it is a systematic analysis of errors made by the learners. Errors can provide an objective perception of how learners learn a second or foreign language. Thus, conducting EA is considered an excellent alternative to describing and explaining the forms of language learners' errors. To sum up, the findings of the study revealed that some Malay learners still face difficulties in

using SVA and copula *be* correctly. They were found to commit different types of errors, namely omission, overinclusion, misselection, and blend errors in their writing. Some of them were still unable to construct grammatically acceptable sentences. Therefore, remedial actions must be taken by the learners, teachers, educators, and curriculum designers to improve this situation as the English language plays many vital roles in the process of knowledge acquisition and transfer, as well as in ensuring effective communication in various local and global contexts.

ACKNOWLEDGEMENTS

REFERENCES

Ang, L. H., Hajar, A. R., Tan, K. H., & Khazriyati, S. (2011). Collocations in Malaysian English learners' writing: A corpus-based EA. *3L: The Southeast Asian Journal of English Language Studies*, *17* (Special Issue), 31-44. http://journalarticle.ukm.my/3043/

Ang, L. H., Tan, K. H. & He, M. (2017).A Corpus-based Collocational Analysis of Noun Premodification Types in Academic Writing. *3L: The Southeast Asian Journal of English Language Studies*, *23*(1), 115-131. https://doi.org/10.17576/3l-2017-2301-09

Ang, L. H. & Tan, K. H. (2018). Specificity in English for Academic Purposes (EAP): A corpus analysis of lexical bundles in academic writing, *3L: The Southeast Asian Journal of English Language Studies*, *24*(2), 82-94. https://doi.org/10.17576/3l-2018-2402-07

Ang, L. H. & Tan, K. H. (2019). From Lexical Bundles to Lexical Frames: Uncovering the extent of phraseological variation in academic writing. *3L: The Southeast Asian Journal of English Language Studies*, *25*(2), 99-112. https://doi.org/10.17576/3l-2019-2502-08

Anthony, L. (2018). *AntConc* (3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available at: www.laurenceanthony.net/

Candling, R. B. (2001). *Vocabulary and language teaching*. New York: Longman Publishers.

Cheong, K. C., Hill, C., Fernandez-Chung, R., & Leong, Y. C. (2016). Employing the 'unemployable': employer perceptions of Malaysian graduates, *Studies in Higher Education*, *41*(12), 2253-2270. https://doi.org/10.1080/03075079.2015.1034260

Corder, S. P. (1974). *Techniques in Applied Linguistics*. Oxford: Oxford University Press.

Corder, S. P. (1981). *EA and Interlanguage. London.* Oxford: Oxford University Press.

Dahlmeier, D., Ng, H.T., Wu, S.M. (2013). Building a large annotated corpus of learner English: The NUS corpus of learner English. *Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications*. Stroudsberg: Association for Computational Linguistics, 22-31.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. New York: Oxford University Press.

Ellis, R. (1994). *The Study of Second Language Acquisition*. New York: Oxford University Press.

Flowerdew, L. (2010). Using corpora for writing instruction. In A. O'Keeffe, & M. McCarthy (Eds.). *The Routledge Handbook of Corpus Linguistics* (pp. 444-457). New York: Routledge.

Government of Malaysia (GoM) (2016). Strategy Paper 10: Transforming Education System, 1-36. Putrajaya: Economic Planning Unit. Available at: http://www.rmk11.epu.gov.my/pdf/strategy-paper/Strategy

Hamilton, C. E. (2015). The contribution of systemic functional grammar to the EA framework. *TESOL International Journal*, *10*(1), 11-28.

Hazita Azman. (2016). Implementation and challenges of English language education reform in Malaysian primary schools. *3L: The Southeast Asian Journal of English Language Studies*, *22*(3), 65-78. https://doi.org/10.17576/3l-2016-2203-05

Ilhaamie Abdul Ghani Azmi, Rosmawani Che Hashim, & Yusmini Md Yusoff (2018). The employability skills of Malaysian university students. *International Journal of Modern Trends in Social Sciences*, *1*(3), 1–14. http://www.ijmtss.com/PDF/IJMTSS-2018-03-09-01.pdf

James, C. (2013). *Errors in language learning and use: Exploring EA*. Harlow: Addison Wesley Longman Limited.

Kim, M. (2010). Translation EA: A systemic functional approach. In C. Coffin, T. Lillis, & K. O.'Halloran (Eds.), *Applied Linguistics Methods: A Reader* (pp. 84-94). London: Rouledge.

Kong, Y. W. (2013). *A study of verb errors in written English by Chinese ESL students*. [Unpublished MA dissertation]. Universiti Malaya.

Lim, Y. M., Lee, T. H., Yap, C. S., & Ling, C. C. (2016). Employability skills, personal qualities, and early employment problems of entry-level auditors: Perspectives from employers, lecturers, auditors, and students. *Journal of Education for Business*, *91*(4), 185-192. https://doi.org/10.1080/08832323.2016.1153998

Manokaran, J., Ramalingam, C., & Adriana, K. (2013). A corpus-based study on the use of past tense auxiliary 'be' in argumentative essays of Malaysian ESL learners. *English Language Teaching, 6*(10), 111-119. https://doi.org/10.5539/elt.v6n10p111

Marlyna Maros. Tan, K. H. & Khazriyati, S. (2007). Interference in learning English: Grammatical errors in English essay writing among rural Malay secondary school students in Malaysia. *Jurnal e-Bangi*, 2(2), 1-15. http://journalarticle.ukm.my/1539/

Martin, J.R. (2013). *Systemic functional grammar: a next step into the theory – axial relations*. Beijing: Higher Education Press.

Matthiessen, C. M., & Halliday, M. A. (2009). *Systemic functional grammar: A first step into the theory*. Beijing: Higher Education Press.

McDowell, L. (2020). EA: A methodological exploration and application. In P. Clements, A. Krause, & R. Gentry (Eds.), *Teacher efficacy, learner agency*. Tokyo: JALT. https://doi.org/10.37546/JALTPCP2019-53

Nadzrah Abu Bakar, Norsimah Mat Awal, & Nor Hashimah Jalaluddin. (2011). Investigating Malay Language Writing Proficiency Level Among Upper Secondary School Students. *GEMA Online Journal of Language Studies*, *11*(2), 39-51. http://ejournal.ukm.my/gema/article/view/55/49

Rohani Salleh, Md Anwar Md Yusoff, Haryanni Harun, & Mumtaz Ali Memon. (2015). Gauging Industry's Perspectives on Soft Skills of Graduate Architects: Importance vs Satisfaction. *Global Business and Management Research: An International Journal*, *7*(2), 95-101. Available at: http://gbmrjournal.com/pdf/vol.%207%20no.%202/V7N2-12.pdf

Roslina Abdul Aziz, & Zuraidah Mohd Don. (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. https://doi.org/10.17576/gema-2019-1903-09

Satake, Y. (2020). How error types affect the accuracy of L2 error correction with corpus use. *Journal of Second Language Writing*, *50*, 1-14. https://doi.org/10.1016/j.jslw.2020.100757

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10*, 201- 231.

Siti Hamin Stapa & Mohd Mustafa Izahar. (2010). Analysis of errors in subject-verb agreement among Malaysian ESL learners. *3L: The Southeast Asian Journal of English Language Studies, 16*(1), 56-73. http://ejournals.ukm.my/3l/article/view/1005/917

Siti Zaidah Zainuddin, Pillai, S., Dumanig, F. P., & Phillip, A. (2019). English language and graduate employability. *Education + Training Journal*, *61*(1), 79-93. https://doi.org/10.1108/et-06-2017-0089

Tarone, E. (1988). *Variation in interlanguage*. London: Arnold.

Tribble, C. (2015). Teaching and language corpora: Perspectives from a personal journey. In A. Leńko-Szymańska, & A. Boulton (Eds.). Multiple affordances of language corpora for data-driven learning (pp. 37–62). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.69.03tri.

Ting, S. H., Ernisa Marzuki, Chuah, K. M., Jecky Misieng, & Jerome, C. (2017). Employers' views on the importance of English proficiency and communication skill for employability in Malaysia. *Indonesian Journal of Applied Linguistics*, *7*(2), 315-327. https://doi.org/10.17509/ijal.v7i2.8132

UCREL (n.d.). *POS CLAWS Tagger*. [online] Lancaster University. Available at: http://ucrel.lancs.ac.uk/claws/trial.html.

Wong, B. E. (2012). Acquisition of English tense and agreement morphology by L1 Malay and L1 Chinese Speakers. *3L: The Southeast Asian Journal of English Language Studies*, *18*(3), 5-14.

Xie, Q. (2019). EA and diagnosis of ESL linguistic accuracy: Construct specification and empirical validation. *Assessing Writing*, *41*, 47-62. https://doi.org/10.1016/j.asw.2019.05.002