A Corpus-Based Study on the Use of 'Be' Verbs in Indonesian ESL Students' Writing (A Case Study)

Corry Caromawati – *Institut Teknologi Nasional Bandung*

ABSTRACT

Error Analysis (EA) is beneficial for identifying the student's current language knowledge state in order to provide information to the teacher to help their students become better language users. As English verbs are considered to be one of the most challenging grammatical items to learn by non-native English speakers, some research in Malaysia have investigated EA to identify errors in the use of "verb be". This topic, however, remained untouched in Indonesia's context. Thus, this study investigated the types of errors in the use of "verb be" in the writing of 84 university students in a private university in Bandung, Indonesia. The results showed that Indonesian students tend to use "verb be" correctly when they write common phrases. Additionally, it showed that the errors tend to distributed almost evenly throughout the seven error categories. These results suggest that formulaic sequences can help English learners avoid errors and also suggest that another EA on the use of "verb be" should use a smaller number of error categories

Keywords: *Error Analysis*, *Corpus*, *EFL students' writing*, *verb be*

INTRODUCTION

It is commonly agreed that learning English as a foreign language is challenging for multilingual learners such as Indonesians. Throughout the process, errors are inevitable. Some errors are acceptable, but some might shift the attention of the readers from the actual meaning of what it is written or spoken (Shaughnessy, 1977). As a consequence, some errors can possibly cause confusion and the failure of the message delivery, particularly on written text where there is zero chance for the writers to negotiate the meaning. Studying errors, particularly made by adult language learners is also beneficial to understand the process of their language learning (Burt, 1975). In fact, the learner's errors provide important information to the teacher as they show the state of their knowledge (Brown, 2000; Wu & Garza, 2014). Therefore, errors need to be

anticipated. Considering their importance, it is not surprising if the study on errors of English learners in EFL setting have been started since decades ago. Take Corder (1971) as an example.

Shaughnessy (1997) claims that of the English grammatical system, errors are most likely found in the use of verbs. Supporting the claim, Huang (2001) put the verbs as one of the top six grammatical errors commonly found in students' writing. She claimed that these errors occurred due to the students' overgeneralization, ignorance of rule restrictions, simplifications, incomplete application of rules and L1 negative transfer.

As Malaysian and Indonesian share almost the same L1 background, i.e. Melayu, Young (2001) put them in the same group. She describes that there are some grammatical errors categories regarding verbs commonly made by Malaysian/ Indonesian students, they cover *time, tense and aspect, to be, non-finite form, modal Auxiliaries, active and passive, question forms and question tags, negatives, complex sentences, preposition, adjectives and adverbs* (pp. 286- 291). Of those errors, little research has been conducted, particularly in Indonesian setting. In the literature, Wahyudi (2012) investigated the use of subject-verb agreement in an Indonesian learner of English, while Hasyim (2002) provided a logical explanation why Error Analysis (EA) is important. In Malaysia, on the other hand, EA is more popular. It can be seen from the amount of research found in the literature. Two interesting studies are those conducted by Manokaran, et.al. (2013) and Jishvithaa, J. et al. (2013). These study investigated the use of auxiliary "be", the former focused on past tense and the later focused on present tense.

In their study, Manokaran, et.al. (2013) used WordSmith to analyze the errors and resulted in the errors classification of the past tense 'be' verbs into seven types: Tense Shift, Omission, Agreement, wrong verb Form, Addition, Misformation, and Misordering. It also provided a finding that students faced difficulties to use them in their writing. On the other hand, Jishvithaa, J. et al. (2013) investigated the types and frequency of the present simple 'be' verbs in Malaysian students' writing. They classified the errors into three categories: *tense shift, agreement,* and *misuse*. The error is regarded as a *tense shift* error category when there is a change from past to present time, or vice versa. The *agreement* category, on the other hand, involves problems in agreement between singular or plural subject with the 'be' verbs. The *misuse* category is when learners know the rules but produce something else although it is incorrect or they may have overlooked the sentence (P. 6).

This paper looks and discusses a corpus study on Indonesian students' writing on the use of present simple 'be' verbs both as copular (main) and auxiliary (helping) verb (Shaughnessy, 1977 P. 94). Auxiliary verbs are used either in present continuous tense active or in passive voice in present time. For the aforementioned purpose, this paper seek answers to these following questions:

1. What are the frequencies of the error free units of the 'be' verbs used in the students' writing?
2. How different is the distribution of error types in the students' writing regarding the use of 'be' verbs?

METHOD

This study applied corpus-based research method as it used the corpora to 'expound, test, or exemplify theories and descriptions that were formulated before large corpora became available to inform language study' (Tognini-Bonelli 2001: 65 cited in McEnery, Xiao & Tono, 2006: 8). The corpus was a selection of descriptive texts written by 84 first year students from different departments of a private university in Bandung, Indonesia, All of the participants come from different cities in Indonesia ranging from *Parahyangan Land* where people speak Sundanese, Central Java region where people speak Javanese, and some other cities with their own local languages. All learners involved have at least a local language, but they use Indonesian as their lingua franca in their daily basis.

The texts were imported to text files (.txt) in order to enable it run in the analysis tool. Before being run, the data was analyzed for the error free and erroneous units. The erroneous units were based on the use of verb be and later were classified into seven types of errors: Tense Shift, Omission, Agreement, wrong verb Form, Addition, Misformation, and Misordering. These types were used in the study conducted by Manokaran, et.al. (2013). Figure 1 summarizes the categories of errors, while the description of each category is explained afterwards.
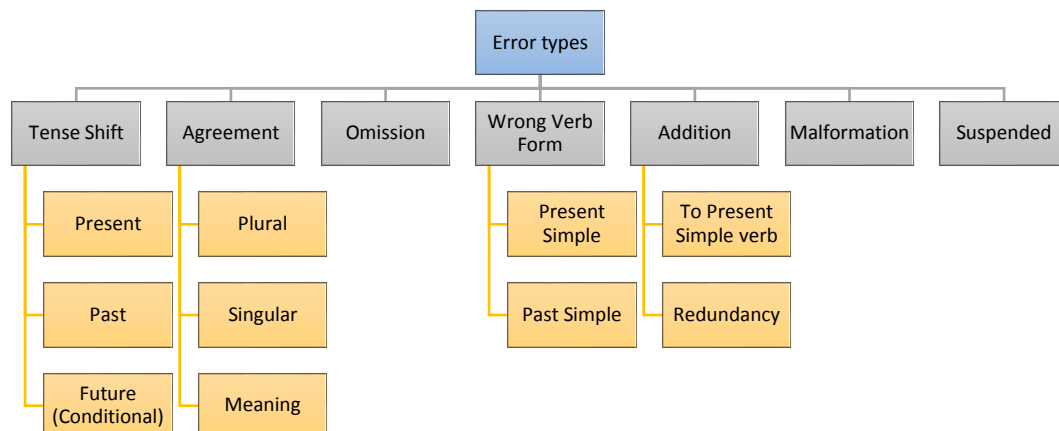
Error types

| Tense Shift | Agreement | Omission | Wrong Verb Form | Addition | Malformation | Suspended |

- Tense Shift: Present; Past; Future (Conditional)
- Agreement: Plural; Singular; Meaning
- Wrong Verb Form: Present Simple; Past Simple
- Addition: To Present Simple verb; Redundancy

Figure 1 Error type scheme

The descriptions of each category and subcategory is explained below:

1. Tense Shift

Present Time Situation

An error on this category is identified when the students used past simple 'be' verb ('was' or 'were') instead of the present 'be' verbs ('am', 'is', and 'are') for present situations.

Example from the corpus:

 I think the use Internet for learning English [GECTS] was quite good, but the internet just to read or search for material about the English language.

The correct form should be:

I think the use Internet for learning English <u>is</u> quite good, but the internet just to read or search for material about the English language.

Past Time Situation

An error is identified when the students used present 'be' verbs ('am', 'is', and 'are') instead of past simple 'be' verb ('was' or 'were') for past situations.

Example from the corpus:

The first time I learned English when I was in the third grade of elementary school, the first lesson I got was the alphabet and numbers. At that moment I was glad to learn English, because English [GECTS]is a new thing I learned.

The correct form should be:

…… At that moment I was glad to learn English, because English <u>was</u> a new thing I learned.

Future Time Situation (conditional)

An error is identified when the students used present 'be' verbs ('am', 'is', and 'are') for a first conditional form, or for something which is suggested for the future time.

Example from the corpus:

Using the Internet in the language lab [GETS][FTT]is very useful. (Note: we do not possess an internet connection yet in our lab, but we are planning to)

The correct form should be:

Using the Internet in the language lab <u>will be</u> very useful.

2. Agreement

Plural

An error is classified into this category when students failed to make the verbs agree with the subject in terms of plurality. In form 1, they used singular 'be' verbs for a plural noun as the subject.

Example from the corpus:

Yes, because the facilities [GECAG]is complete.

The correct form should be:

Yes, because the facilities <u>are</u> complete.

Singular

When the students failed to use a plural 'be' verb for a singular noun as a subject, an error is classified into this category.

Example from the corpus:

I think English lesson in ITENAS [GECAG]are good enough.

The correct form should be:

I think English lesson in ITENAS <u>is</u> good enough.

Meaning

Students failed to make the subject and the 'be' verb agree in terms of meaning so that the error causes confusion to the readers.

Example from the corpus:

I think teaching is in very good ITENAS and able to teach students…

The possible correct form should be:

I think teachers in ITENAS are very good …

3. Omission

An error is classified into this category when they failed to use a 'be' verb for a sentence requiring it.

Example from the corpus:

Introduction about my English Education [GECOM]enough.

The correct form should be:

Introduction about my English Education <u>is</u> enough.

4. Wrong Verb Form

This category classifies an error when students failed to make a correct verb form required for the sentence, but instead they used 'be' verbs.

Present Simple verb

Example from the corpus:

In this globalization era, language [GECVF]is more variety.

The correct form should be:

In this globalization era, language <u>has</u> more variety.

Past Simple Verb

Example from the corpus:

I've been learnt English since elementary, ...[GECVP] I'm not care about it.

The correct form should be:

I've been learnt English since elementary, ... I <u>did </u>not care about it.

5. Addition

An error belongs to this category when its use is an additional to what the correct form is supposed to be.

Students add a 'be' verb to a present simple predicate

Example from the corpus:

[GECAD][PSV]I'm agree.

The correct form should be:

I <u>agree</u>.

Students add a 'be' verb before a present or past participle as a post modifier of a noun.

Example from the corpus:

Everyone [GEAAD][PP]is taking an aducation course have a goals and expectation,as well as me.

The correct form should be:

Everyone <u>taking</u> an aducation course have a goals and expectation,as well as me.

6. Malformation

This category is used when students failed to use a proper punctuation such as additional coma in between the 'be' verb and the rest of the sentence, and the omission of apostrophe.

Example from the corpus:

..., also [GECMF]its my fault cause ...

The correct form should be:

..., also it's my fault cause ...

7. Suspended

This category is used for a verb 'be' used without a subject.

Example from the corpus:

... [GECSP]is better if not often using internet.

The correct form should be:

... <u>it</u> is better ...

In identifying and classifying the errors, a code was inserted in the text, some have more than one codes when there were more than one possible problems. The list of the codes used for this study can be found in Appendix 1. LeCompte and Schensul (1999: 55) in Ellis (2005: 253) define *codes* as names or symbols use to stand for a group of similar items, ideas, or phenomena that the researcher has noticed in his or her data set'. The annotated corpora were afterwards imported and analyzed on *AntConc* by making the concordances (see Figure 2 as an analysis illustration). The system itself, according to Granger, et al. (1998) is hierarchical, consisting one major category code and a series of subcodes. Therefore, the codes applied to tag the errors in the corpus stand for the main category and subcategory. The results of the analysis were used as the answers to the research questions of this study.
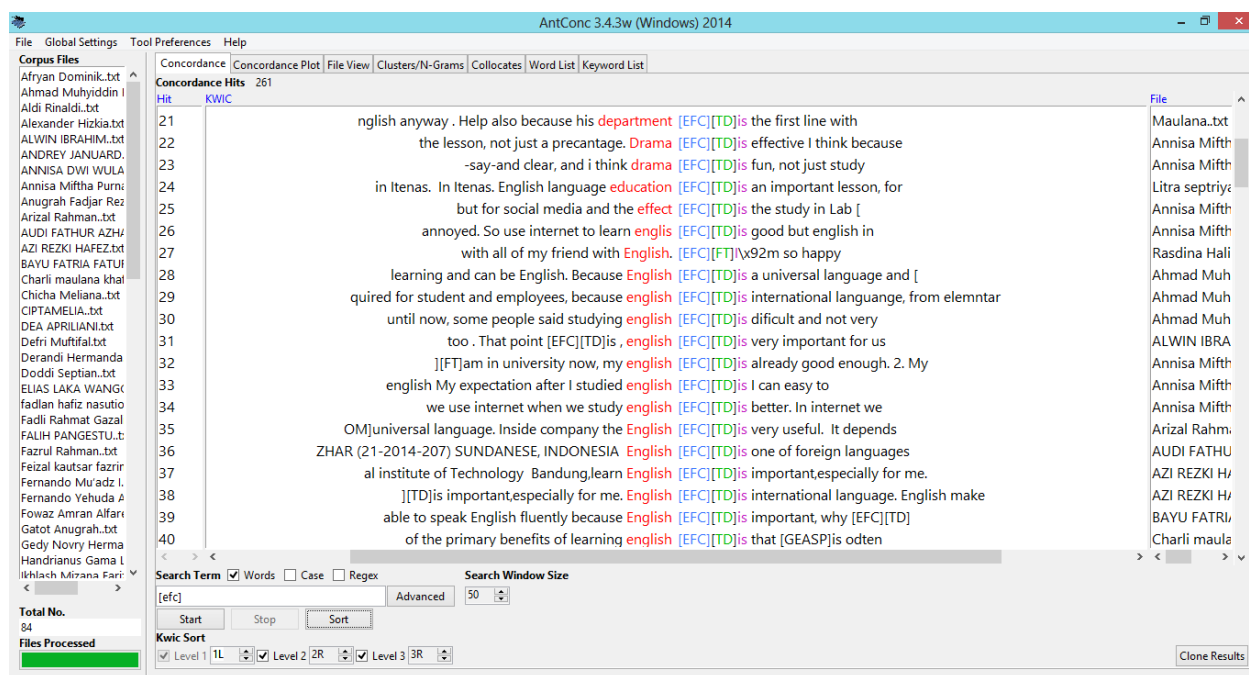
Figure 2. The analysis run through AntConc

RESULTS AND DISCUSSION

The 84 students wrote different numbers of words with all total sum up to 20,815 tokens. The attempts of the present simple 'be' verbs (am, is, and are), error free tags of copular and auxiliary verb "be" were calculated. The results are shown in Table 1 below.

Table 1. The attempts of the 'be' verbs and error free tags

	Σ attempts	Error Free Copular	Error Free Auxiliary	Σ error free	%
Am	45	18	4	22	49%
Is	304	212	13	225	74%
Are	69	29	9	38	55%
Total	418	259	26	285	68%

Of the 20,815 tokens, students merely attempted to use the target verbs approximately 418 times (2%). 261 concordance lines are formed on AntCons, and they show that: the highest

frequency of the 'be' verb as used goes to 'is' which students use to successfully to complement English (language), internet, my hope, it, the lecturer, my name, my opinion, there, and one of them as the subjects of the sentences they made (see *appendix 1*). The attempts on using 'is' was identified in 304 occurrences, which is the most frequent among the three verbs. Of the attempts, 74% of them are error free. The verb 'am' was tried to use in the frequency of 45, and 49% are error free. The verb 'are', on the other hand, were attempted 69 times, and 68% of them are considered successful. This result is quite astonishing even though the learners use them in relatively simple sentences. Additionally, from the table above, it is obvious that learners tend to succeed using the verbs as copular instead of as auxiliaries, which resulted in 26 concordance lines, since they might be more familiar how to use them with some adjectives and nouns than when combined with participles (past and present). Consequently, teachers have a challenge to help students use get familiar with and use passive voice in their sentences.

To answer the second RQ, the analysis from the corpus error tagging process created and resulted an error scheme in the discovery of the types of errors as follow:

Table 2. The results of analysis

		Verb 'be' Present Simple						
		Copular	%	Auxiliary	%	Total	Σ	%
Tense Shift	Past	55	67%	5	6%	60	82	29%
	Present	11	13%	9	11%	20		10%
	Future	2	2%	0	0%	2		1%
Agreement	Singular	4	14%	6	21%	10	28	5%
	Plural	11	39%	4	14%	15		7%
	Meaning	3	11%	0	0	3		1%
Omission		6	0.75	2	0.25	8	8	4%
Wrong Verb								
Form	Present Simple Verb	12	31%	14	36%	26	39	12%
	Past Simple Verb	2	5%	6	15%	8		4%
	Future Tense Verb	1	3%	1	3%	2		1%
	Voice	0	0%	3	8%	3		1%
Addition	Present Simple	3	43%	4	22%	7	18	3%
	Participles	2	11%	5	28%	7		3%
	Redundant	3	17%	1	6%	4		2%
Malformation		11	69%	5	31%	16	16	8%
Suspended		17	89%	2	11%	19	19	9%
TOTAL		143		67		210	210	

100%

Based on the data shown in Table 2, the errors regarding the shift of time occurred more frequently than other errors, it makes up 29% of the total. Students apparently failed to use 'be' verbs in past time aspect (was and were) but using the present forms instead. Additionally, they also used the past 'be' verbs for the present situation (10%), but only a few failed to use first conditional form (1%) since there were not too many attempts done to use this form.

The second error type students made is constructing the right verb forms. Present simple is proved to have the most frequent erroneous forms (12%) involving the 'be' verbs. Students

sometimes used a 'be' verb before a base form instead of the correct present simple verb, or used the base form of 'be' verb itself instead of the present simple forms (is, am, are). Past simple (4%), future (1%), and voice (active and passive) (1%) are other errors students performed involving the target verbs; they occasionally used the 'be' verbs as in passive, but present participle was used in the sentences that made them have active aspect.

Error in agreement is the third most frequent problematic category found in the students writing that makes up to 13% in total. The students' failure to make the 'be' verbs agree with the subject is not only regarding singularity (5%) and plurality (7%), but also match the meaning between the subject and the verb itself (1%) (*see* the concordances).

The suspending 'be' verbs found in the corpus interestingly reached 9% from the total errors. Students failed to have a clear subject that would go them. They might consider the previous noun which is object of preposition or adverb of time as the subject of sentences. Additionally, some erroneous sentences sound highly influenced by Indonesian language. For example, the sentence "*I think [GECSP]is very helpful for us to learn english language more deeper .*" would sound make more sense in Indonesian language "*Saya rasa sangat membantu bagi kami untuk belajar bahasa Inggris lebih dalam.*"

The next fault is under malformation category which reached approximately 8%. This type is closely related to the use of punctuation, i.e. coma and apostrophe. Students did not put an apostrophe in between the subject and the contracted 'be' verb, or they put a coma after the verb itself which blocked it with the rest of the sentence. The later form is, I believe, also strongly influenced by students' lingua franca (Indonesian language). Students tried to translate literally their mother tongue into English. For instance, "*Our hope in English [GECMF]is, we can understand English more, especially in speaking.*" Which is literally derived from "*Harapan kami dalam bahasa Inggris adalah, kami dapat menegrti bahasa Inggris lebih banyak, terutama kemampuan berbicara.*"

Erroneous sentences also occurred when students use an unnecessary 'be' verb (8%) which if dropped, would have made a better sentence. They used it before a present simple verb (3%), participle as post-modifier (3%), and redundancy (2%).

The last type of errors discovered in the corpus is omission (4%). The omission occurred more in the 'be' verbs as copular (75%) than as auxiliary (25%). This fact is somewhat worrying because 'be' verbs as copular is considered as a grammatical item taught for students in the beginner level and has been introduced since students studied in primary school. Thus, it is going to be a 'challenge' for the teachers to 're-teach' this grammar point in the university level.

English teachers in EFL setting, particularly in Indonesia, always have challenges to improve students' writing performance, especially to apply grammar in use. As writing is very essential in a language learning process, observation and analyses on students' writing are needed to reveal the kinds of errors students' make and the possible factors that might cause them. This hopefully will enable teachers to find better ways to anticipate and treat them to make some progress in teaching writing.

The grammatical errors in the use of present simple 'be' verbs found in the students' writing should be considered as a problem. Therefore, it is important for teachers to teach the grammatical items properly and give feedback when encountering the errors effectively so that students would be able to correct and improve their writing. If teachers can give clearer explanation in their teaching and feedback on the types of errors students' make, the learning and teaching process would be developed to give better result on students' work and English proficiency. Additionally, it is not only students who will get the benefit, teachers and the management would get the value of it as well. Together we will be able to construct better syllabus and develop more appropriate teaching materials. Broadly, the findings of this study might also be able to be used as reference for in-service teachers who teach Indonesian students and encounter the same problem, or pre-service teachers who would like to anticipate problems in teaching the target 'be' verbs to Indonesian learners.

CONCLUSION

This study has proved that Indonesian students, in lower proficiency level, have some problems dealing with the use of the 'be' verbs. In addition to being used in simple sentences, errors are found in different types: *tense shift (past, present, and future), agreement (plural, singular, and meaning), omission, wrong verb form (present simple and past simple), addition (to present simple verb and redundancy), malformation,* and *suspended.* Some types are recognized to be common errors English learners mentioned in the previous studies while *suspended* and *malformation* are

considered strongly influenced by students' native language, Indonesian. This study can provide a clear way for teachers who teach Indonesian students to teach the 'be' verbs in present simple form.

Appendix 1

The description of the codes can be described as the following:

Error Free	Tense shift	Agreement
Main Categories:	Main categories:	Main categories:
EFC = Error Free Copular Verb	GECTS = Grammar Error Copular Tense Shift for the 'be' verb used as a copular verb.	GECAG= Grammatical Error Copular Agreement
EFA = Error Free Auxiliary Verb	GEATS = Grammar Error Auxiliary Tense Shift for the 'be' verb used as an auxiliary verb.	GEAAG = Grammatical Error Auxiliary Agreement
Subcategories:	Subcategories:	Subcategories:
FT = First ('am')	PTT = Past time situation	SG = When the subject is singular
TD = Third (is)	PRT = Present time situation	PL = when the subject requires a plural verb
ST = Second and Third (are)	FT = Future time Situation	MNG = When there is a mismatch between the verb and the subject in terms of meaning
Omission	**Wrong Verb Form**	**Addition**

GECOM = Grammatical Error Copular Omission

Main categories:

Main categories:

GEAOM = Grammatical Error Auxiliary Omission

GECVF = Grammatical Error Copular Verb Form

GECAD = Grammatical Error Copular Addition

GEAVF = Grammatical Error Auxiliary Verb Form

GEAAD = Grammatical Error Auxiliary Addition

Subcategories:

Subcategories:

PSV = Present simple verb (third person singular)

PSV = Present Simple Verb

PTV = Past tense verb form

PP = Past or present participle as a post modifier

FTV = Future tense verb form

RDNT = Redundant

VC = voice (active or passive)

Suspended	Malformation
GECSP = Grammatical Error Copular Suspended	GECMF = Grammatical Error Copular Malformation
GEASP = Grammatical Error Auxiliary Suspended	GEAMF = Grammatical Error Auxiliary Malformation

REFERENCES

Ahour, T., & Mukundan, J. (2012). *Errors and Variations of TESL Students' Written Description*. Pertanika Journal of Social Sciences & Humanities, 20(1), 55.

Burt, M. K. (1975). Error analysis in the adult EFL classroom. *TESOL quarterly*, 53-63.

Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *IRAL-International Review of Applied Linguistics in Language Teaching*, *9*(2), 147-160.

Corder, S. P. (1981). *Error Analysis and Interlanguage*. Oxford: Oxford University Press.

Ellis, R. & Barkhuizen, G. (2005). *Analysing Learner Language.* Oxford University Press

Hasyim, S. (2002). Error analysis in the teaching of English. *kota lama*, *4*(1), 62-74.

Hawkins, R., & Chan, Y. H. (1997). *The partial availability of Universal Grammar in second language acquisition: The failed functional features hypothesis*. Second Language Research, 13(3), 187-226.

Heift, T. and Schulze, M. (2007). *Errors and Intelligence in Computer-assisted Language Learning*. Taylor & Francis Routledge

Manokaran, et.al. (2013). *A Corpus-Based Study on the Use of Past Tense Auxiliary 'Be' in Argumentative Essays of Malaysian ESL Learners*. English Language Teaching; Vol. 6, No. 10.

Muftah, M., & Rafic-Galea, S. (2013). *Error Analysis of Present Simple Tense in the Interlanguage of Adult Arab English Language Learners*. English Language Teaching, 6(2), 146-154. http://dx.doi.org/10.5539/elt.v6n2p146

Shaughnessy, Mina (1977). *Errors and Expectations.* New York. Oxford University Press

A learner corpus-based study on error associations1

Swan, M. and Bernard, S. (2001). *Learner English (second Edition)*. UK. Cambridge University Press

Jishvithaa, J. et al (2013). *Teaching Grammar: The Use Of The English Auxiliary "BE" Present Tense Verb Among Malaysian Form 4 And Form5 Students*. Advances in Language and Literary Studies Vol. 4 No. 2; July 2013

Wahyudi, R. (2012). Error analysis on subject-verb agreement: the case of a university student in Indonesia. *Research on Humanities and Social Sciences*, *2*(4), 20-25.