

Exploring Grammatical Structure: A Corpus Analysis on the Use of “To Be” in Vocational Students’ WhatsApp Group Conversations using AntConc

**Muhammad Arba'in¹, Itsna Millatul Himmayati², Dwi Hastuti³,
Desi Nosita⁴, Hanung Triyoko⁵**

MA Al-Fathimiyah, Lamongan, Jawa Timur, Indonesia¹

SMP Pesantren Afaada, Boyolali, Jawa Tengah, Indonesia²

SMK N 3 Salatiga, Salatiga, Jawa Tengah, Indonesia³

SD Tahfizhul Qur'an As Surkati, Salatiga, Jawa Tengah, Indonesia²

UIN Salatiga, Salatiga, Jawa Tengah, Indonesia²

arbainforty40@gmail.com¹, itsnamilla11@gmail.com², jffausta@gmail.com³,

nositadesi@gmail.com⁴, hanungina@uinsalatiga.ac.id⁴

Abstract: The study addresses issues related to the usage of “to be” in unstructured written conversations among vocational high school students through the WhatsApp group platform. The focus is an analysis using AntConc to understand conversational patterns involving “to be”. The research aims to identify sentence variations and common errors among the students. The source of evidence is the corpus of unstructured written conversations among the students on the WhatsApp group platform. The analysis involves employing a descriptive method from corpus linguistic research to understand sentence variations and detect common errors related to the usage of “to be”. The students employ the use of “to be” in several variations of grammar structure such as nominal sentence, passive sentence, progressive tense and be going to. However, observations also reveal some errors such as omission, misordering, and miss formation. The findings suggest a need for more in-depth instruction regarding the placement of the verb “to be” in sentences. It would include insights into conversational patterns and the prevalence of errors. The implications of the study highlight the importance of more intensive lessons for students. The recommendation is that teachers should provide more intensive instruction in these areas.

Keywords: Grammatical structure; Corpus analysis; Vocational students; WhatsApp conversation; AntConc.

Mengeksplorasi Struktur Tata Bahasa: Analisis Korpus Penggunaan “To Be” dalam Percakapan Grup WhatsApp Siswa SMK menggunakan AntConc

Abstrak: Penelitian ini membahas isu-isu yang berkaitan dengan penggunaan “to be” dalam percakapan tertulis yang tidak terstruktur di antara siswa SMK melalui platform grup WhatsApp. Fokusnya adalah analisis menggunakan AntConc untuk memahami pola percakapan yang melibatkan “to be”. Penelitian ini bertujuan untuk mengidentifikasi variasi kalimat dan kesalahan umum di antara siswa. Sumber datanya adalah korpus percakapan tertulis yang tidak terstruktur di antara siswa di platform grup WhatsApp. Analisis melibatkan penggunaan metode deskriptif dari penelitian linguistik korpus untuk memahami variasi kalimat dan mendeteksi kesalahan umum yang terkait dengan penggunaan “to be”. Para siswa menggunakan “to be” dalam beberapa variasi struktur tata bahasa seperti kalimat nominal, kalimat pasif, bentuk progresif dan be going to. Namun, pengamatan juga mengungkapkan beberapa kesalahan seperti kelalaian, kesalahan urutan, dan formasi yang hilang. Temuan ini menunjukkan perlunya instruksi yang lebih mendalam mengenai penempatan kata kerja “to be” dalam kalimat. Ini akan mencakup wawasan tentang pola percakapan dan prevalensi kesalahan. Implikasi dari penelitian ini menyoroti pentingnya pelajaran yang lebih intensif bagi siswa. Rekomendasinya adalah bahwa guru harus memberikan pengajaran yang lebih intensif di bidang-bidang ini.

Kata kunci: Struktur tata bahasa; Analisis korpus; siswa SMK; percakapan WhatsApp; AntConc.

1. Introduction

In the field of language usage, digital communication platforms have become a significant arena for linguistic exploration and analysis. It includes verbal communication that is essential in human interactions, as people participate in speaking activities in various situations and at any time (Suryadi et al., 2024). The study delves into grammatical structures, specifically focusing on the use of "to be" in the context of vocational student interactions on WhatsApp, a widely used platform for informal discourse among diverse demographics (Lee, 2023).

This research begins an exploration of grammatical nuances evident in WhatsApp group conversations of vocational students. The focal point revolves around "to be" which consists of be, is, am, are, was, were and been. It is a fundamental element in linguistic structure, to distinguish frequency, contextual application, and syntactic variation in this specific demographic (Ibrohimova & Valiyev, 2019).

The use of AntConc, a powerful corpus analysis tool developed by Laurence Anthony, the study aimed to dissect the nuances of language use embedded in these digital interactions (Cahyati et al., 2021). By harnessing the linguistic abilities of the corpus, we sought to uncover the intricacies and underlying grammatical patterns manifested in the spread of "to be" by vocational students on WhatsApp.

The findings of this investigation hold promise in offering valuable insights not only into the linguistic tendencies of vocational students in informal digital settings but also their language acquisition processes. In addition, the implications derived from this analysis have the potential to inform tailored pedagogical approaches to align language teaching methodologies with students' informal digital language practices, thereby encouraging improved language proficiency in vocational contexts.

2. Materials and Methods

The methods should be provided in sufficient detail to allow other authors to replicate or reproduce them. The methods section may include the research design and setting of the study; the participants' characteristics (population and sample); the description of the materials (research instruments); the description of all processes and methodologies used (data collection procedures); and the data analysis. This study adopts a qualitative approach with a focus on qualitative descriptive corpus linguistics

research (Hasko, 2012). It is a methodology for conducting in-depth investigations of linguistic phenomena, grounded in the context of natural language use.

The goal of qualitative corpus analysis is to explore the meanings and functions of linguistic forms found in a specific corpus and how they interact with diverse ecological characteristics of language used (Hasko, 2012). Corpus linguistics is an empirical approach that investigates language use in its natural context with different types of corpora as its database (Voelkel & Kretzschmar, 2021). The methodological issues include considerations on corpus linguistic approaches, types and criteria of corpora, steps of corpus analysis, such as tokenization and tagging, and finally types of analysis.

2.1. Data Collection

In the beginning, the researchers collected a corpus of WhatsApp group conversations involving vocational students. It involved the students requesting their consent to use their chat logs as research data. The researchers also considered using publicly available conversations or those obtained with proper permissions. Given the potential volume of data, the researcher sampled the corpus to make analysis manageable. The researchers sampled conversations from different time periods of vocational courses to ensure representativeness. It was the conversation from September 26th 2023 to October 16th 2023.

2.2. Data Analysis

After that, the researchers did the pre-processing. It involved cleaning the data by removing non-textual elements (e.g., name, contact, images, etc) and tokenizing the text into words or units. The researchers have obtained informed consent from participants and are handling the data in an ethical manner, respecting privacy and confidentiality.

In qualitative descriptive corpus linguistics research, the primary goal is to provide a detailed and rich description of linguistic patterns and features in the corpus. Researchers engage in a close reading of the texts, identifying recurrent themes, linguistic constructions, or patterns. The researchers used qualitative coding techniques to categorise and label linguistic elements of interest. This may involve coding for different forms of "to be" (e.g., am, is, are, was, were and been) and the context in which they are used.

Understanding the context in which "to be" is used is crucial. The researchers then examine the surrounding text to determine the role and function of "to be" in various conversations. It included analysing whether it's used in a copular

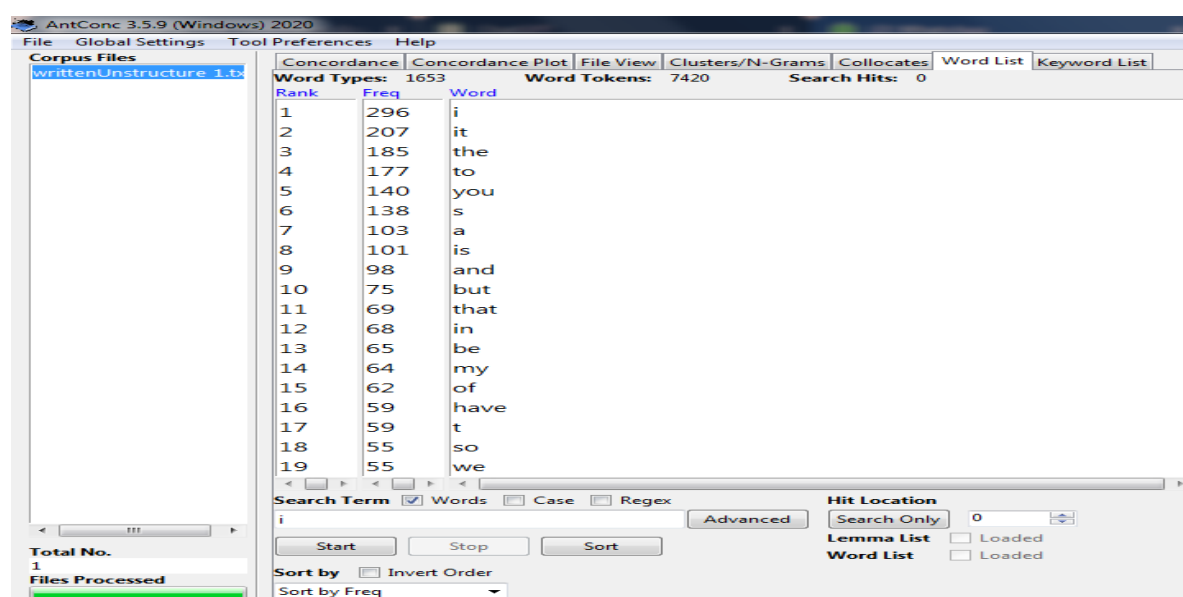
sense, as an auxiliary verb, or in other linguistic functions. After coding and analysing the data, the researchers interpreted the findings. Explain the significance of the identified patterns and how they reflect the grammatical structure in WhatsApp group conversations among vocational students. The researchers also discuss any errors from Standard English grammar.

3. Result and Discussion

3.1. Type and Token

Understanding the concepts of "type" and "token" in language analysis is essential. With a grasp of both concepts, we can conduct a more in-depth analysis regarding the diversity of vocabulary and the frequency of word

occurrences in a text. A token is like a single word or part of a language in a text (Hussein, 2018). The number of tokens in a computerised database is how many words there are in total. On the other hand, the number of types is how many different and unique words there are (Hussein, 2018). So, a token is any language thing that shows up in a text, no matter what kind it is. But a type is a maths idea that looks at the special words in a studied group of words. If we compare how many tokens there are in the data or group of words to how many types there are, we can learn about how many different words are used in that group of words. The researchers show the result of data about type and token on the Picture 1 below:



Picture 1. Corpus Analysis of Type and Token

From Picture 1, it can be seen that the corpus analysis on the use of "to be" in vocational students' WhatsApp group conversations, conducted using AntConc, reveals valuable insights. The corpus comprises 1653 unique word types and a total of 7420 tokens. Among the frequently occurring words, the top-ranking word is "I" with a frequency of 296, indicating a significant presence of self-reference in the conversations. The second position is held by the word "It" with a frequency of 207, suggesting a focus on various entities or objects in the discussions. Following closely in the third position is the word "The," occurring 185 times, possibly indicating the introduction or discussion of specific topics or subjects within the group.

This distribution of word frequencies provides a preliminary understanding of the linguistic patterns and content emphasis within the vocational students' WhatsApp group. The

prominence of personal pronouns like "I" suggests a high level of individual engagement or self-expression. The frequent use of "It" implies a considerable discussion around different objects or concepts, while the recurrence of "The" indicates a recurring focus on specific subjects or topics of conversation. Further analysis could delve into the context of these frequently used words, shedding light on the thematic elements and communication dynamics within the group

3.2. Word List of Be

The word list from the AntConc application is a list of words that appear in a text corpus along with their frequencies of occurrence. In the context of language or text analysis, a word list can provide information about which words are most frequently used or dominant in a document or group of documents. The researchers show the result of data about word list of "be" on the Picture 2. below:

Picture 2. Word List of “be”

AntConc 3.5.9 (Windows) 2020			
File Global Settings Tool Preferences Help			
Corpus Files	Concordance	Concordance Plot	File View Clusters/N-Grams Collocates Word List Keyword List
writtenUnstructure 1.txt	Word Types: 7 Word Tokens: 249 Search Hits: 0		
Rank	Freq	Word	
1	101	is	
2	65	be	
3	38	are	
4	26	was	
5	11	am	
6	5	were	
7	3	been	

Based on Picture 2, it can be said that the word "is" dominates the usage in the corpus with a frequency of 101 times. Followed by "be" with a frequency of 65, and "are" with a frequency of 38. Furthermore, "was" appears in fourth place with a frequency of 26, followed by "am" in fifth place with a frequency of 11. "Were" and "been" each appear in sixth and seventh place with frequencies of 5 and 3. This provides a brief overview of the distribution of "be" usage in the corpus.

3.3. Is

The word "is" is a form of the verb "to be" used in the context of third person singular in English. The use of "is" specifically relates to singular subjects, such as he, she, it, or objects considered singular (Lubis et al., 2022). It is also used as a progressive form, such as the word "is," to convey that an action is happening at a specific time or in the process. Typically, this form is combined with verbs ending in -ing. This

construction is employed to discuss actions and conduct, but it is not commonly used to express emotions.

In examining this corpus, the term "is" within a WhatsApp group of high school students' conversations reveal interesting multiple purposes, and its application across different situations yields valuable insights into the wide range of expressions and sentence constructions prevalent in day-to-day communication.

In corpus linguistics, the word "is" is a form of the verb "to be" and is often referred to as a copula or linking verb. The usage of "is" in corpus linguistic studies involves analysing its frequency, collocations, and patterns within a large collection of texts or a corpus. In the provided corpus linguistic analysis, the data includes concordance lines, the type of structure, and the frequency of specific patterns involving the word "is". The result about the word "is" can be seen in the Table 1. below:

Table 1. The Result of Corpus Analysis of “is”

No	Concordance	Kinds Structure	of Subject Categorization	Grammatical Acceptance		Freq
				Yes	No	
1.	sambal roa is a chili sauce made from smoked flying fish	Nominal	<ul style="list-style-type: none"> Single Noun (58) Noun Phrase (28) 	27	2	86
2.	Simply, I think it is based on human's behaviour...	Passive	<ul style="list-style-type: none"> Single Noun (4) Noun Phrase (5) 	8	1	9
3.	Sharing is caring?? Let's share our experience!!!	Present progressive	<ul style="list-style-type: none"> Single Noun (4) Gerund (2) 	5	1	6
Total						101

After find the number of the use “is”, then the researcher did the error analysis of the use of “is” that can be seen in the Table 2. below:

Table 2. Kinds of Error in Corpus Analysis of “is”

No	Kinds of Error	Concordance
1.	Addition	<ul style="list-style-type: none"> Spices is always be good food.
2.	Misformation	<ul style="list-style-type: none"> Spices is always be good food. Gambling is always have bad effect actually. I think, accent is depend on how often they speak english.
3.	Misordering	<ul style="list-style-type: none"> Be a pilot is my dream when i was 10 years old.

The phrase “is always be,” is a misformation and addition. Corrections are made to address two errors. Firstly, the verb “is” is inaccurately used and is replaced with the plural form “are” to match the subject “spices.” Secondly, the unnecessary word “be” after “always” is eliminated to enhance the sentence structure, resulting in the more accurate statement, “Spices are always good food.”

Moving on to the next sentence, “Gambling is always have bad effects actually,” belongs to misformation. The error in the verb choice is identified, and the correct form “has” is substituted for “have.” The revised sentence reads, “Gambling always has bad effects actually.” The optional use of “actually” is noted as potentially unnecessary, adding further clarity.

In the third sentence, “I think, accent is dependent on how often they speak English,” a misformation is identified, leading to the replacement of “depend” with “depends” to match the singular subject “accent.” The refined sentence is, “I think, accent depends on how often they speak English,” and the use of a comma after “I think” is deemed optional for increased conciseness.

Lastly, the sentence “Being a pilot was my dream when I was 10 years old,” involves the category of misordering. An error in sentence structure is addressed by expanding the verb “Be” into a complete clause, resulting in the more accurate statement, “Being a pilot was my dream when I was 10 years old.” These corrections collectively emphasize the significance of precise verb usage, appropriate sentence structures, and the thoughtful incorporation of language elements to enhance overall communication clarity.

3.4. Be

In this corpus analysis, the usage of the word “be” in unstructured written conversations from a high school student WhatsApp group demonstrates intriguing syntactic and semantic variations. The English word “be” holds diverse functions, and its usage in various contexts provides profound insights into the array of expressions and sentence structures in everyday communication (Harahap, 2020). The result about the word “be” can be seen in the Table 3. below:

Table 3. The Result of Corpus Analysis of “be”

No	Concordance	Kinds of Structure	Subject Categorization	Grammatical Acceptance		Freq
				Yes	No	
1	with my dad But now i just wanna be a millionaire???? Hi guys... what are you doing?	Verbal	<ul style="list-style-type: none"> Single noun (12) 	10	1	12
2	Scorpio, Nov Scorpio? Yes. But sometimes I can be a person who is indifferent and doesn't glish-british-dan-english-american/ https://youtu.be/66aG5P0kQpU?feature=shared Try to watch	Nominal	<ul style="list-style-type: none"> Single noun (43) Noun Phrase (1) 	43	2	44
3	hole not a door and the hole will be blocked by cement?? I see... Is it Mushroom?? ??	-	-	-	1	1
4		Passive	<ul style="list-style-type: none"> Single noun (5) Noun Phrase (3) 	8	0	8
Total						65

After find the number of the use “be”, then the researcher did the error analysis of the use of “be” that can be seen in the Table 4. below:

Table 4. Kinds of Error in Corpus Analysis of “Be”

No	Kinds of Error	Concordance
1	Misordering	glish-british-dan-english-american/https://youtube/66aG5P0kQpU?feature=shared Try to watch
2	Misordering	used to break it down?? yeah, it could be.. but there's another answer actually Heart Try
3	Misordering	for you to practice english to make you be able to speak english fluently?
4	Misordering and addition	Spices is always be good food?

The first concordance, “glish-british-dan-english-american/https://youtube/66aG5P0kQpU?feature=shared Try to watch”, is not a sentence. It is a link to a youtube video and it cannot be categorised as a sentence. It does not fulfil the minimum structure of a sentence that contains the subject and predicate or verb.

The sentence "used to break it down?? yeah, it could be. but there's another answer actually "Heart Try" includes a Misordering error. The structure seems a bit confusing, with the words "Heart Try" sounding like a separate phrase not directly related to the preceding sentence. There might be a deficiency in arranging the words that makes it challenging to understand clearly.

The third grammatical error is from “for you to practise english to make you be able to speak english fluently?”. The concordance consists of phrases only with four prepositional phrases “for you”, “to practise english”, “to make you be able”, “to speak English”. The sentence misses its subject and verb.

The fourth grammatical error from a question sentence. The concordance “Spices is

always be good food?” is ended by a question mark. It should be started with “is” so that the concordance could be “is spices always good food?”. Beside that there are two to be in that sentence. The to be are “is” and “be”. In English grammatical rules, it is not allowed. Therefore “be” should be omitted from the sentence.

3.5. Are

In the rules of English grammar, “are” is the plural form of the verb to be” (Ismi et al., 2021). The word 'are' is used in the present progressive for the subjects 'you,' 'we,' and 'they' as a grammatical element indicating that an activity or event is currently taking place at the time of speaking (Gu, 2020). It is used in the context of clauses or sentences that indicate the existence, state, or relationship of an object or subject in the present time. For example, in the sentence “They are students”, “are” indicates that “they” are students or currently in the state of being students. The result about the word “are” can be seen in the Table 5. below:

Table 5. the Result of Corpus Analysis of “are”

No	Concordance	Kinds of Structure	Subject Categorization	Grammatical Acceptance		Freq
				Yes	No	
1.	They are master students in UIN Salatiga	Nominal	<ul style="list-style-type: none"> Noun (9) Pronoun (18) Conjunction (2) 	27	2	29
2.	They are affecting people mind to do sin ??	Present progressive	Pronoun (4)	4	-	4
3	Strange natural disasters are also known to occur on the 13 th	Passive	Noun (2)	2	-	2
4.	What exactly are we going to do in here?	Be going to	Pronoun (2)	2	-	2
5.	when i was a little Well, most dads are like that, right?	Be like that	Noun (1)	1	-	1
Total				38		

The researchers' dataset includes 29 sentences, with a prime example such as “They are master students at UIN Salatiga”, demonstrating the prevalence of nominal

sentences. In this context, the use of “are” provides clarity in conveying information about the identity or attributes of a group of individuals (Syafar, 2014). When delving deeper into the

analysis of subject types in these nominal sentences, we can observe a rich variety. Nouns emerge as subjects 9 times, pronouns dominate with 18 occurrences, and conjunctions connect thoughts in 2 cases. This analysis not only reveals the frequency of sentence forms but also offers insights into the distribution of subject types within the corpus under examination.

Next, the application of “are” in the present progressive reflects situations or events unfolding during the conversation (Ismi et al., 2021). For instance, the sentence “They are affecting people’s minds to do sin?” illustrates an ongoing action, highlighting events that remain relevant at the time of speaking. This showcases an awareness of time in verb usage, enriching the expression of ideas or occurrences. In this context, there are four nominal sentences that appear, and all of their subjects are in the form of pronouns.

Transitioning to passive voice sentences indicates a tendency to spotlight the object receiving the action. The use of the verb ‘be,’ followed by the past participle, constitutes the formation of passive sentence structures (Malabar & Mustapa, 2022). An example sentence like “Strange natural disasters are also known to occur on the 13th” portrays the use of “are” as part of a passive construction, shifting the focus onto the event itself rather than the doer. This introduces structural variation in sentences, adding complexity and diversity to expression. In

this particular context, two passive sentences are evident, and the entirety of their subjects assumes the structure of nouns.

Furthermore, the use of “be going to” presents an aspect of future planning or intention. Sentences such as “What exactly are we going to do in here?” demonstrate the use of “are” in the context of future plans or objectives. This adds a temporal dimension to communication, allowing speakers to plan and organize actions in the future. In this particular context, the manifestation of two instances involving the “be going to” form is evident, with all respective subjects adopting the pronoun form.

Lastly, in the expression “be like that”, “are” is used to state general characteristics of a group of people. The sentence “Well, most dads are like that, right?” illustrates the use of “are” to describe common attributes attributed to a group of individuals. This creates a sense of generalization involving the speaker and the listener in a shared understanding of specific characteristics.

Overall, this corpus analysis reveals that the use of the word “are” in high school student conversations encompasses diverse contexts, from situational explanations to the depiction of general characteristics. This variation reflects the richness of English language structures and semantics in the context of everyday communication. The error analysis of the use of “are” that can be seen in the Table 6. below:

Table 6. Kinds of Error in Corpus Analysis of “are”

No	Kinds of Error	Concordance
1.	Omission	There are several similar thing
2	Misordering	Are you belong to Gryffindor’s student?

Based on the analysis presented in Table 6, two grammatical errors can be identified in the nominal forms of the example sentences. The first error is evident in the sentence “There are several similar thing”, where the use of the noun “thing” does not adhere to proper grammatical rules and should be replaced with the plural form “things”. Such grammatical errors are specifically referred to as “omission”, indicating the omission or oversight of elements that should be present in a sentence. In this context, omission manifests when the plural form is not used correctly, resulting in grammatical errors that can impact the clarity and accuracy of communication.

The second error identified related to the sentence “Are you belong to Gryffindor’s student?” can be categorized as “misordering” or incorrect sequencing. Misordering itself refers to the improper arrangement or sequencing of elements within a sentence, resulting in a

sentence structure that deviates from the grammatical norms of a language. In this example, misordering occurs because the verb “belong” is placed incorrectly in relation to the subject, affecting the syntactic integrity of the sentence. The focus of this research is on the use of “be”, so the correct sentence should not include the phrase “belong to” but rather be simply “Gryffindor’s student”. Therefore, the correct sentence is “Are you Gryffindor’s student?”.

3.6. Was

The verb “was” is the past tense form of the verb “to be”. The verb “was” is used to describe an action or state of being that occurred in the past. It is used when the subject is singular (I, he, she, it) or when the subject is a singular noun. “Was” is used for the first and third persons singular, and were is used otherwise. For example, “I was happy” or “She was at the store.” (Murphy, 2000).

In the corpus analysis of the use of the word “was” in unstructured written conversations from a high school student WhatsApp group, two main patterns reflecting variations in the usage of this word are observed. The result about the word “was” can be seen in the Table 7. below:

Table 7. The Result of Corpus Analysis of “was”

No	Concordance	Kinds of Structure	Subject Categorization	Grammatical Acceptance		Freq
				Yes	No	
1	When I was a child, the most horror movie was Friday the 13	Nominal	Pronoun (15) Noun (5)	20	-	20
2	The latest was released in 2009 right?	Passive voice	Pronoun Noun	5	1	6
Total						26

The sentence "When I was a child, the most horror movie was Friday the 13th" is an example of a nominal sentence in the past tense. In this sentence, the subject is "the most horror movie," and the verb is "was." The sentence is in the past tense because it describes a past event or state of being. The sentence also includes a subordinate clause, "When I was a child," which provides additional information about the time frame in which the speaker is referring to. The use of "was" in the sentence indicates that the speaker is referring to a specific point in the past when Friday the 13th was the scariest movie in their opinion. Overall, the sentence is a simple statement about the speaker's experience as a child and their perception of horror movies during that time.

The sentence "The latest was released in 2009 right?" is an example of the passive voice in the past tense. In the passive voice, the subject of the sentence is the receiver of the action. In this case, the subject is "The latest," and the verb is "was released." This construction emphasizes the action of the release rather than the entity performing the action. The use of "was released" indicates that the focus is on the release of the latest version, and the specific agent or performer of the action is not mentioned in the sentence. The addition of "right?" at the end of the sentence turns it into a question, seeking confirmation or clarification about the release date of the latest version. The error analysis about the word “was” can be seen in the Table 8. below:

Table 8. Kinds of Error in Corpus Analysis of “was”

No	Kinds of Error	Concordance
1	Miss formation	Prof. Dumbledore was pass away, yesterday

The sentence "Prof. Dumbledore was pass away, yesterday" contains a formation error. The correct sentence should be "Prof. Dumbledore passed away yesterday." The error lies in the use of the verb "pass away" in the past tense. The correct past tense form of "pass away" is "passed away." Therefore, the sentence contains a formation error in the past tense of the verb "pass away"

3.7. Am

The verb "am" is the first person singular present tense form of the verb "to be." (Murphy,

2000). It is used to describe or tell us the condition of the subject, such as age, nationality, job, or other traits. It is used when the subject is "I." For example, "I am happy" or "I am an engineer.". The verb "am" is used to indicate a present state or action of the subject. It is important to use the correct form of the verb "be" to ensure that the sentence is grammatically correct and conveys the intended meaning. The result about the word “am” can be seen in the Table 9. below:

Table 9. The Result of Corpus Analysis of “am”

No	Concordance	Kinds of Structure	Subject Categorization	Grammatical Acceptance		Freq
				Yes	No	
1	I am a professional chef at home	Nominal sentence	Pronoun	9	-	9
2	What am I?	Question	Pronoun	2	-	2
Total						11

The sentence "I am a professional chef at home" is an example of a nominal sentence in the

present tense. In this sentence, the subject is "I," the linking verb is "am," and the subject

complement is "a professional chef at home." The present tense "am" is used to describe the present state of being or profession. It indicates that the speaker's current profession is a professional chef, and this is true in the present time. The sentence follows the pattern of formation for nominal sentences in the present tense, where the subject is followed by the linking verb "am" and the subject complement. This construction is used to make statements about the speaker's identity, profession, or state of being in the present time.

The sentence "What am I?" is an example of a present tense interrogative sentence. It is formed by using the verb "am" (a form of the verb "be") with the subject "I" to ask a question about the present state of being or action. In this case, the question is seeking information about the speaker's current state or action. The structure of the sentence is formed by placing the verb "am" before the subject "I" to create the question "Am I?" or by adding a question word to the sentence, such as "What," to ask "What am I?" This construction is used to seek information or clarification about the present state of being or action of the speaker.

The researchers do not find the error in using "am" at the students' conversation through

WhatsApp. The analysis indicated that the usage of 'am' aligned contextually with the communicative norms observed within the students' digital interactions. This suggests that their use of 'am' was appropriate and contextually fitting within the informal discourse of their WhatsApp conversations. As such, no grammatical or contextual issues were found in the students' application of 'am' in this specific communication context.

3.8. Were

In the grammatical structure of the English language, "were" is a form of the verb "be" used to indicate events that occurred in the past within a statement. Specifically, "were" is applied to plural subjects such as "they", "we", and "you" (in plural form) in sentences referring to past events (Azhar, 1992).

In the corpus analysis of the use of the word "were" in unstructured written conversations from a high school student WhatsApp group, two main patterns reflecting variations in the usage of this word are observed. The result about the word "were" can be seen in the Table 10. below:

Table 10. The Result of Corpus Analysis of "were"

No	Concordance	Kinds of Structure	Subject Categorization	Grammatical acceptance		Freq
				Yes	No	
1.	I think their accent were so sexy	Nominal	Noun (3) Pronoun (1)	2	2	4
2	More than a dozen towns were destroyed and 71 people died	Passive	Noun	1	-	1
Total						5

Firstly, "were" in nominal sentences, like "I think their accent was so sexy", expresses attributes or characteristics of the subject an accent considered sexy. This verb choice indicates understanding of the subject with specific features, illustrating "were" as the past tense of "be", depicting the object's state at a particular time. Notably, among four nominal sentences, three feature noun subjects, and one uses a pronoun, adding diversity to explored sentence structures.

Secondly, "were" is used in passive voice sentences, as seen in the example "More than a dozen towns were destroyed and 71 people died". This shift in focus from the doer to the object emphasises events in the past, highlighting town destruction and casualties. In this context, "were"

is crucial for structuring sentences that emphasise the impact of actions. One of the four nominal sentences employs "were" in passive voice, emphasising the significant impact of destructive events with a noun-based subject.

This corpus analysis illustrates the two primary uses of "were" in high school student conversations, namely, in nominal contexts to express attributes or characteristics of the subject, and in passive voice sentences to highlight the object receiving the action. This variation reflects the syntactic and semantic richness of the English language in conveying various shades of meaning in everyday conversations. The result about the error analysis of the word "were" can be seen in the Table 11. below:

Table 11. Kinds of Error in Corpus Analysis of "were"

No	Kinds of Error	Concordance
1.	Misordering	According police sitting next to him, were "Fuck"
2	Misordering	I think their accent were so sexy?

The sentence indicates a form of error commonly referred to as “misordering”. This error arises due to a lack of alignment in the sequence of words and sentence structure with established grammatical norms. In this context, the arrangement of words and sentence elements is suboptimal, creating an impression of less coherence and non-compliance with language standards. For instance, the use of the word “were” after the phrase “police sitting next to him” might lead to confusion in conveying the overall meaning of the sentence.

3.9. Been

The use of the word “been” occurs in the constructions of present perfect progressive and past perfect progressive (Duffley & Morin, 2022). It is a part of grammar indicating the duration or continuity of an activity from the past to the present moment or a specific reference point (Azhar, 1992). In both constructions, “been” functions as the past form of the auxiliary verb

“be”, depicting the ongoing nature of an action or state within a particular time frame. The second error is “I think their accent were so sexy?”. The sentence contains a form of error commonly referred to as “misformation”. This error arises due to the use of the verb “were”, which is not aligned with the context of the sentence. The correct verb should be “was”, considering the singular subject “accent” that requires a singular form of the verb. This error creates a discrepancy with grammatical norms, emphasizing that the sentence should read, “I think their accent was so sexy?”

“be”, depicting the ongoing nature of an action or state within a particular time frame.

In the corpus analysis of the use of the word “been” in unstructured written conversations from a high school student WhatsApp group, two main patterns reflecting variations in the context of using this word are observed. The result about the word “been” can be seen in the Table 12. below:

Table 12. the Result of Corpus Analysis of “Been”

No	Concordance	Kinds of Structure	Subject Categorization	Grammatical acceptance		Freq
				Yes	No	
2.	Hi, Marelli. You've been inside here.	Nominal	Pronoun (2)	2	-	2
3	Maybe i've been watching it thousand times	Progressive	Pronoun (1)	1	-	1
Total				3		

Initially, it is crucial to note that in the type of sentences under consideration, two instances employ “been”, utilizing pronouns as subjects (Duffley & Morin, 2022). One example is “Late to be here... Hi, Marelli. You've been inside here. Let's talk about anything again?”. In essence, the use of “been” in this context signifies the continuation of an action or state from the past to the present. When stating “You've been inside here”, it implies that the subject was previously inside, and this remains relevant at the current moment. The utilization of “been” establishes a seamless connection between past occurrences and their ongoing relevance.

Secondly, Hundt & Smith (2009) explains that “been” is used in progressive sentences, like in “Maybe I've been watching it a thousand times”. In this context, “been” is part of the Present Perfect Progressive construction, indicating an ongoing action from the past to the present (Duffley & Morin, 2022). This emphasizes the extended duration of a specific activity, such as watching something repeatedly. The use of “been” vividly portrays the progressive nature of the activity over a considerable period. It's worth noting that this paragraph starts with a

sentence in the progressive tense, adding diversity to the explored structures.

4. Conclusion and Suggestions

Through the analysis of unstructured written discourse in the WhatsApp group of vocational school students (SMK), we can observe that the use of “to be” in students’ communication has shown positive progress. The conversation results reflect a sufficient understanding of the use of “to be” in sentences. The students employ the use of “to be” in several variations of grammar structure such as using “to be” on nominal sentence, passive sentence, progressive tense and be going to. However, observations also reveal some errors, particularly related to subject-verb agreement, word order, and omission in singular and plural contexts.

The main conclusion of this research is that while students have made progress in understanding the basic structure of “to be”, there is a need for more attention to specific aspects that still contain errors. Therefore, teachers are advised to provide more in-depth lessons, focusing on the placement of “be” in sentences,

appropriate contextual usage, and mastery of detailed aspects such as subject-verb agreement.

The importance of a profound understanding of "to be" not only affects students' written communication skills but also enriches their overall understanding of language structure. Hence, a contextual and interactive learning approach is required so that students can better comprehend the use of "to be" in everyday situations. The hope is that this will encourage students to integrate their knowledge more effectively in both written and oral communication activities.

References

- Azhar, B. S. (1992). *Fundamental of English Grammar*. Longman.
- Cahyati, N., Silvana Sinar, T., & Yusuf, M. (2021). Using AntConc Software in Analyzing Modality: Case Study in EFL Tertiary Students' Hortatory Exposition Text in Indonesia. In *LingPoet: Journal of Linguistics and Literary Research* (Vol. 2, Issue 1). <http://talenta.usu.ac.id/lingpoet>
- Duffley, P., & Morin, G. (2022). It's been a while since I've been to church: The use of the Present Perfect after the conjunction since. *Australian Journal of Linguistics*, 42(2), 184–205. <https://doi.org/10.1080/07268602.2022.2124844>
- Gu, W. (2020). *The Use of Participles and Gerunds*. ERIC - Education Resources Information Center. <https://files.eric.ed.gov/fulltext/ED608205.pdf>
- Harahap, F. S. (2020). Students' Errors on the Use of Simple Present in Writing Daily Activities. *Journal Language League*, X(2), 594–603.
- Hasko, V. (2012). Qualitative Corpus Analysis. In *The Encyclopedia of Applied Linguistics*. Wiley. <https://doi.org/10.1002/9781405198431.wbal0974>
- Hundt, M., & Smith, N. (2009). The present perfect in British and American English: Has there been any change, recently? *ICAME Journal*, 33.
- Hussein, N. O. (2018). *The Effect of Using Authentic Materials on Developing Undergraduate EFL Students' Communicative Competence*. <https://www.researchgate.net/publication/336013535>
- Ibrohimova, M. B., & Valiyev, Q. O. (2019). The Role of the Verb "to be" in English. *Мировая Наука*, 4(25), 52–56. science-j.com
- Ismi, I., Nurdin, N., & Kuliachana, A. (2021). Teaching Present Progressive Tense Through Whole Brain Method to the Eight Grade Students of SMP Negeri 2 Palu. *Datokarama English Education Journal*, 1, 44–55. <https://doi.org/10.24239/dee.v1i2.17>
- Lee, C. E. C. (2023). WhatsApp Use in a Higher Education Learning Environment: Perspective of Students of a Malaysian Private University on Academic Performance and Team Effectiveness. *Education Sciences*, 13, 244. <https://doi.org/10.3390/educsci13030244>
- Lubis, ainun, Humairo, A., Dwi Agustina, R., Sundari, S., Havalza Panjaitan, Y., & Manajemen Pendidikan, J. (2022). Analysis Of Student's Errors In Writing Simple Present Tense At Eighth Grade EDU MANAGE. In *Edu Manage-Jurnal Manajemen Pendidikan Islam EDU MANAGE* (Vol. 1, Issue 1). <https://www.jurnal.staini.ac.id/index.php/edumanager>
- Malabar, F., & Mustapa, L. (2022). Analyzing students' passive voice difficulties. *Indonesian EFL Journal*, 8(2), 155–164. <https://doi.org/10.25134/ieflj.v8i2.6439>
- Murphy, R. (2000). *English Grammar in Use. Grammar reference*.
- Suryadi, S., Muhyidin, A., & Syafrizal, S. (2024). Using Discourse Markers to Enhance Students' Speaking Skills. *Ideguru: Jurnal Karya Ilmiah Guru*, 9(3), 1181–1186. <https://doi.org/10.51169/ideguru.v9i3.1002>
- Syafar, A. (2014). Teaching the Present Continuous Tense Through Short Dialogue. In *Journal of English Language Teaching Society (ELTS)* (Vol. 2, Issue 1).
- Voelkel, S., & Kretzschmar, F. (2021). Introducing Linguistic Research. In *Introducing Linguistic Research*. Cambridge University Press. <https://doi.org/10.1017/9781316884485>