IDEAS

Journal on Language Teaching and Learning, Linguistics and Literature

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Volume 14, Number 2, December 2026
pp. 7900 - 7925

Copyright © 2026 The Author
IDEAS is licensed under CC-BY-SA 4.0 License



Issued by English study program of UIN Palopo

Recurrent Multi-word Sequences: Syntactic Structure and Function in EFL Student Writing Texts

Rasuna Rasid Talib[1], Zulkifli Tanipu[2], Sri Widyarti Ali[3], Muzdalifah Mahmud[4], Indah Nuraini Gusasi[5]

[1,2,3,4,5]English Language Education Study Programme, Faculty of Letters and Culture, Universitas Negeri Gorontalo

Article Info

Received: 2026-08-25
Revised: 2026-09-14
Accepted: 2026-09-18

Keywords:
Recurrent Ulti-word Sequences;
EFL Writing;
Syntactic Structure;
Discourse Function;
Learner Corpus

DOI:
10.24256/ideas. v14i2. 12429

Corresponding Author:
Indah Nuraini Gusasi
Indh.gusasiiii@gmail.com
English Language Education Study Programme,
Faculty of Letters and Culture,
Universitas Negeri Gorontalo

Abstract

This study analyzed Recurrent Multi-word Sequences (RMS) in the writing of EFL students in the English Language Education Study Programme at Universitas Negeri Gorontalo, focusing on their syntactic structures and discourse functions. Using a qualitative approach, the study examined 98 writing texts containing 88,576 tokens from third-semester students in the Class of 2020 and 45 texts containing 25,969 tokens from fifth-semester students in the Class of 2021. The RMS were identified using AntConc based on frequency and range, and the most frequent sequences were examined through concordance lines. The analysis focused on two syntactic structures, Noun Phrase (NP) and Verb Phrase (VP), and two discourse functions, Discourse Organizers (DO) and Stance Expressions (SE). Results showed different syntactic patterns between the two groups: Class of 2020 was dominated by NP-structured RMS (66.7%), while Class of 2021 was dominated by VP-structured RMS (66.3%). In terms of function, DO was dominant in both groups, accounting for 79.5% of RMS in Class of 2020 and 64.2% in Class of 2021, while SE accounted for 20.5% and 35.8%, respectively. The findings indicate that Class of 2020 use RMS to refer to and describe topic-related information, whereas Class of 2021 use RMS verb-based sequences for expressing possibility, evaluation, conclusions, and relationships between ideas. The findings suggest that RMS provide useful patterns for describing how EFL students construct academic meaning through recurring syntactic structures and discourse functions.

1. Introduction

Writing is one of the most important skills for EFL students because it allows them to communicate ideas, opinions, and knowledge in a clear and organized way (Marlina, 2024). In academic writing, students are expected not only to use correct grammar and vocabulary but also to produce texts that are coherent, natural, and well-structured (De Smedt et al., 2022). However, many EFL students still experience difficulties in organizing ideas cohesively and using appropriate language patterns in writing, resulting in texts that

sound less natural and less connected (Nora Fitria et al., 2024; Ismi et al., 2025). For this reason, examining the language patterns that students frequently use in their writing has become an important area of research in EFL writing studies.

English Department students at Universitas Negeri Gorontalo still face difficulties in writing texts, especially in organizing ideas clearly and using appropriate grammatical structures. Their writing often lacks cohesion and natural flow, even when they understand basic grammar and vocabulary (Syarif, 2022; Aljoza & Almustafa, 2024). This problem may be caused by limited use of effective language patterns, such as recurrent multi-word sequences (RMS), which are important for building fluent and well-structured texts.

Recurrent multi-word sequences (RMS) are strings of words that repeatedly occur together across multiple texts and registers in language corpora (Cortes, 2013a; Douglas Biber et al., 2004). Recent studies in EFL writing research have highlighted that learners' language development can be understood more clearly through patterns that appear repeatedly in their texts. According to (Li & Fang, 2022), RMS provide a useful window into how students naturally combine words to form meaning in real writing situations.

RMS are also considered reliable indicators of cohesion, fluency, and syntactic control (Douglas Biber et al., 2004). (Casal & Yoon, 2023) also found that students who frequently use RMS tend to produce more coherent and structured texts. These RMS often reflect the way learners build cohesion, showing how ideas are linked across sentences. Therefore, analyzing RMS helps researchers understand how students develop their writing skills from both a cohesive and structural perspective.

In line with this, RMS are closely related to syntactic structure and function because they show how words are commonly organized and used in real communication. In terms of structure, RMS can appear as patterns like noun phrases (the importance of the study) or verb phrases (is based on the), while in terms of function, they help express meaning such as transitions (on the other hand) or organizing ideas (the purpose of this study is). Recent studies show that mastering these patterns helps learners write more fluent and cohesive texts. However, many UNG students still struggle to use these patterns effectively, which causes their writing to lack cohesion and sound less natural (Aljoza & Almustafa, 2024; Hamdani & Abid, 2025). Therefore, analyzing RMS in terms of syntactic structure and function can help explain and improve students' writing problems.

Previous research on RMS has also been conducted in the English Language Education Study Programme (ELESP) of Universitas Negeri Gorontalo, such as the study by Mahading (2024). However, the study did not examine in depth how RMS are used in terms of syntactic structure and function in students' writing. It is still unclear whether students' ability to use RMS appropriately in these aspects can make their writing more complex and varied.

Therefore, this study aims to explore this issue more deeply by focusing on how RMS reflect both structure and function in writing. Another important gap addressed in this research is that it specifically investigates the writing of students from the ELESP at Universitas Negeri Gorontalo. Even so, understanding the relationship between RMS and syntactic structures & functions is crucial for portraying students' writing ability. Recent findings (Appel et al., 2024) show that analyzing RMS allows researchers to identify consistent patterns that reflect both linguistic development and functional choices.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

These patterns help explain how students' writing evolves, not only in terms of vocabulary, but also in how sentences are organized and linked. As emphasized by (Yang et al., 2022), examining RMS within syntactic functions provides a more balanced view of behavior in writing, particularly in relation to recurrent syntactic patterns and discourse functions. because it integrates form and meaning. Therefore, the present research uses RMS analysis to understand how students' writing reflects their syntactic structure and the function created through their writing. This research aims to examine how students use RMS in their writing in terms of syntactic structures and functions, as well as to provide a clearer description of students' writing based on the patterns of syntactic structures and functions reflected through their use of RMS.

Limitation of the study

Several limitations of this study should be acknowledged as a basis for future research. The data were drawn from two groups of students from a single study programme and institution, and the two corpora differed considerably in size, which limits the generalizability of the findings and makes raw frequency comparisons inappropriate without considering corpus size. In addition, the analysis focused specifically on noun phrases and verb phrases for syntactic structure and discourse organizers and stance expressions for function. Future researchers are therefore encouraged to investigate larger and more balanced corpora involving students from different semesters, proficiency levels, and institutions.

Future studies could also examine other syntactic structures and functional categories, as well as use longitudinal designs to observe how students' RMS use develops over time. Interviews, questionnaires, or classroom observations could further be used to investigate the factors that influence students' choices and development of RMS, while future research may also examine the relationship between RMS use and the overall quality of students' academic writing.

2. Method

This study employs a comparative corpus-based research design to examine the use of Recurrent Multi-word Sequences (RMS) in EFL students' writing. A corpus-based approach is used because the study identifies recurrent sequences through frequency and range and examines their distribution in two learner corpora. The study compares the RMS identified in the writing of third-semester and fifth-semester students in terms of syntactic structure and function. The comparison is intended to describe similarities and differences in the language patterns found in the two corpora rather than to measure students' overall writing ability or language development.

The analysis consists of two complementary stages. First, AntConc is used to identify RMS based on frequency and range criteria. Second, the identified RMS are examined through concordance and contextual analysis to determine their syntactic structures and discourse functions. The first stage provides corpus-based evidence concerning the recurrence and distribution of the sequences, while the second stage provides a qualitative interpretation of how the sequences are used in context. Therefore, the study combines corpus-based quantitative information with qualitative contextual interpretation within a comparative design.

The data in this study come from a specialized corpus that contains samples of students' academic writing. This corpus represents real writing produced by EFL learners in an authentic classroom context. By using corpus-based analysis, this study looks closely at how students naturally use RMS in their writing and how these sequences make their texts more cohesive and grammatically stronger. The main focus is on identifying the most frequent and varied RMS used by both groups of students.

This study applies a two-layered analysis to examine students' use of RMS in their writing. The first layer focuses on identifying RMS. Because the two learner corpora differ considerably in size, the frequency and range criteria were considered in relation to corpus size. The initial extraction used a minimum frequency of five occurrences and a minimum range of three texts to identify recurrent four-word sequences. However, raw frequency alone may favor sequences from the larger corpus.

Therefore, normalized frequencies were also calculated to allow a more balanced comparison between the two corpora. Range was retained as an additional criterion because it indicates how widely an RMS is distributed across the students' texts rather than whether it occurs repeatedly in only one or a few texts. The selected criteria were therefore used as identification criteria, while normalized frequency and distribution across texts were considered when interpreting differences between the two corpora.The researcher analyzes the frequency (how often the sequences appear) and range (how widely they are distributed across students' texts) in two groups: third-semester students (2020 academic year) and fifth-semester students (2021 academic year). This analysis helps to identify the most common RMS used by students and to compare their usage across different academic levels.

The second layer focuses on examining how these identified RMS are used in terms of syntactic structure and function. In this study, syntactic structure is limited to noun phrase patterns, while function is focused on stance expressions. This study focuses on noun phrases and verb phrases in terms of syntactic structure, as well as stance expressions and discourse organizers in terms of function. These categories were selected because they appeared most frequently in the students' writing texts that will be analyzed in this study. Noun phrases and verb phrases help show how students construct grammatical patterns and develop sentence complexity, while stance expressions and discourse organizers show how students' express opinions, attitudes, and organize ideas in their writing. These functions are important because they help reveal students' writing styles and how they build meaning in their texts.

Although other structures and functions were also found in the analysis, this study only focuses on these four categories to keep the discussion more focused and systematic. Analyzing all RMS categories at once would make the study too broad and difficult to discuss in depth. By limiting the analysis to the most frequently used structures and functions, the researcher can provide a clearer description of how students use RMS in their writing. Through this analysis, the study aims to compare whether third-semester or fifth-semester students use RMS in a more complex, organized, and formal way.

Source of Data

The data of this study consist of a corpus of students' writing texts of third-and fifth-semester students of ELESP, which were collected from two separated corpora:

Corpora 1: EFL students in the third semester (2020 academic year). Data collected from these corpora consisted of written assignments.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Corpora 2: EFL students in their fifth semester (2021 academic year). Data collected from these consisted of written assignments.

The two learner corpora were constructed from students' writing produced in their respective academic courses and writing tasks. Therefore, differences in course, task, topic, and genre may influence the recurrent language patterns found in each corpus. For example, topic-specific assignments may encourage students to repeatedly use terminology related to a particular subject, whereas more general academic tasks may encourage broader academic expressions. Consequently, differences between the two corpora are interpreted as differences in observed RMS patterns within the respective corpus contexts rather than as evidence that semester level alone caused the differences. This consideration is particularly important because the study uses a cross-sectional comparison of two different student cohorts.

3. Result

The findings are presented in a systematic way to answer the two research questions. First, this chapter provides an overview of the RMS identified in the corpus. It then describes how the identified RMS are used in terms of syntactic structure, particularly noun phrase and verb phrase patterns. Finally, it explains how the RMS function as stance expressions and discourse organizers in students' writing. The discussion of these findings is presented in the following sections.

Coding Validation

To improve the reliability of the qualitative classification, the identified RMS were examined through concordance lines and their surrounding contexts. The classification of syntactic structure and function was based on the definitions and categories provided by the selected theoretical framework. The researcher first classified the RMS independently and then rechecked the classifications against their concordance contexts. Ambiguous cases were reviewed by comparing the sequence with the definitions and structural examples in the framework. This process was used to reduce inconsistent classification and ensure that the final categories reflected the actual use of the RMS in the students' texts.

Overview of RMS Identified in the Corpus

The Recurrent Multi-Word Sequences (RMS) were extracted using the N-Gram feature in AntConc (version 4.4.0). The analysis focused on four-word sequences. Because the two corpora were different in size, the extraction settings were adjusted for each group. Using the same extraction criteria for corpora of different sizes could produce unreliable results, such as identifying too many unimportant sequences in the smaller corpus or missing important patterns in the larger corpus.

Therefore, the extraction parameters were adapted to match the size of each corpus, following the framework of Biber et al. (2004). Unlike Wray (2002), who focuses more broadly on formulaic language as prefabricated sequences, Biber provides specific corpus-based criteria such as frequency and range for identifying RMS. Therefore, Biber's framework was more suitable for this study because it supports the systematic identification and comparison of RMS across students' texts.

For the Class of 2020, the corpus contained 98 texts with a total of 86.15 occurrences per 10,000 words. A minimum frequency of 5 occurrences and a minimum range of 3 texts were used to identify the RMS. Meanwhile, the Class of 2021 corpus consisted of 45 texts with 23.49 occurrences per 10,000 words. Since this corpus was smaller, the same minimum frequency of 5 occurrences and a minimum range of 3 texts were applied to ensure that enough recurrent sequences could be identified for further analysis.

After the extraction process, the identified RMS were classified into two categories. First, they were analyzed based on their syntactic structure, namely Noun Phrase and Verb Phrase. Second, they were classified according to their discourse function as Discourse Organizers or Stance Expressions. Before discussing these categories in detail, it is important to provide an overview of the total number of RMS identified in each corpus. This overview helps explain the amount of data analyzed and provides a foundation for interpreting the structural and functional patterns discussed in the following sections. Table 3.1 presents the corpus information and the total number of RMS extracted from each group before the classification process.

Table 3.1. Overview of Corpus Specifications and Total RMS Identified per Cohort

Group	Number of Texts	Total Tokens	N-Gram Size	Min. Co-occurrence	Min. Range	Total RMS
Class of 2020 (3rd Semester)	98	88,576	4 words	5	3	78
Class of 2021 (5th Semester)	45	25,969	4 words	5	3	18

As shown in Table 3.1, Class of 2020 yielded a considerably larger number of RMS (78 sequences) compared to Class of 2021 (18 sequences), a difference that is largely attributable to the substantial difference in corpus size between the two groups 88,576 tokens versus 25,969 tokens, a ratio of more than three to one. The higher token count and greater number of texts in the 2020 corpus naturally produce a wider range of recurrent sequences at the specified thresholds, since a larger corpus offers more opportunities for the same four-word combination to recur across different texts and writers. This difference in raw output does not by itself indicate a difference in writing quality or sophistication between the two cohorts; rather, it sets the scale against which the structural and functional distributions discussed in the following sections must be read, since proportions and percentages rather than raw counts provide the more meaningful basis for cross-cohort comparison.

In addition to frequency, each of the RMS tables presented in the following sections also reports the range of every sequence, that is, the number of distinct texts in which the bundle occurs at least once, as opposed to the total number of times it occurs across the corpus as a whole. Range is a standard companion measure to frequency in corpus-based lexical bundle research (Biber et al., 2004), because frequency alone cannot distinguish between a sequence that is genuinely widespread across many student writers and one that is highly frequent only because it happens to be repeated many times within a small number of individual texts.

A bundle with a high frequency but a low range, for instance, may reflect one or two writers' idiosyncratic phrasing rather than a pattern shared across the cohort, whereas a bundle with a comparatively high range relative to its frequency indicates that the sequence is used consistently by a broad cross-section of students. Reporting range alongside frequency therefore allows the following tables to be read not only for how often a sequence recurs, but also for how representative that recurrence is of the cohort as a

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

whole. With this overview established, the following sections turn to a detailed analysis of each group's RMS, first in terms of syntactic structure and then in terms of discourse function.

Syntactic Structure of RMS in Students' Writing

After identifying the Recurrent Multi-word Sequences (RMS) in the corpus, the next step was to analyze their syntactic structure. This analysis answers the first research question by examining how students use RMS as grammatical patterns in their writing. Following the framework of Biber et al. (2004), the identified RMS were classified based on their syntactic structure. In this study, the analysis focuses on two structural categories: noun phrase patterns and verb phrase patterns. These categories were selected because they were the most frequently found in the students' writing corpus. After classifying the identified RMSs into noun phrase and verb phrase patterns, the distribution of these syntactic structures for each cohort is presented in the following figures. The charts provide a visual comparison of the frequency of each syntactic pattern in the students' writing.

Figure 3.1. Distribution of Syntactic Structures in RMS occurrences/ frequency of the Class of 2020

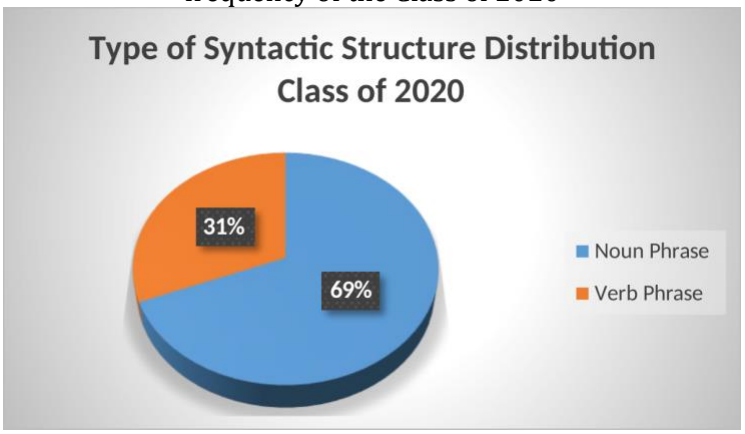
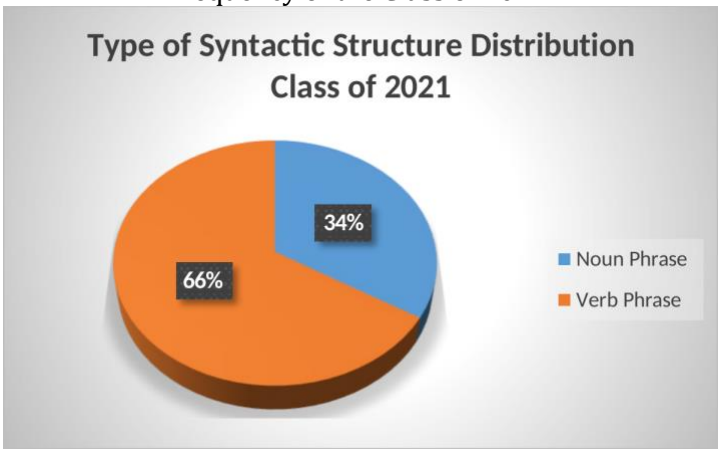


Figure 3.2. Distribution of Syntactic Structures in RMS occurrences/ frequency of the Class of 2021



As shown in Figure 3.1, the Class of 2020 demonstrates a greater use of noun phrase structures, accounting for 62% of the occurrences/frequency identified RMSs, while verb phrase structures represent 38%. This indicates that noun phrase constructions were more frequently employed by students in this cohort. The dominance of noun phrases suggests that the students tended to construct recurring multiword sequences centered around nominal expressions, which are commonly used to present concepts, objects, or abstract ideas in academic writing. In contrast, Figure 3.2 shows that the Class of 2021 exhibits a different pattern.

Verb phrase structures account for 66% of the occurrences/frequency identified RMS, whereas noun phrase structures make up only 34%. This suggests that students in this cohort relied more heavily on verb phrase constructions than noun phrase constructions. The higher proportion of verb phrases indicates a greater tendency to use recurring expressions centered on actions or processes, reflecting a different pattern of syntactic preference compared to the Class of 2020.

Noun Phrase Patterns

Noun Phrase (NP) - structured RMS represents sequences in which the main grammatical element is a noun. These sequences usually consist of a head noun together with other grammatical elements, such as determiners, adjectives, prepositional phrases, or other modifiers that provide additional information. In academic writing, NP-structured RMS help writers introduce, explain, and develop important ideas in a clear and organized way. The frequent use of these sequences in students' writing is often considered a sign that students are developing their ability to produce more complex noun phrases, which is an important feature of academic writing. It also shows that they are becoming more frequent at organizing information into clear and meaningful units (Biber et al., 2004; Yang et al., 2022).

In Class of 2020, a total of 52 NP-structured RMS were identified, with a combined frequency of 671 occurrences across the corpus, a figure that already signals how central nominal phrasing is to this cohort's writing. These sequences reflect the topical and conceptual vocabulary of the students' writing assignments, which centred heavily on themes of children's literature, creative writing, and educational contexts for students with special needs.

Representative NP bundles include in the form of (f=69), children s literature is (f=32), of children s literature (f=28), literature and creative writing (f=26), and the results of the (f=19). Read together, these five bundles alone trace the shape of the assignment: institutional or procedural description (in the form of), disciplinary definition (children s literature is, of children s literature), thematic linkage (literature and creative writing), and empirical reporting (the results of the).

The high frequency of these sequences suggests that students in this cohort frequently used topic-related noun phrases in their writing. This indicates that they were familiar with the key concepts and vocabulary commonly used in their field of study. Table 3.2 presents the complete list of NP-structured RMS identified in Class of 2020, ranked by frequency, so that the relative weight of each construction within the cohort's writing can be easily identified.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Table 3.2. Frequency Distribution of Noun Phrase-Structured RMS in the Class of 2020 Corpus

Recurrent Multi-word Sequence	Frequency	Range
children s literature is	32	8
of children s literature	28	7
literature and creative writing	26	10
president of the country	25	9
blind junior high school	22	3
creative writing is a	20	11
the results of the	19	10
students with special needs	18	5
The blind junior high	18	3
Between literature and creative	15	9

Beyond Beyond the raw counts in Table 4.2, it is instructive to see how the two most frequent bundles actually surface in running text, since frequency alone cannot show whether a sequence is used mechanically or meaningfully. The following concordance lines, drawn directly from the student texts, illustrate how in the form of and the results of the the two highest-ranking NP bundles in this cohort appear in authentic writing contexts:

"... The relationship between literature and creative writing is in an imaginative work, which means..."

"... The general description of children's literature is that children's literature is specially made for children's reading..."

These two examples show that the NP bundles used by the Class of 2020 mainly help students introduce and explain concepts related to their field of study. For example, literature and creative writing of is used to describe the specific difference of that, while children's literature is used to introduce or present information about children's literature. Most of these sequences have a referential function because they help writers identify and describe important ideas or information. This pattern reflects the descriptive and explanatory nature of the writing tasks completed by this group. After examining the NP bundles used by the Class of 2020, the next section explores the Class of 2021 use of NP bundles in their writing.

In Class of 2021, a total of 7 NP-structured RMS were identified, with a combined frequency of 32 occurrences. The NP patterns identified in this group demonstrate a variety and thematic generalizability. Sequences such as one of the most (f=7), for a long time (f=5), an example of a (f=4), and the purpose of this (f=4) represent general academic NP constructions applicable across a range of argumentative and analytical writing contexts. Table 4.3 presents the NP-structured RMS identified in Class of 2021.

Table 3.3. Frequency Distribution of Noun Phrase-Structured RMS in the Class of 2021 Corpus

Recurrent Multi-word Sequence	Frequency	Range
one of the most	7	6
an example of a	4	3
country in the world	4	3
country is the most	4	3

Recurrent Multi-word Sequence	Frequency	Range
the purpose of this	4	4

Concordance lines from Class of 2021 illustrate the deployment of the two most frequent NP sequences, one of the most and the purpose of this, in context:

> *"...And there are many festivals and traditions in Chinese culture. One of the most important is the Lunar New Year..."*
>
> *"...The purpose of this essay is to discuss both perspectives and will continue to the conclusion..."*

These two concordance lines show that one of the most and the purpose of this are used general academic expressions. The phrase one of the most is used to introduce or compare items within a broad category, while the purpose of this is commonly used to introduce the main objective or argument of a piece of writing. These expressions are not limited to one specific topic and can be used in many different academic contexts. This implies that fifth-semester students are using RMS with general academic writing patterns and are able to use them to organize and present their ideas effectively. After discussing the noun phrase patterns used by both groups, the next section examines the second syntactic category, namely Verb Phrase patterns.

Verb Phrase Patterns

Verb Phrase (VP)-structured RMS represents sequences in which the main grammatical element is a verb. These sequences may include main verbs, modal verbs, linking verbs, and other elements that complete the meaning of the verb. In academic writing, VP-structured RMS help writers express opinions, explain ideas, describe processes, and present logical reasoning. Therefore, examining these sequences helps show how students use verb patterns to communicate their ideas clearly and appropriately in an academic context. It also provides an understanding of their ability to use different types of verb constructions in their writing. (Biber et al., 2004).

Class of 2020, 26 VP-structured RMS were identified, with a combined frequency of 302 occurrences. The most frequent sequences include *is one of the* (f=25*), it can be concluded* (f=21), *can be concluded that* (f=20), *to be able to* (f=16), and *there is also a* (f=16). The high frequency of it can be concluded and can be concluded that, which appeared 41 times in total, shows that students frequently used common expressions to write conclusions in their academic texts. This suggests that they were already familiar with one of the common conventions of academic writing, even at this stage of their study.

In addition, the frequent use of being able to see (f=15) and not being able to (f=15) reflects the topics discussed in their writing. These expressions were commonly used because many of the writing assignments focused on issues related to visual impairment, particularly the experiences and abilities of students with visual disabilities. Table 3.4 presents the complete list of VP-structured RMS for Class of 2020.

Table 3.4. Frequency Distribution of Verb Phrase-Structured RMS in the Class of 2020 Corpus

Recurrent Multi-word Sequence	Frequency	Range
is one of the	25	19
it can be concluded	21	13
can be concluded that	20	13
is a type of	16	7

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Recurrent Multi-word Sequence	Frequency	Range
there is also a	16	8
to be able to	16	13
being able to see	15	3
is a form of	15	8
not being able to	15	3
Be in the form	10	6

To illustrate how these bundles related to conclusions and ability constructions actually function in student writing, the following concordance lines show it can be concluded and not being able to, two of the cohort's most frequent VP bundles, in their original textual context:

> *"...Based on the discussion presented in the previous sections, it can be concluded that literature for children with special needs must be adapted to accommodate their sensory and cognitive needs..."*

> *"...Thus, being blind means not being able to see in the conventional sense, though this does not equate to a complete absence of perceptual experience..."*

These examples show that the VP bundles used by the Class of 2020 serve two purposes. First, they help students connect ideas and draw conclusions in their writing. Second, they are used to discuss the specific topics given in the writing assignments. The frequent use of it can be concluded and can be concluded that in many students' texts shows that they commonly relied on these expressions to write conclusions. It suggests that they depended on the same expressions rather than using varied verb phrase patterns. The next section examines the Class of 2021 use of VP bundles in their writing.

In Class of 2021, 11 VP-structured RMS were identified, with a combined frequency of 63. The most frequent VP bundles are is one of the (f=14), to be able to (f=5). A notable feature of the Class of 2021 VP set is the presence of evaluative and epistemic sequences such as is the most important (f=4), it is also possible (f=5), and make it difficult for (f=4), which encode the writer's assessment and judgement regarding the propositions being presented. This suggests a deployment of modalizing and evaluative VP structures in fifth-semester writing, indicative of argumentative positioning. Table 4.5 presents the VP-structured RMS identified in Class of 2021.

Table 3.5. Frequency Distribution of Verb Phrase-Structured RMS in the Class of 2021 Corpus

Recurrent Multi-word Sequence	Frequency	Range
is one of the	14	10
is an example of	5	3
it is also possible	5	3
to be able to	5	4
can be concluded that	4	4
have the right to	4	3
is the most important	4	4
it can be concluded	4	4
make it difficult for	4	3

Concordance lines from Class of 2021 illustrate the evaluative and contrastive functions of two of this cohort's characteristic VP bundles, on the other hand and it is also possible:

"... Mandarin is one of the most widely used languages in the country..."

"...However, it is also possible for values to contradict and conflict with each other, especially in complex modern industrial societies..."

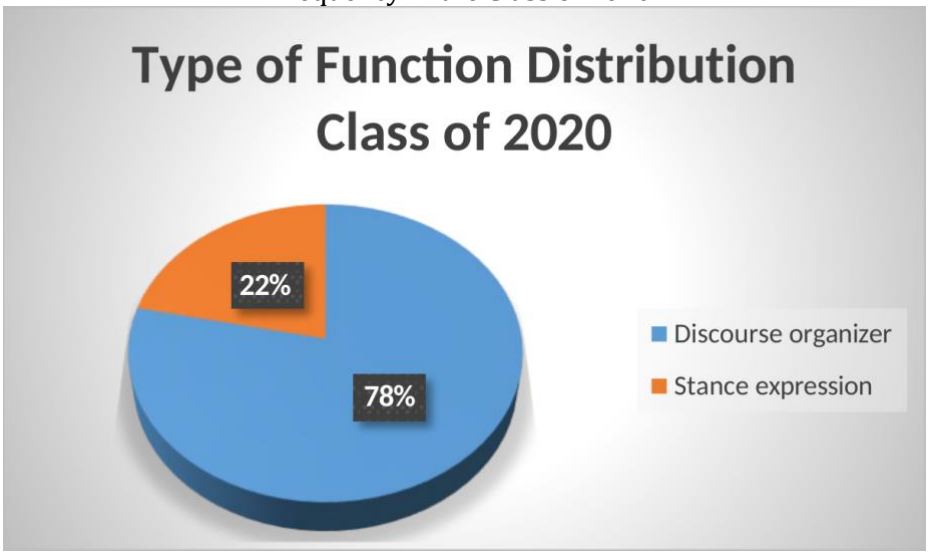
Class of 2021 examples show that the writer is identifying something as part of a larger group or category and it is also possible hedging a claim about competing values. These concordance lines demonstrate that fifth-semester students use VP bundles not only to present ideas and conclusions, but also to express different perspectives and indicate

Functional Categories of RMS in Students' Writing

The second research question examines how students use Recurrent Multi-word Sequences in terms of their discourse functions in writing. Following the functional taxonomy adopted from Biber et al. (2004) as adapted in this study, all identified RMS were classified into two functional categories: Stance Expressions (SE) and Discourse Organizers (DO). Stance Expressions are sequences that encode the writer's epistemic, attitudinal, or evaluative position toward the information being presented.

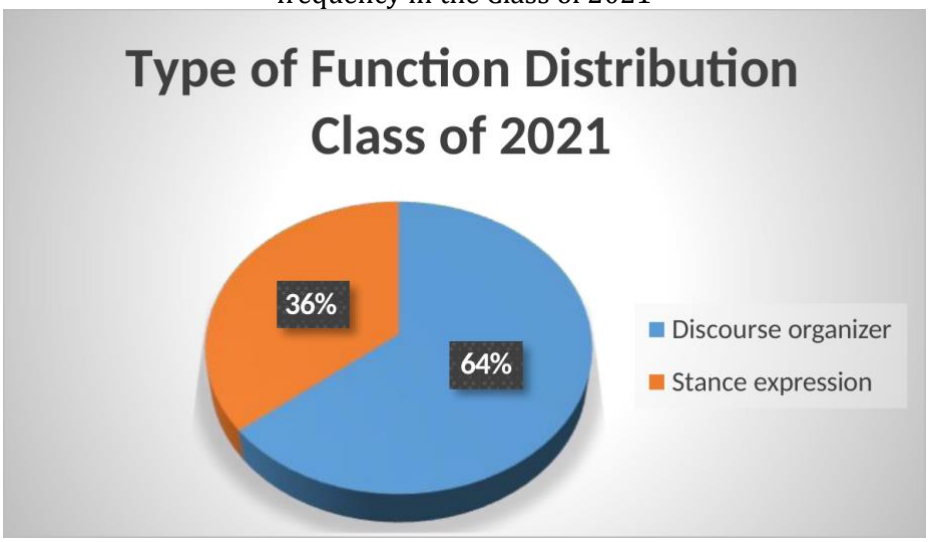
Discourse Organizers, on the other hand, are sequences that primarily serve to structure the flow of information, connect textual units, or guide the reader through the logical organization of the text. Where the structural analysis in Section 3.2 asked what grammatical form students' RMS take, this functional analysis asks what communicative work those sequences are doing a distinction that is essential, since a bundle's syntactic category does not by itself determine its rhetorical role. After identifying the functional categories of the identified RMSs, the distribution of discourse organizers and stance expressions for each cohort is presented in the following figures. These charts provide a visual comparison of the functional patterns found in the students' writing.

Figure 3.3. Distribution of Functional Categories of RMS occurrences/ frequency in the Class of 2020



ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Figure 3.4. Distribution of Functional Categories of RMS occurrences/ frequency in the Class of 2021



As illustrated in Figure 3.3, discourse organizers account for 76% of the identified RMSs in the Class of 2020, whereas stance expressions represent only 24%. This indicates that discourse organizers were used considerably more frequently than stance expressions in this cohort. The predominance of discourse organizers suggests that the students primarily employed recurring multiword sequences to organize ideas, connect information, and structure the flow of their writing, while stance expressions were used less frequently to convey personal attitudes or evaluations.

As shown in Figure 3.4, the Class of 2021 also demonstrates a higher proportion of discourse organizers, which account for 64% of the occurrences/frequency identified RMS. In comparison, stance expressions comprise 36% of the total. Although discourse organizers remain the dominant functional category, the proportion of stance expressions is relatively higher than that of the Class of 2020. Overall, the results indicate that students in both cohorts relied more frequently on discourse organizers to structure and connect ideas in their academic writing than on stance expressions to express evaluation or personal perspective.

Stance Expressions

Stance Expressions are multi-word sequences that signal the writer's subjective positioning toward the content of their writing. They may encode certainty, possibility, necessity, evaluation, or attitudinal commitment, and are therefore central to the academic practice of argumentation. Research in academic discourse has consistently demonstrated that the ability to deploy stance expressions appropriately is a key marker of writing development in EFL contexts, as it indicates the writer's growing capacity to move beyond descriptive prose toward evaluative and argumentative writing (Biber et al., 2004; Cortes, 2013).

In Class of 2020, 16 stance expression sequences were identified, with a combined frequency of 210 occurrences. The most frequent SE bundles include *children s literature is* (f=32), *it can be concluded* (f=21), *can be concluded that* (f=20), *to be able to* (f=16), *and being able to see* (f=15). The high frequency of *children s literature is* shows that students in this cohort often focused on the main topic of their writing assignments. They

frequently used this noun phrase to express their ideas and opinions about the nature and importance of children's literature. In addition, the frequent use of *it can be concluded* and *can be concluded that* shows that students were familiar with common expressions used to write conclusions in academic writing. However, they tended to rely on the same expressions repeatedly instead of using a wider variety of phrases to express their conclusions or viewpoints. Table 3.7 presents the complete list of SE-functioning RMS for Class of 2020.

Table 3.7. Frequency Distribution of Stance Expression RMS in the Class of 2020 Corpus

Recurrent Multi-word Sequence	Frequency	Range
it can be concluded	21	13
can be concluded that	20	13
to be able to	16	13
being able to see	15	3
not being able to	15	3
of writing this paper	10	5
above it can be	9	6
that can be used	9	6
This type of literature	9	3
Children s literature has	8	4

The following concordance lines illustrate how children s literature is and can be concluded that, the two dominant SE bundles identified above actually function in the writing of Class of 2020 students:

> *"...Nowadays children's literature has become public consumption and does not close the boundaries..."*
>
> *"...From the discussion of the data presented above, it can be concluded that the adaptation of literary texts for students with special needs requires careful attention to both content and format..."*

These examples show that SE bundles in Class of 2020 are primarily deployed to assert evaluative propositions about their essay topics and to signal inferential conclusions: children s literature has opened a generalized value-claim about the subject matter, while can be concluded that closes an argument built on the evidence discussed earlier in the text. The frequent recurrence of such RMS suggests that while students demonstrate an awareness of academic stance conventions, their range of stance-marking strategies remains somewhat limited and formulaic.

In Class of 2021, 8 stance expression sequences were identified, with a combined frequency of 34 occurrences. The SE bundles identified in this group including it is also possible (f=5), to be able to (f=5), can be concluded that (f=4), have the right to (f=4), is the most important (f=4), and make it difficult for (f=4) demonstrates that fifth-semester students use a varied range of stance expression. Instead of relying on only a few common expressions, they use different stance expressions to communicate their ideas more effectively. These expressions are used to show possibility (it is also possible), make evaluations (is the most important), and express rights or obligations (have the right to, make it difficult for). This suggests that fifth-semester students are able to use a variety of

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

stance expressions in their academic writing. Table 4.8 presents the SE-functioning RMS identified in Class of 2021.

Table 3.8. Frequency Distribution of Stance Expression RMS in the Class of 2021 Corpus

Recurrent Multi-word Sequence	Frequency	Range
One of the most	7	6
For a long time	5	4
can be concluded that	4	4
It is also possible	5	3
have the right to	4	3
is the most important	4	4
it can be concluded	4	4
To be able to	5	4

Concordance lines from Class of 2021 illustrate varied evaluative functions of two representative SE bundles, to be able to and make it difficult for, in this cohort's writing:

> "...Having good and correct language skills can encourage a child to be able to continue reading books independently as they grow older..."
>
> "...However, unnecessary changes to the register of a text will make it difficult for readers to follow the flow of the dialogue..."

The phrase to be able to frames a developmental capability and make it difficult for expresses a negative evaluative judgment about a specific writing choice both markedly analytical uses of stance. These concordance lines reflect a argumentatively engaged and evaluatively deployment of stance expressions among fifth-semester students, suggesting competence in using multi-word sequences to convey intellectual positioning and critical judgment. Having examined Stance Expressions in both cohorts, the discussion now turns to the second functional category, Discourse Organizers.

Discourse Organizers

Discourse Organizers are multi-word sequences that help organize ideas and connect different parts of a text. They guide readers by showing how information is related and how the text is structured. These sequences can be used to introduce topics, show transitions, indicate order or contrast, refer to previous information, and connect sentences or paragraphs. In students' writing, the use of Discourse Organizers reflects their ability to organize ideas clearly and logically. A wider and more appropriate use of these sequences generally indicates that students are able to produce well-structured and coherent academic writing that is easy for readers to follow (Halliday & Hasan, 1976; Biber et al., 2004).

In Class of 2020, 62 discourse organizer sequences were identified, with a combined frequency of 763 occurrences, making this the dominant functional category for this cohort by a wide margin. The most frequent DO bundles include *in the form of* (f=69), *of children s literature* (f=28), *literature and creative writing* (f=26), *is one of the* (f=25), and *the results of the* (f=19). The sheer number and frequency of DO sequences in Class of 2020

suggests that students rely heavily on multi-word sequences to anchor their writing to specific topics and to structure information within and across paragraphs.

However, the topically bound nature of many of these sequences such as *blind junior high school* (f=22*), the blind junior high* (f=18), and *students with special needs* (f=18) indicates that a significant portion of the discourse organizing function is driven by subject-specific vocabulary repetition rather than generalizable academic organizational patterns. Table 3.9 presents a selection of the most representative DO-functioning RMS identified in Class of 2020.

Table 3.9. Frequency Distribution of Discourse Organizer RMS in the Class of 2020 Corpus

Recurrent Multi-word Sequence	Frequency	Range
in the form of	69	26
is one of the	25	19
the results of the	19	10
in the learning process	18	9
Is a type of	16	7
There is also a	16	8
Is a form of	15	8
On the other hand	14	10
In the process of	11	8
At the same time	8	5

To this pattern in authentic text, the following concordance lines show in the process of and the results of the DO bundle, respectively in their original context:

> *“The Region tribe is one of the areas in the northern part of Central Region Province near the Region border..."*
>
> *"...The writers also realize that there are still many mistakes in the process of writing this paper, particularly with respect to presenting the results of the analysis in a fully systematic manner..."*

These examples show that the DO bundles used by the Class of 2020 have two main functions. First, they help students keep the discussion focused on the main topic, as shown by in the process of. Second, they help introduce or connect information, as seen in the results of the, which is used to present the students' research findings. This pattern is consistent with the frequent use of on the other hand (f=14), which helps students show contrast between ideas. The frequency of topic-related sequences suggests that students mainly organize their writing by referring to the specific topic of the assignment. The next section examines the pattern found in the 2021 cohort.

In the Class of 2021, 10 discourse organizer sequences were identified, with a combined frequency of 61 occurrences. The DO bundles used by this group, including is one of the (f=14), on the other hand (f=8), one of the most (f=7), as well as the (f=6), and the purpose of this (f=4), are general and typical of academic writing. These sequences can be used in many different writing tasks and academic contexts. For example, the purpose of this is commonly used to introduce the aim of a piece of writing, on the other hand is used to show contrast between ideas, and as well as the is used to add supporting

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

information within an argument. Table 4.10 presents the DO-functioning RMS identified in Class of 2021.

Table 3.10. Frequency Distribution of Discourse Organizer RMS in the Class of 2021 Corpus

Recurrent Multi-word Sequence	Frequency	Range
is one of the	14	10
on the other hand	8	8
As well as the	6	4
for a long time	5	4
is an example of	5	3
an example of a	4	3
in the form of	4	4
the purpose of this	4	4

Concordance lines from Class of 2021 illustrate the conventionally academic discourse organizing functions performed by for a long time and is an example of, two representative DO bundles from this cohort:

"...This colonization lasted for a long time until it passed through World War II..."

"...The definition of status is a person's standing, position, or state in society. Middle class is an example of a person's financial status..."

The phrase for a long time is used to describe a period of time that can be applied to many different topics, while is an example of is used to introduce a general example. These concordance lines show that fifth-semester students use DO bundles not only to organize their writing but also to introduce essays, show contrast, add information, and describe time relationships.

4. Discussion

The findings presented in Section 4.1 in relation to existing theoretical frameworks and empirical research on multi-word sequences, syntactic structure function, and academic writing competence in EFL contexts. The discussion draws on foundational work in corpus linguistics (Douglas Biber et al., 2004) research on lexical bundles in academic discourse (Cortes, 2013a; Yang et al., 2022) and the theoretical framework of cohesion in written language (Halliday, M. A.K. & Hasan, 1976) to illuminate what the observed patterns reveal across the two cohorts.

Students' use of RMS in Terms of Syntactic Structure.

The findings show class of 2020 mostly used Noun Phrase (NP) sequences while Class of 2021 mostly used Verb Phrase (VP) sequences. This is meaningful because syntax is the system that shows how learners combine words into complete and organized ideas (Ha, 2022; Yang et al., 2022) The following paragraphs explain this pattern deeper, drawing on the theories of multi-word sequences and syntax.

The strong NP use in Class of 2020 fits (Douglas Biber et al., 2004) description of referential bundles, which are sequences that mainly name and describe a topic rather than build an argument. In other words, Class of 2020 students tend to build their writing by "mentioning and explaining something." Their frequent use of noun phrases shows that

they often construct their texts around a particular topic, object, concept, or piece of information. Many of these RMS are also closely connected to the topics of their assignments, such as children's literature is and students with special needs.

Therefore, these RMS do not simply appear as grammatical patterns; they help students maintain and explain the topic throughout their writing. This can also be explained through Nattinger and DeCarrico, 1994 idea that learners store and reuse ready-made phrases because doing so lowers the mental effort needed to build a sentence from scratch. Since Class of 2020 students are still in an earlier stage of academic writing, they likely lean on these ready-made noun phrases as a practical way to talk about their topic. This supports the view that formulaic phrases help reduce the cognitive load of writing, especially for less experienced writers (Wray, 2013)

This pattern becomes clearer when looking at the actual bundles' students used. Sequences such of children s literature and children s literature is, the two most frequent NP bundles in this cohort, work mainly to name or point to something rather than to argue a point. This topic-centered pattern helps explain why NP has a higher percentage in Class of 2020. Rather than mainly using RMS to express what they think about the information, students tend to organize their writing by repeatedly referring to the concepts or information they are discussing. According to (Yang et al., 2022) this kind of NP-heavy use is typical of descriptive and content-focused writing rather than analytical writing.

In contrast, Class of 2021 showed VP dominance, with sequences such as it can be concluded, it is also possible, and to be able to appearing frequently. This suggests that Class of 2021 verb phrases to "do something with the information." For example, these sequences allow students to express that something may happen, that something is important, or that something causes difficulty. They can also use expressions such as on the other hand to connect and compare ideas. Thus, their writing present how they express a view about that information, showing its relationship with other ideas, and making an evaluation. These sequences do an argumentative and analytical language. This fits Cortes, (2013, 2015) observation that the native-like and flexible use of recurring verbal sequences.

Looking closer at the VP sequences also shows a range of communicative purposes in Class of 2021. Bundles like is the most important and have the right to add an evaluative or judgmental tone, while it is also possible expresses possibility rather than certainty. Therefore, the high percentage of VP in Class of 2021 can be understood not simply as a difference in grammatical structure, but as a reflection of how students use RMS to perform different communicative actions in their writing. They use the RMS to express relationships, evaluation, possibility, and personal positioning. This variety suggests that fifth-semester students are using verb phrases to reason and evaluate (Cortes, 2013a; Douglas Biber et al., 2004)

Compared with previous research, an important contribution of the present study is that it does not only identify recurrent multi-word sequences according to their frequency, but also shows how their syntactic structure is related to the way EFL students construct academic meaning. Previous studies have shown that formulaic sequences are associated with academic writing proficiency, rhetorical functions, and the development of phraseological competence (Appel et al., 2024; Casal & Yoon, 2023; Yan et al., 2024) For example, (Appel et al., 2024) found that higher-proficiency L2 writers used more phrase-frames in their academic essays, while (Yan et al., 2024) found that more advanced EFL

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

learners used more varied phrase-frames and showed greater control of their structures and functions.

The present study adds to this body of research by showing the use of NP-dominant pattern in Class of 2020 and a VP-dominant pattern in Class of 2021, and by connecting these structural patterns with different ways of constructing academic discourse. In particular, the NP-dominant pattern is mainly associated with referring to and describing information, whereas the VP-dominant pattern provides resources for expressing possibility, evaluation, conclusion, and relationships between ideas. Therefore, the important finding is not simply that the two cohorts use different grammatical forms, but that their RMS choices appear to reflect different use of RMS to perform academic communicative purposes. This extends previous RMS research by bringing together the three dimensions of RMS, syntactic structure, and academic communicative use in an Indonesian EFL student-writing context (Appel et al., 2024; Casal & Yoon, 2023; Yan et al., 2024)

The findings also have direct implications for EFL academic-writing pedagogy. The results suggest that RMS should not be taught only as vocabulary items or fixed expressions that students memorize. Instead, they can be taught as functional building blocks that help students perform specific academic writing purposes. For example, NP-based sequences such as in the form of can be introduced when students learn how to describe, classify, or introduce information, while VP-based sequences such as it can be argued that, this suggests that, and it is possible to can be introduced when students learn how to make claims, express possibility, interpret information, and develop arguments.

This approach is supported by recent research showing that formulaic sequences can serve rhetorical purposes and can be incorporated into corpus- and genre-based L2 writing instruction ((Casal & Yoon, 2023) Similarly, (Appel et al., 2024) found a relationship between the use of phrase-frames and L2 academic writing proficiency, while (Yan et al., 2024) showed that learners' use of formulaic patterns becomes more varied and structurally controlled as proficiency increases.

These findings suggest that EFL writing instruction can benefit from explicitly helping students notice not only what an RMS looks like, but also what the RMS does in an academic text. Therefore, teachers could use a four-step approach: first, students notice frequent RMS in authentic academic texts; second, they identify the syntactic structure and communicative function of each sequence; third, they practice using the sequences in guided sentences or short paragraphs; and finally, they apply them independently in their own academic writing. Such instruction would help students move beyond memorizing phrases and toward using RMS appropriately to describe information, connect ideas, express evaluation, and construct arguments.

This is particularly relevant to the present findings because the two cohorts indicates that students may need earlier and more explicit exposure to VP-based academic frames that support analytical and argumentative writing. Recent work on RMS learning also emphasizes the value of meaning-focused activities for developing learners' knowledge of RMS, supporting the integration of RMS instruction into meaningful writing tasks rather than isolated memorization activities (Puimège, 2024). Therefore, for teaching, this suggests that introducing useful VP frames, such as it can be argued or this suggests that, earlier in students' academic writing training may help their skills in using RMS. Having discussed the syntactic dimension of RQ1, the discussion now turns to the

second research question, which looks at the discourse functions of RMS in students' writing.

Students' Use of RMS in Terms of Functions

The findings show that both cohorts relied more on Discourse Organizer (DO) sequences than on Stance Expression (SE) sequences. This connects to the concept of cohesion, which refers to the grammatical and lexical ties that hold a text together as a unified whole (Halliday, M. A.K. & Hasan, 1976) The following paragraphs explain what this pattern reveals about students' functional competence, drawing on the cohesion and RMS theories.

The strong presence of DO sequences in both cohorts is consistent with earlier corpus-based studies on student academic writing (Cortes, 2013a; Douglas Biber et al., 2004) Discourse organizers work to connect ideas and guide readers through a text, which is a central part of what makes writing cohesive (Halliday, M. A.K. & Hasan, 1976; Ken Hyland, 2008) Because both cohorts rely heavily on DO sequences, this shows that students at both levels already understand the basic need to organize and connect their ideas.

In Class of 2020, many DO sequences are tied closely to the assignment topic, such as in the form of, is one of the, and is a form of. When we look more closely at these RMS, many of them remain strongly connected to particular topics. This means that Class of 2020 students do use RMS to organize information, but they often do so by repeating or maintaining information that is directly related to the topic of their writing. In other words, these sequences work more like topic vocabulary than general organizing tools. This matches (Cortes, 2013) distinction between bundles that are specific to one subject area and those that can be used more broadly. The findings also suggest that part of this organizing function is driven by the repetition of topic-specific vocabulary. This also reflects a common issue in EFL cohesion, where learners may repeat the same lexical items too often instead of varying their cohesive devices (Ahmed, 2010) This suggests that third-semester students' organizing their writing depends heavily on how familiar they are with the topic.

Class of 2021, on the other hand, used DO sequences that are general and could apply to almost any topic, such as on the other hand, the purpose of this, and for a long time. This suggests that, the DO used in Class of 2021 show RMS using in general academic pattern. These students appear to use RMS as tools for organizing their arguments. These sequences work as genuine organizational tools, since they can introduce contrast, state purpose, or mark time in any kind of academic text. For example, on the other hand can be used to compare ideas or introduce another point of view, the purpose of this can be used to explain a purpose, and is an example of can be used to introduce an example. Therefore, these RMS can be applied across different topics rather than being closely tied to one particular assignment. This fits (Halliday, M. A.K. & Hasan, 1976) view that using varied and appropriate cohesive devices is a sign of a mature writing style.

Turning to Stance Expressions, the RMS found in Class of 2021 is meaningful because stance markers show a writer's opinions, judgments, and level of certainty (Ken Hyland, 2008). RMS such as it is also possible, is the most important, and make it difficult for reflect the writer's evaluation of the content. On the other hand, class of 2020 frequently used expressions such as it can be concluded and can be concluded that. This shows that these students know how to express a conclusion in an academic way, but they tend to rely on

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

the same patterns repeatedly. For Class of 2021, they used varied expressions, such as it is also possible, is the most important, have the right to, and make it difficult for. These expressions allow students to show different meanings, including possibility, evaluation, importance, rights, and judgment. This kind of evaluative language is associated with an academic voice and strong argumentative skills (Douglas Biber et al., 2004) This suggests that fifth-semester students are using the RMS to comment on and judge as well. In addition, the range of stance expressions in Class of 2021 suggests that these students are using the RMS to state conclusions, to express possibility, evaluation, importance, and judgment.

In Class of 2020, the frequent use of it can be concluded and can be concluded that (f=21 and f=20) shows that these students are signaling conclusion. However, the repeated use of these expressions also shows that their way of expressing stance is still limited and formulaic. They have limited RMS for showing different types of evaluation or positioning. However, because these two phrases are repeated so often on their own, they function like a fixed formula than a flexible stance tool. This matches the idea that overreliance on the same formulaic sequence, without variation, can limit how naturally and effectively a writer expresses different shades of meaning. (Ken Hyland, 2008; Wray, 2013)

An important contribution of the present study, when compared with previous RMS research, is that it shows that the RMS use is not only a matter of how many sequences students use, but also of how flexibly they use those sequences for different purposes. Recent research on EFL phrase-frames similarly shows that learners use varied formulaic patterns and employ them for a range of functions, while lower level learners show more limited use of referential and discourse-organizing expressions (Yan et al., 2024) The present study adds to this finding by showing a similar functional difference in an Indonesian EFL student-writing context, but with a findings that topic-bound and broadly applicable discourse organization, as well as between repeated conclusion formulas and more varied stance expressions. Thus, the important finding is that Class of 2020 and Class of 2021 do not simply about the amount of DO or SE sequences they use, also the flexibility of their RMS use. Class of 2020 tends to use RMS to maintain the topic and repeatedly signal conclusions, whereas Class of 2021 shows a functional RMS for organizing ideas, expressing possibility, making evaluations, and positioning claims. This extends previous research can show how functional flexibility in RMS use may provide a detailed picture of students' academic writing (Yan et al., 2024; Zhou & Li, 2024)

Considering these functional findings point to the structural findings discussed in Section 4.2.1. Fifth-semester students use varied of stance expressions and applicable discourse organizers, while third-semester students rely on repeated, topic-bound sequences. Both patterns reflect the gain experience with academic writing, their use of RMS that topic-specific phrases also flexible, evaluative, and generalizable.

These findings have important pedagogical implications for EFL academic-writing instruction. The results suggest that RMS should be taught not simply as lists of expressions to memorize, but as functional resources that help students perform specific academic actions. For example, teachers can introduce on the other hand as a resource for creating contrast, the purpose of this for explaining purpose, it can be concluded for signaling a conclusion, and it is also possible for expressing possibility. Students should therefore be taught to recognize the function of an RMS before practicing how to use it in their own writing. One possible classroom approach is to use a four-stage cycle: noticing,

functional analysis, guided practice, and independent use. First, students identify frequent RMSs in authentic academic texts. Second, they discuss what each sequence does in the text, such as organizing, contrasting, concluding, evaluating, or expressing certainty and possibility. Third, they practice using the sequences in short sentences and paragraphs. Finally, they apply them independently when writing their own academic texts.

This type of instruction is supported by recent research showing that formulaic language can be incorporated into corpus- and genre-based L2 writing pedagogy and can help learners reflect on how formulaic patterns signal rhetorical intentions (Casal & Yoon, 2023) A recent review also shows that formulaic sequences can be learned through meaning-focused activities, supporting the use of meaningful writing tasks rather than isolated memorization (Puimège, 2024) Therefore, the present findings suggest that EFL academic-writing teachers should explicitly teach the function and appropriate context of RMS, especially discourse-organizing and stance sequences, so that students can develop a wider and more flexible repertoire for organizing arguments and expressing their positions.

This has a practical implication for teaching, since explicitly noticing and practicing different RMS functions in class, rather than leaving them to be picked up naturally, can help students build this flexibility earlier (Cortes, 2013a; Ucar, 2017) In particular, the present findings suggest that instruction should give students an early exposure to a range of stance and discourse-organizing RMS, rather than relying heavily on a small number of repeated formulas.

5. Conclusion

This This study examined the use of Recurrent Multi-word Sequences (RMS) in the academic writing of third-semester students from the Class of 2020 and fifth-semester students from the Class of 2021 at Universitas Negeri Gorontalo. The study focused on two aspects: syntactic structures, namely Noun Phrases (NP) and Verb Phrases (VP), and discourse functions, namely Discourse Organizers (DO) and Stance Expressions (SE). The findings showed differences in how the two groups used RMS to construct meaning and organize their academic writing.

In terms of syntactic structures, the Class of 2020 mainly used Noun Phrase (NP) sequences, which accounted for 66.7% of their total RMS frequency. These sequences were mostly used to identify, describe, and provide information about the topics discussed in their writing. In contrast, the Class of 2021 predominantly used Verb Phrase (VP) sequences, accounting for 66.3% of their total RMS frequency. These sequences helped students explain ideas, present conclusions, express possibilities, and evaluate information. This difference suggests that the Class of 2020 tended to use RMS to describe and refer to specific topics, while the Class of 2021 used them more to express actions, arguments, and evaluations.

Regarding discourse functions, Discourse Organizers were the most frequently used category in both groups. They accounted for 79.5% of RMS in the Class of 2020 and 64.2% in the Class of 2021. These sequences helped students connect ideas, organize information, and maintain the flow of their writing. However, the use of Stance Expressions increased from 20.5% in the Class of 2020 to 35.8% in the Class of 2021. This indicates that the fifth-semester students used RMS more frequently to express opinions, judgments, possibilities, and evaluations. Examples such as on the other hand, it is also possible, and is the most important show how RMS can help writers construct arguments and communicate their perspectives.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Overall, the findings indicate different patterns of RMS use between the two groups. The Class of 2020 relied more on noun-based and topic-related expressions, whereas the Class of 2021 showed greater use of verb-based sequences and expressions that conveyed evaluation and stance. Although these differences do not necessarily mean that all fifth-semester students have better writing abilities than third-semester students, they suggest a shift in how RMS was used across the two groups.

Therefore, RMS plays an important role in academic writing because it helps students organize information, connect ideas, construct arguments, and express their viewpoints. Understanding these patterns can also help students become more aware of how recurring word sequences contribute to clear and effective academic writing.The study focused on two syntactic structures, namely Noun Phrase (NP) and Verb Phrase (VP), and two discourse functions, namely Discourse Organizers (DO) and Stance Expressions (SE).

6. References

Ahmed, A. H. (2010). Students' Problems with Cohesion and Coherence in EFL Essay Writing in Egypt: Different Perspectives.

Alasmary, A. A. (2018). Multiword Sequences and Language Learning Pedagogy: Bridging the Research-Practice Gap. English Language Teaching, 12(1), 204. https://doi.org/10.5539/elt.v12n1p204

Alawerdy, A. S., & Alalwi, F. S. (2022). Enhancing English as a foreign language university students' writing through explicit instruction of conjunctions as cohesive devices: An experimental study. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.1053310

Albelihi, H. (2022). A CORPUS-BASED INVESTIGATION OF PHRASAL COMPLEXITY IN A CORPUS-BASED INVESTIGATION OF PHRASAL COMPLEXITY IN L1, L2 PHD STUDENTS, AND EXPERT WRITERS IN EDUCATION L1, L2 PHD STUDENTS, AND EXPERT WRITERS IN EDUCATION FIELD FIELD. https://digitalrepository.unm.edu/educ_llss_etds/

Aljoza, A. Y., & Almustafa, I. M. (2024). EFL Academic Writing Challenges among First Year Students: Evidence from North and East Syria. TESOL Today, 1(2), 79–90. https://doi.org/10.53898/tesol2024125

Appel, R., Geluso, J., & Feng, H. H. (2024). An examination of phrase-frames in L2 english academic writing: Exploring relationships with writing quality. System, 123. https://doi.org/10.1016/j.system.2024.103349

Barabadi, E., Robatjazi, M. A., & Bayat, M. (2020). A phraseological examination of research articles in the field of environment using key phrase frame. Eurasian Journal of Applied Linguistics, 6(1), 81–100. https://doi.org/10.32601/ejal.710217

Biber, D., Gray, B., Poonpon, K., https://dialnet.unirioja.es/servlet/autor?codigo=1781192, https://dialnet.unirioja.es/servlet/autor?codigo=2891577, & https://dialnet.unirioja.es/servlet/autor?codigo=2891578. (2011). Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development? TESOL Quarterly: A Journal for Teachers of English to Speakers of Other Languages and of Standard English as a Second Dialect, ISSN 0039-8322, Vol. 45, No. 1, 2011, Págs. 5-35, 45(1), 5–35. https://dialnet.unirioja.es/servlet/articulo?codigo=3610674

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman grammar of spoken and written English. In Perspectives in Public Health (Vol. 130, Number 5). Longman. https://research.lancaster-university.uk/en/publications/longman-grammar-of-spoken-and-written-english/

Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). Chapter 8. Complex noun phrases. Grammar of Spoken and Written English, 567–650. https://doi.org/10.1075/Z.232.C8

Casal, J. E., & Yoon, J. (2023). Frame-based formulaic features in L2 writing pedagogy: Variants, functions, and student writer perceptions in academic writing. English for Specific Purposes, 71, 102–114. https://doi.org/10.1016/J.ESP.2023.03.004

Chen, Y.-H., & Baker, P. (2010). LEXICAL BUNDLES IN L1 AND L2 ACADEMIC WRITING. 14(2), 30–49. http://llt.msu.edu/vol14num2/chenbaker.pdf

Choemue, S., & Bram, B. (2021). Discourse markers in academic and non-academic writings of thai efl learners. Studies in English Language and Education, 8(3), 1209–1226. https://doi.org/10.24815/siele.v8i3.20122

Cortes, V. (2013a). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. Journal of English for Academic Purposes, 12(1), 33–43. https://doi.org/10.1016/j.jeap.2012.11.002

Cortes, V. (2013b). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. Journal of English for Academic Purposes, 12(1), 33–43. https://doi.org/10.1016/j.jeap.2012.11.002

De Smedt, F., Landrieu, Y., De Wever, B., & Keer, H. Van. (n.d.). Title: Do Cognitive Processes and Motives for Argumentative Writing Converge in Writer Profiles? Do Cognitive Processes and Motives for Argumentative Writing Converge in Writer Profiles?

Douglas Biber, Susan Conrad, & Viviana Corted. (2004). Lexical Bundles in University Teaching and Textbooks.

Durrant, P., & Brenchley, M. (2023). Development of Noun Phrase Complexity Across Genres in Children's Writing. Applied Linguistics, 44(2), 239–264. https://doi.org/10.1093/APPLIN/AMAC032

Ellis, R. (2012). Language Teaching Research and Language Pedagogy. Language Teaching Research and Language Pedagogy. https://doi.org/10.1002/9781118271643

Ha, M. J. (2022). Syntactic Complexity in EFL Writing: Within-Genre Topic and Writing Quality 1. In Computer Assisted Language Learning Electronic Journal (CALL-EJ) (Vol. 23, Number 1).

Halliday, M. A.K. & Hasan, K. (1976). 1976_Cohesion in English .pdf. 374.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Hamdani, M. H., & Abid, N. (2025). INDONESIAN EFL STUDENTS' WRITING CHALLENGES: A QUALITATIVE META-ANALYSIS. Getsempena English Education Journal (GEEJ), 12(1), 14–36. https://ejournal.bbg.ac.id/geej

Hardiyanti, I., & Mahading, P. (2024). Individual Diversity in English as Foreign Language Development: Aptitude, Exposure, and Behaviour.

Ismi, N., & Anggreni, A. (2025). Cohesion and Coherence in EFL Academic Writing: A Case Study of Tadulako University. Journal on Interdisciplinary Studies in Humanities, 8, 2025. https://doi.org/10.34050/elsjish.v8i2.45245

James R. Nattinger and Jeanette S. DeCarrico. (1994). LEXICAL PHRASES AND LANGUAGE TEACHING Oxford: Oxford University Press, 1992. Pp. xvi + 218. Studies in Second Language Acquisition, 16(2), 254–254. https://doi.org/10.1017/S0272263100013000

John W. Cresswell. (2017). Research Design. https://www.sagepub.com/shop/buy-a-book/research-design-5-255675

Ken Hyland. (2008). As can be seen: Lexical bundles and disciplinary variation. https://studylib.net/doc/8902747/as-can-be-seen--lexical-bundles-and-disciplinary-variation?utm_source=chatgpt.com

Li, Y. K., & Fang, A. B. (2022). Synergistic effects of multiword sequences structure, function, frequency and association on raters' evaluations of essay quality. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.1026658

Mariel Mangaya, P. G. (2024). EXAMINING DISCOURSE STRATEGIES AND STANCE MARKERS IN ESL STUDENTS' REFLECTIVE ESSAYS PSYCHOLOGY AND EDUCATION: A MULTIDISCIPLINARY JOURNAL Examining Discourse Strategies and Stance Markers in ESL Students' Reflective Essays. https://doi.org/10.5281/zenodo.10966166

Miles, M. B., & Huberman, A. M. (1994). Qualitative Data Analysis. An Expanded Sourcebook, 3, 338pp. https://doi.org/10.1080/0140528790010406

Norahmi, M., Dona Retsi, O., Ira Nova, S., Martha Meka, L. C., Haryani, T., Rahmaniah, S., Flory Aristia, C., Munawarah, S., & Fitriyani, W. (2026). A DESCRIPTIVE ANALYSIS OF NOUN PHRASES IN EFL STUDENTS' ESSAYS. In Journal of English Educational Study (JEES (Vol. 9).

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. Annual Review of Applied Linguistics, 32, 130–149. https://doi.org/10.1017/S0267190512000098

Puimège, E. (2024). Learning L2 formulaic sequences from meaning-focused activities. ITL - International Journal of Applied Linguistics, 175(2), 163–186. https://doi.org/10.1075/itl.23014.pui

Quoc Tuan, N., Thi Bao Trang, N., & Vu Quynh Nhu, N. (2023a). USE OF COHESIVE DEVICES IN PARAGRAPH WRITING BY EFL STUDENTS AT ENGLISH LANGUAGE CENTERS IN VIETNAM. In VNU JOURNAL OF FOREIGN STUDIES (Vol. 39, Number 3).

Quoc Tuan, N., Thi Bao Trang, N., & Vu Quynh Nhu, N. (2023b). USE OF COHESIVE DEVICES IN PARAGRAPH WRITING BY EFL STUDENTS AT ENGLISH LANGUAGE CENTERS IN VIETNAM. In VNU JOURNAL OF FOREIGN STUDIES (Vol. 39, Number 3).

Rahma Annisa, D., & Opinia, L. (2024). Enhancing EFL Learners' Writing Skills: Challenges, Strategies, and Pedagogical Insights. https://ejournal.baleliterasi.org/index.php/JEs

Schneider, E. W. (2001). Review of Biber, Johansson, Leech, Conrad & Finegan (1999): Longman Grammar of Spoken and Written English. English World-Wide, 22(1), 137–143. https://doi.org/10.1075/EWW.22.1.11SCH

Sinclair, J. (1991). Corpus, concordance, collocation. Oxford: Oxford University Press. 179. https://books.google.com/books/about/Corpus_Concordance_Collocation.html?id=Vcx5AAAAIAAJ

Sugiyono. (2023). METODE PENELITIAN KUALITATIF . https://inlislite.ipdn.ac.id/opac/detail-opac?id=21562&utm_source=chatgpt.com

Thao, T. Q., & Bon, P. Van. (2023). English Majors' Difficulties in Using Phrasal Verbs in Academic Writing. VNU Journal of Science: Education Research. https://doi.org/10.25073/2588-1159/vnuer.4738

Thi, N. K., Vo, D. Van, & Nikolov, M. (2023). Investigating syntactic complexity and language-related error patterns in EFL students' writing: corpus-based and epistemic network analyses. Language Learning in Higher Education, 13(1), 127–151. https://doi.org/10.1515/cercles-2023-2009

Ucar, S. (2017). A Corpus-based Study on the Use of Three-word Lexical Bundles in the Academic Writing by Native English and Turkish Non-native Writers. English Language Teaching, 10(12), 28. https://doi.org/10.5539/elt.v10n12p28

Wang, W., Chen, W., Shi, S., Tong, L., & Duan, M. (2024). A Corpus-Based Study on Cohesion Development in Writing by EFL Novices. Journal of Language Teaching and Research, 15(3), 772–786. https://doi.org/10.17507/jltr.1503.10

Wray, A. (2013). Formulaic language. Language Teaching, 46(3), 316–334. https://doi.org/10.1017/S0261444813000013

Yan, J., Li, Q., & Jiang, J. (2024). The development of phrase-frames in EFL learners' essays: Variability, structures, and functions. International Journal of Applied Linguistics (United Kingdom), 34(4), 1689–1708. https://doi.org/10.1111/IJAL.12589;WGROUP:STRING:PUBLICATION

Yang, Y., Yap, N. T., & Ali, A. M. (2022). A Review of Syntactic Complexity Studies in Context of EFL/ESL Writing. International Journal of Academic Research in Business and Social Sciences, 12(10). https://doi.org/10.6007/ijarbss/v12-i10/14843

Yue, C. (2024). We make viewpoints in different ways: the use of stance adverbials across academic levels and disciplines. Cogent Arts & Humanities, 11(1). https://doi.org/10.1080/23311983.2024.2435169

Zhou, H., & Li, T. (2024). "In the present study": An exploration of prepositional phrase-frames in Chinese EFL learners' theses and dissertations of applied linguistics. Journal of English for Academic Purposes, 70. https://doi.org/10.1016/J.JEAP.2024.101400