

Formulaic language use of advanced English learners in writing

January, 2023

ABSTRACT
FORMULAIC LANGUAGE USE OF ADVANCED ENGLISH LEARNERS IN WRITING

January, 2023

Vocabulary is a significant constituent of the language learning process: lexical phrases, in particular, are important when it comes to developing learners' language skills. Since formulaicity is a complex topic, not enough research has been done to shed light on some of its features and there are few articles on the connection between FL (formulaic language) and writing. In this study, I analyze formulaic language use in the writing of four advanced English language learners and explore their habits, awareness and attitudes towards using it via text analysis of their written works and consecutive interviews. The findings reveal that the students utilize a high number of formulaic elements; however, they are not always conscious of FL use. As for the pedagogical implications, more emphasis should be put on these language elements when teaching.

Keywords: formulaic language use, multiword elements, EFL writing, text analysis

TABLE OF CONTENTS

ABSTRACT	2
CHAPTER 1: INTRODUCTION.....	4
CHAPTER 2: REVIEW OF LITERATURE.....	5
Definition of formulaic language.....	5
Significance and functions of formulaic language	5
Formulaic language in Language Teaching	6
Formulaic language classification	8
CHAPTER 3: METHODOLOGY	10
Research aim and question	10
Setting and participants.....	10
Research design	12
Instruments and materials	13
Data collection procedure	13
Data analysis	14
CHAPTER 4: RESULTS AND DISCUSSIONS	16
The results of text analysis	16
The interview results.....	19
Results of students' habits	19
Results of students' awareness.....	21
Results of students' attitudes towards FL	10
CHAPTER 5: CONCLUSION	25
Findings and discussions	26
Limitations of the study	26
Pedagogical implications	26
Suggestions for future research	26
REFERENCES	27
APPENDICES	25

1 Introduction

Vocabulary is crucial in English as a Second Language (ESL) learning because “it underpins all other language skills”. A wide range of lexical expressions can help a student to master English more quickly (Egamova & Sharofova, 2022, p. 190). Vocabulary accounts for various elements, but one of the main ones is formulaic language (FL). The frequency of FL is approximately 80% in general (Altenberg, 1998) and 50% in writing (Erman & Warren, 2000), which is the focus of this study. Thus, FL use is significant for learners of English, especially at advanced levels. However, at the same time, it also appears to be extremely hard to acquire (Lundell, 2021) primarily due to a misconception: ESL teachers and students tend to think about vocabulary in terms of singular words instead of phrases (Schmitt, 2010) and it leads to a lack of emphasis on it during education.

The rationale behind focusing the study on writing is that writing may be more complex than other language skills. It involves logical and analytical competencies to make a text coherent and the ability to use appropriate grammar and vocabulary structures simultaneously (Kilich, 2015). Therefore, obtaining new data about FL in writing appears to be of great linguistic and pedagogical importance. After intense research in applied linguistics over the last decade in terms of FL, showing its significance and offering the latest recommendations with regard to writing, this study will explore the habits, awareness and attitudes of ESL learners and provide lexical analysis of their texts, which indicates the current situation and may be applied to future studies.

The aim of the study is to reveal how much FL the subjects of the study used in their writing and also to give an insight into their habits, attitudes and consciousness. First, relevant theoretical issues related to FL use and writing are discussed in the literature review, followed by the research questions. Finally, conclusions are drawn, along the study is presented afterwards.

2 Review of literature

The review of the literature part introduces the definition of FL, its significance and functions in general and writing mainly. The pedagogical implications based on various studies are also listed in this chapter. Additionally, the classification of FL elements used in the study is demonstrated.

2.1 Definition of formulaic language

Applied linguistics has provided a new perspective on languages: a significant part of their account for specific patterns or *formulas*. Currently, there are multiple definitions of FL. For Sinclair (1991, p. 110), expressions such *sooner or later* or *make a decision* refer to “single choices, even though they might appear to be analyzable into segments”. Wood (2020, p. 30) broadens the term. It states that FL is a “multiword language phenomena which holistically represent a single meaning or function, and are likely mentally stored and used as unanalyzed wholes, as are single words”. Erman and Warrens (2000, p.31) account for the above definitions but also highlight that there are native speakers who primarily use multiword constructions: “A combination of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization”. As can be seen, the concept of FL is vague, but there are overlapping terms.

Based on the existing literature, the aforementioned definitions include three key elements: unity, mentality and nativity. Firstly, it is evident that formulaic sequences are inextricably linked to the agreement of different words, both lexical (nouns, verbs, adjectives, and adverbs) and functional (conjunctions, prepositions, articles, pronouns), as seen in the examples. Secondly, a mental concept appears because the term is also related to brain function, namely memorizing and speech production. Thirdly, FL expressions are primarily used by native speakers of a language and, therefore, it poses a problem for its learners – in the case of this research, English.

2.2 Significance and functions of formulaic language

The complexity of the term indicates its significance. Although scholars classify FL elements differently (see 2.4), they still provide a large percentage of the frequency of the phrases. As an example, according to Altenberg (1998), formulaic expressions account for over 80% of the vocabulary of an adult native speaker and thus make up a considerable part of English. Narrowing the topic to research, FL has been acknowledged to account for 52.3 per cent of L1 written English discourse (Erman & Warren, 2000). These statistics refer to the authentic nature of the FL phenomenon.

Since the concept is of paramount importance, multiword sequences serve essential functions for ESL learners. Martinez and Schmidt (2012) listed some of them in the article. The first reason is that FL “phrases realize many referential, communicative, and textual functions in discourse”. In writing, these are expressing concept (*take into account* = considering something), signposting discourse organization (*on the other hand* = signaling transition) and efficiently conveying ideas by including technical phraseology in the text (using a *figure of speech* in a non-literal environment) (Schmidt & Carter, 2004). For ESL learners, such phrases may help them to express concepts comprehensibly and explicitly. Besides, FL sequences can realize the writer’s needs in various contexts, including academic and scientific ones (Biber et al., 2004). Therefore, using multiword phrases might provide a person with a relevant advanced vocabulary in different linguistic scenarios.

Another reason multiword terms are beneficial in writing is that they facilitate understanding a text written by a non-native person. Millar (2011) found a connection between the written production of L2 learners and its processing by native speakers of English. Thus, if a learner used non-conventionalized formulaic elements, L1 readers needed much time to analyze them. Research on self-paced reading tasks also proved that FL phrases are processed more quickly than their non-formulaic analogues by L1 and L2 readers (Conklin & Schmidt, 2008). By including FL elements in texts, ESL learners have more chances to be understood correctly by readers, in particular native speakers.

Finally, using FL expressions affects the impression of ESL learners’ language production. There is a proven connection between FL use and language proficiency, especially “at advanced and very advanced levels”, as it gives a student more confidence to express himself in the language (Lundell, 2021). For instance, those who scored higher on an ESL proficiency test used more FL constructions (Ohlrogge, 2009). In addition, L2 speakers seemed more advanced when using multiword elements in writing (Boers et al., 2006). To conclude, if a learner consciously includes more FL phrases in texts, that might indicate a higher English level; it can also make a better impression on a reader.

2.3 Formulaic language in Language Teaching

Nevertheless, FL is ubiquitous in the language; it takes much time to master for English learners and appears extremely difficult (Lundell, 2021). The reasons for this problem have yet to be sufficiently studied, and one of the research goals was to collect more data. Based on what is available now, there were pedagogical and accessibility issues ten years ago. Firstly, most EFL teachers and learners think about vocabulary in terms of single words and not multiword elements

(Schmitt, 2010), although they are ubiquitous in the language and perform essential functions. It can be seen that an incomplete picture of vocabulary elements and a lack of emphasis on FL may lead to significant issues for ESL learners. Secondly, since this concept of language needed to be taken into proper attention, virtually no classroom materials capitalizing on FL elements were available for commercial purposes (Wood, 2002 as cited). Two decades ago, there were only so many educational sources for teachers and students to learn more.

However, after realizing the contradiction between the importance of FL and the lack of appropriate teaching and learning methods, some research regarding pedagogical implications has been done. For example, Meunier (2012) compared the situation in English language teaching a decade later, after Wood's findings. It became clear that education has taken a step forward in this regard: digital sources now provide access to authentic materials and, thus, FL. The author also referred to another research in which cutting-edge activities promoting the role of multiword phrases were suggested – *shadowing* and *dictogloss*. *Shadowing* is reading out loud a text simultaneously with a speaker. *Dictogloss* is a task during which learners collect keywords from a dictated text and then work in groups aiming to reconstruct its gist (Ellis, 2008, as cited). These activities help students find separate units of multiword expressions and then try to create whole phrases.

Regarding the issues that need to be addressed, the author lists questions about categories of FL that should be emphasized more, high or low frequent ones, and efficient tasks for learning and practising these sequences. Moreover, it has been stated that there is still a gap between digitally developed and lagging in this respect countries, which impacts access to new educational methods. In summary, the study reflects the use of FL in general and not particularly in writing, demonstrating a positive tendency in education.

In terms of the connection between FL use and writing skills, Kilich (2015) researched the exam conditions where the students needed to perform an opinion essay. The findings showed no significant relationship between FL use and students' coherence and total writing scores. Furthermore, it contradicted the previous research results showing that FL phrases positively contribute to the final results of proficiency exams (see 2.2). The reason for that might be that other essential linguistic elements, such as good content grammar and vocabulary, impact the quality of writing in conjunction with FL. However, research does not diminish the importance of multiword phrases; the point was to highlight that FL should not be considered separately from the rest of linguistic objects. Concerning the findings about the general use of FL, learners tend to use those expressions that were often repeated and correctly presented in learning sources.

Let us now turn to the pedagogical implications listed in Kilich's article. Most importantly, the administration of educational establishments should carefully approach the choice of studying materials. First, it should include a wide range of multiword expressions which ESL can use in classes and proficiency tests. Then, FL should be incorporated into the curriculum along with grammar and vocabulary as they complement each other. There should also be writing models provided where it is noticeable that FL is an excellent addition to an already advanced level of English rather than its basis. Lastly, material developers should offer more activities, including writing, to help students apply the elements to grammar and meaningful context and create a complete linguistic picture for ESL learners. In conclusion, this article shows a new perspective on the approach to learning multiword sequences: they should be studied along with other elements of English.

2.4 Formulaic language classification

40 FL terms were identified 20 years ago (Wray and Perkins, as cited in Wood, 2020); the main categories are collocations, idioms, metaphors, proverbs, phrasal verbs and compounds. However, the elements may overlap (Wood, 2020).

Table 1

Main categories of FL

Category	Examples
Collocations	- Collocated pairs of words: <i>senior manager, single parent</i> - Syntactic functions: <i>red herring</i> - Pragmatic functions: <i>how are you?</i>
Idioms	<i>Spill the beans, burn one's candle at both ends</i>
Metaphors	- Vehicle (unusual manner) + topic (referent of the vehicle) scheme: <i>time is a healer</i> - Simile: <i>life is like a box of chocolate</i>
Proverbs	<i>A stitch in time saves nine, like death and taxes</i>
Compounds	- Closed forms: <i>hardcore, laptop</i> - Hyphenated forms: <i>open-handed</i> - Open forms: <i>real estate</i>
Phrasal verbs	- Verb + preposition: <i>give up, let down</i>

	- Verb + particle: <i>show off, cave in</i> - Verb + particle + preposition: <i>run out of</i>
Lexical phrases	- Poly words: <i>so far so good</i> - Institutionalized expressions: <i>nice meeting you</i> - Phrasal constraints: the ____er the ____er - Sentence builders: I think that X

(Wood, 2020)

Another crucial category is lexical bundles which are “recurrent sequences of three or more words that occur in high frequency across texts” (Biber et al., 1999). The examples vary from text to text, but linking phrases like *on the one hand* and *on the other hand*, which may be repeated, is one of them.

There are several ways to determine FL phrases, such as frequency statistics in corpora and internet search engines. The most prominent is the native-speaker checklist consisting of 5 criteria for estimating if a sequence is formulaic. Moreover, in this checklist, non-native phrases were accepted, for example, *thanks god* instead of *thank god*. However, more research needs to be done to define FL more explicitly because the FL classification, as stated before, is relatively obscure (Wood, 2020).

In conclusion, the ubiquity and significance of multiword constructions are evident but there needed to be more emphasis on it years earlier. Also, they are considered one of the most challenging components of English vocabulary, making the situation even more complicated. Nevertheless, the use of FL by ESL students must change after new research in this field and the offering of cutting-edge educational methods of its acquisition, and this research aimed to gain data about the current level of FL use by advanced learners for whom this skill is vitally important (Lundell, 2021).

3 Methodology

The methodology section, most importantly, covers the aim and questions of the study. It also provides information about its setting and participants, including their profiles, research design and the instruments used. Finally, the collection and analysis of data are likewise described in this chapter.

3.1 Research aim and questions

This research aims to assess the FL use of advanced ESL learners and refer to three relevant constructs: habits, awareness and attitudes. Thereby, the study attempts to answer the following research questions:

1. What characterizes the formulaic language use of advanced EFL learners when writing short tasks?
2. What are these advanced learners' habits, beliefs and attitudes towards their formulaic language use?

3.2 Setting and participants

The research was conducted at the Faculty of Humanities at one of the top universities in Hungary. The chosen major for the research was English and American studies BA, with English in Media specialization; the program's language is English. To enroll in this major, Hungarian applicants must prove their language skills by providing the final exam results; international applicants need to submit the standardized test results and do the interview where they are asked questions about their motivation and already gained subject knowledge in the field. It is also a good indicator of a student's English language skills. In addition, the level of English for students must be at least B2 to enter ELTE, according to the Common European Framework of Reference (CEFR).

During the first year of university, students must take academic writing and proficiency exams to continue their studies. The academic writing exam requires them to write a guided summary in 60 minutes, which two independent teachers check on a 5-grade scale; the criteria covered coherence and cohesion, content, logic, grammar, vocabulary and punctuation. A student aimed to get at least 2 in each section to pass the test. The proficiency exam, divided into two days, includes tasks in speaking in groups, multiple choice questions in vocabulary and sentence completion in grammar; the maximum score is 100. One must have at least 65 points in the oral part and 54 in the written section to pass the exam. In case of failure, it is possible to retake the tests the following year.

It was mentioned before that the research was held with English in Media students. To clarify, after the first year of BA, students can choose the area of English they want to learn more about: Media, Business or Films and Culture. To be accepted, one needs to do an interview with the professors of this discipline. The pool for each is approximately 15 students. However, there are options to pick minors from other departments, but the application process for the specializations is highly competitive.

There are numerous reasons why this university was chosen as a place for research. Firstly, it is considered the best higher institution in Hungary in the field of Humanities and many gifted students, both local and international, do their programs there. Secondly, there is strict control over students' language skills: the entrance requirements, two exams in the first year and an interview for the specializations. Finally, concerning English in Media, its students must write a lot for subjects such as Workshops and Fact-based Journalism; thus, there are excellent samples of advanced writing.

For this particular research, the Popular Science in Social Media class was decided on, as the students were required to submit short written tasks in 200-250 words at least four times a semester based on the content covered: *"As a short written task, you can write a max. 250-word article (based on anything on the handout, the videos or any other source, using new vocabulary)"*. Happiness, stress, longevity, issues of nutrition, the brain, green issues, inventions and the future, the virtual world, geography and travel journalism were among the possible topics for the tasks; for the research, *stress, longevity, issues of nutrition and the brain* were chosen for the lexical analysis. Although the main idea was to summarize and synthesize the sources, students must have expressed themselves freely in the texts, combining academic and creative skills. Therefore, there were no strict criteria for the written tasks.

Surprisingly, four students did the short written tasks almost every week and became participants in the research. Thanks to the weekly basis, tracking the effects of the topic on FL use was possible. Since the semester was becoming harder week by week due to new deadlines and midterms, it is interesting to indicate if the study load impacted students' FL production. Moreover, these four students showed eagerness in the class and studying English in general and appeared to be excellent representatives of advanced ESL learners. Profile information is listed:

Table 2

Profile of participants

	Age (years old)	Learning English experience (years old)	Year of the university (ELTE, English and American Studies BA, English in Media specialization)	Nationality
Participant 1	20	15	2	Hungarian
Participant 2	20	13	3	Hungarian
Participant 3	22	14	3	Hungarian
Participant 4	20	10	2	International

3.3 Research design

The mixed approach was used to determine the characteristics of FL use, beliefs, attitudes and awareness of advanced ESL learners. According to Johnson and Onwuegbuzie (2004), mixed methods include quantitative and qualitative research techniques within one study. Additionally, Creswell (2004) named strengthening one method by combining it with another and allowing new perspectives to emerge as the functions of a mixed-approach study and these are the reasons to use it in this study.

Firstly, the primary constructs and ideas listed before were set up. According to it, the interview questions were created first, and FL categories that needed to be found were asserted secondly. The interviews and text analysis were swapped during the data collection procedure to make the study coherent and obtain further information about using FL in particular short tasks. In addition, it played an essential role in determining if there is a relationship between FL use and other factors, such as attitude to the topic, study load and well-being, and source type. As a result, the first step in this research was to analyze the short essays of the participants on the subject of chosen categories of FL word by word. Secondly, interviews with all of the students were conducted. As mentioned, it included questions regarding the primary constructs and additional ones about the lexical analysis. Finally, the research findings were examined separately and then together to conclude the FL use of advanced ESL learners.

3.4 Instruments and materials

To do their short tasks, the participants used various online sources: *CNN*, *Neater Pets*, *The Conversation*, *Medium*, *Amen Clinics*, *ABS News*, *World Economic Forum*, *Spectrum*, etc. They

also referred to sources from academic journals, such as *the Nutrition Society*, *International Journal of Education Research* and *Cell Metabolism*. Given one field, there was no overlap between the themes. For example, in writing about eating habits, participant 1 analyzed obesity and climate change, while participant 4 proceeded with organic food. Therefore, short written tasks based on various sources were examined, and the type of source was also considered when working with the texts.

The classification provided in Wood's article (2020) was used for the text analysis. In addition, *Reverso Context* and *Cambridge Dictionary* are beneficial in terms of FL, as they provide different cases of using a phrase in the English native environment: films, books, online articles and academics. For this reason, the tools were utilized to define whether a term was formulaic and correct.

A basic phone calculator was taken to determine the frequency of FL elements in the text analysis. In the case of the interview, an online app called "Transcribe – speech to text" was downloaded to decipher the interview data. Descriptive statistics estimate the percentage of formulaic language elements for each participant. However, due to the low number of participants, only these basic descriptions are calculated, and a deeper analysis follows with the help of the interview questions.

3.5 Data collection procedure

Regarding the text analysis, every student in the group agreed to post their short written task for open access on Moodle as they were high quality and appeared engaging to read. Next, the participants were asked for permission for the text analysis and the following interview. After consent was obtained, the texts published by these four students for the same weeks were chosen for the study. It was decided to research only four texts on the topics mentioned before (see 3.2) as they already gave the complete picture of FL use over a while and answered the questions posted. Therefore, the short essays were collected in four separate documents named after the participant.

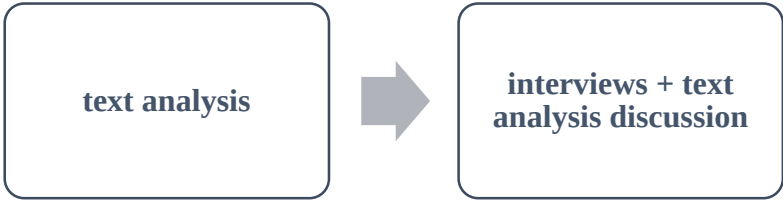
Regarding the elements of FL, general collocations, phrasal verbs, technical lexical phrases (field-specific collocations), linking words (referring to sentence builders of lexical phrases) and lexical bundles (if possible to find) were highlighted in the texts and their frequencies were calculated. These categories were mainly chosen based on the task type. As it was popular science, technical vocabulary and collocations retrieved from the sources were often used and the students also utilized numerous linking phrases to make their short essays logical. Apart from this, phrasal verbs and lexical bundles are the least obscure categories of FL, which are relatively easy to

indicate and analyze. Ultimately, the chosen categories answer the question of seeing the students' creativity when having articles or videos to base on.

The date and time for the interviews were set after the text analysis. Most of the interviews were conducted on campus, and one was done online due to the participant's high study load. The length of each interview was approximately 20 minutes and altogether four interviews were conducted. Firstly, the questions about their habits, awareness and attitudes towards FL use were asked; secondly, the results of the text analysis were shown to the participants, and additional questions about them were raised (see Appendix A for the Interview protocol). The interviews were recorded, transcribed by a phone application and collected in one file (see 3.4).

Table 3

Data collection process



3.6 Data analysis

To conclude, referring to the research design (see 3.3), the text analysis aimed to answer the first research question: to define the FL characteristics of these four advanced ESL learners in writing short tasks. That included calculating the frequency of chosen categories in the texts on four topics. The second part of the research estimated their knowledge and consciousness of FL use and the following questions on why they may use them more or less in specific cases. After the data was collected, the tables with calculations were created for each topic and each participant to see trends of FL use (see Appendix B). In terms of the interview, all the answers were synthesized to find similarities and differences in students' habits, beliefs and attitudes towards multiword phrases; based on it, the conclusions for the sections and general were drawn. Finally, the most representative answers to each construct were put into the tables.

Turning to the results and discussions, it should be mentioned that FL is hard to define, especially in writing; therefore, not enough has been done. This study sheds light on the topic and provides new findings to base on.

4 Results and discussion

This chapter presents the results of the FL use of advanced ESL learners and offers insight into their habits, awareness and attitudes towards it. Most importantly, the study addresses the research questions:

1. What characterizes the formulaic language use of advanced EFL learners when writing short tasks?
2. What are these advanced learners' habits, beliefs and attitudes towards their formulaic language use?

For reference, the mixed methods study was conducted with four advanced English language learners at one of the best universities in Hungary, majoring in English and American studies BA, English in Media specialization. The study consisted of two parts: the text analysis of four short written tasks on the subject of FL elements and their frequency; the next step was to conduct interviews with the participants about the listed constituents and receive feedback regarding the analysis of their short essays.

4.1 The results of the text analysis

Research Question 1: What characterizes the formulaic language use of advanced EFL learners when writing short tasks?

This section provides a comparative analysis of FL elements in four texts. It aims to determine the total frequency of formulaic elements in the short essays, the frequency of collocations, phrasal verbs, technical lexical phrases, linking phrases, lexical bundles (if possible). The results are demonstrated in the tables below.

In the first text, the lowest rate of FL frequency is 31% by participant 4; the highest one equals 40% by participant 3, slightly below the statistics of the native speakers (see 2.2). From the first text, it is noticeable that the writing includes many formulaic elements. Regarding the categories, a significant difference is seen in collocations: the gap between the lowest and the highest figures is 10%.

Table 4.

The frequency of FL use in the text about stress

№ of participant	Collocations, %	Phrasal verbs, %	Technical lexical phrases, %	Lexical bundles, %	Linking phrases, %	Total, %
1	15	0,8	12	2	3	33
2	16	3	10	2	3	35
3	18	0,1	14	3	2	40
4	8	2	17	1	2	31

In the text about longevity, the highest rate is 30% by participant 1, and the lowest one is 15% by participant 4. Zero per cent of results appeared in two categories for the first time: phrasal verbs and lexical bundles.

Table 5.

The frequency of FL use in the text about longevity

№ of participant	Collocations, %	Phrasal verbs, %	Technical lexical phrases, %	Lexical bundles, %	Linking phrases, %	Total, %
1	16	3	8	1	2	29
2	18	1	8	2	1	30
3	11	0	10	3	1	25
4	10	1	3	0	1	15

In the third text, the highest figure of FL frequency is 35% by participant 2, while the lowest is 17% by participant 4. The impressive results concern collocation statistics: it fluctuates from 6% to 27%. A significant difference regards technical lexical phrases, which is 8%.

Table 6.

The frequency of FL use in the text about issues of nutrition

№ of participant	Collocations, %	Phrasal verbs, %	Technical lexical phrases, %	Lexical bundles, %	Linking phrases, %	Total, %
1	8	1	15	0	4	29
2	27	0	4	2	3	35
3	12	0,5	13	0	3	28
4	6	0,6	7	0	4	17

Finally, the highest rate of FL in the text about the brain is 34% of participant 4, and the lowest is 15% of participant 4. The significant difference in percentage concerns collocations and technical lexical phrases, where it is 11% for both cases.

Table 7.

The frequency of FL use in the text about the brain

№ of participant	Collocations, %	Phrasal verbs, %	Technical lexical phrases, %	Lexical bundles, %	Linking phrases, %	Total, %
1	13	2	3	2	6	22
2	17	1	11	2	3	34
3	4	0	12	0	12	28
4	8	2	1	0	4	15

To summarize the statistics, the average percentage of FL use in the short written tasks is approximately 30%, slightly lower than the native speakers' results. The most fluctuating formulaic categories are collocations and technical lexical phrases, with an average frequency of 15% and 10%, respectively. Linking phrases follow the statistics linking phrases with 7%. On the other hand, the most stable sections are phrasal verbs and lexical bundles resulting in 2-3% of frequency.

The text with the lowest FL frequency is the one about longevity. It is suggested that this topic is not as widely covered in social media as others or is the least engaging for the participants to write about. In contrast, in the text about stress, the students showed extensive use of multiword sequences. The reason for that might be that stress surrounds them in university life, and this is something they can discuss for a long time.

In order to reveal some of the reasons for the findings, following the analysis of the results, the participants were asked about the uneven results of the total percentage, collocations and technical phrases. The participants said the topic and efforts made a difference when writing the short tasks. For instance, participant 1 stated that many technical phrases about nutrition were used in the text. However, she did not like the topic and tried to compensate for her lack of interest with advanced constrictions. Regarding efforts, participant 3 said that by the time for writing the text about the brain, she was already tired and could not produce an excellent academic essay, as she had done in the first text. Besides, participant 4 operated with her words rather than phrases from the sources, which is why her percentage of FL use is relatively low. Overall, FL use is not a stable rate, as it depends on various factors.

4.2 The interview results

Research Question 2: What are these advanced learners' habits, beliefs and attitudes towards their formulaic language use?

A separate interview was done with all the participants to investigate the listed constituents. The results for each constituent are divided into separate sections. Also, the tables below the findings show the most demonstrative answers.

4.2.1 Results of students' habits

This section included questions about the process of writing the short tasks, choosing vocabulary to use and its kind – single words or phrases. It aimed to determine if the participants were conscious of FL expressions from the class or their sources at the first stage of creating their texts. The additional questions concerned the amount of digital help when writing short essays (see Table 8 for the main points).

The results show that only half of the participants pay attention to vocabulary when doing the homework. Firstly, two out of four people concentrated on the content more. They read several articles, tried to understand their general meaning and compressed them to 200-250 words, having what they wanted to say in mind. However, after creating an idea, another two participants made notes from their sources, which might be beneficial in terms of FL use: *"I like to have at least 3-4-5 sources. Then I like to put it away for a few days and revisit it"*. By retrieving new expressions from the native or at least native-like context (see 3.4), the students memorize them and, as a result, expand their vocabulary with advanced elements.

The same tendency applied in the case of vocabulary: half of the people were conscious of the language during the classes or reading. Two participants mentioned that they focused on the

style of their texts rather than the phrases themselves, and they aimed to sound more formal or not too sophisticated to make a text more manageable for reading. The others stated that they pay attention to useful phrases, including FL ones, from the class materials and try to proceed with them, or they want a similar tone to the one from the source and, therefore, utilize the same expressions. This example shows that the students had different focuses when writing the short task: for some, the impression on a reader matters more, while others pay attention to FL likewise. It is also possible that the people who emphasize the style more have a more comprehensive range of advanced vocabulary and have developed unconscious habits. As a result, their focus is shifted to other criteria.

Exciting comments were made regarding the use of phrases instead of single words. Firstly, everyone said they tried to use phrases as much as possible primarily because of the word count and being clear about their ideas. Nevertheless, most importantly, one of the participants mentioned the complexity of mirror translation FL sequences from Hungarian to English and the necessity of simply memorizing them. This comment answers why it is hard to acquire FL, as one cannot come to any of them on his own. Another remarkable observation refers to the advantages of using multiword phrases. The student stated that they helped her to be precise about what she was going to write, as English enables her to explain a concept with one word or phrase instead of long sentences; likewise, in Hungarian, where it is possible to express yourself by adding prefixed to the root. This statement highlights the importance of multiword phrases because, as said before, they accelerate writing (see 2.2). Therefore, every participant affirmed the aspiration to utilize phrases rather than single words for the word count and to be straightforward.

According to the interview, the students utilize many digital platforms in their writing. The most frequent ones are translation apps, such as *Google Translate*, helping to find synonyms or translate words from their native language to English. Two people named auto correctors in *Microsoft Word* and *Google Documents* as excellent tools for checking grammatical mistakes. Other helpful applications for writing are *Grammarly* and *Quillbot*, which correct different types of errors, including more clarity and engagement. In the case of dictionaries, only a few were listed: *Merriam-Webster*, *Thesaurus* and *Sztaki* (for translating from Hungarian to English); some use them somewhat. Returning to the availability issues two decades ago and the new opportunities (see 2.3), these observations prove that FL is fully accessible now in the form of online dictionaries, correction and translation apps. The participants take advantage of it.

In conclusion, half of the participants paid attention to vocabulary in classes or while reading: used to make a plan for their essays and included multiword elements there. However, the others did not focus on it, which may indicate having unconscious habits when increasing

vocabulary. Everyone endorses the importance of phrases rather than single words and resorts to the help of digital tools.

Table 8

Findings about students' habits in writing

Category	Code	Definition	Examples
Habits	HAB	Habits in writing the short tasks	P4: <i>"This is, actually, pretty easy [writing]. I just read lots of texts about the topic, and then I explain the idea in my own words."</i> P1: <i>"We are given the class the relevant vocab, so I try to put these phrases and idioms into the text."</i> P3: <i>"I always write these texts in Microsoft Word or Google Documents, so it checks grammar. ... I always never use dictionaries."</i>

4.2.2 Results of students' awareness

This set of questions aimed to estimate to what extent the participants were aware of FL in general, about their conscious attitude to multiword sequences in Popular Science in Social Media class and in writing the short task. The concept of awareness was exciting to research, as, according to the new pedagogical recommendations (see 2.3), more emphasis should be made on FL in language education (see Table 9 for the most demonstrative answers).

In terms of associations with multiword phrases, half of the students mentioned the authentic aspect of it. For instance, it was stated that proverbs and idioms help a person better understand a native speaker's culture of a target language overall. Another participant compared learning English and Italian: in both cases, FL phrases helped her to talk about things only L1 speakers of that language can understand. One more important comment refers to a personal issue: FL can be a weakness of English learners due to the inability to use mirror translation. It refers to its concept of nativity, as FL, which its learners have to memorize, is initially produced by L1 language speakers. Surprisingly, syntax was listed as one of the associations with multiword sequences, as the person did not know much about multiword constructions. In summary, most participants have at least an idea about FL and its significance in language learning.

Let us now move on to the intended FL use in writing. Three-quarters of the participants highlighted that they consciously utilized formulaic elements instead of single words. Although the answers slightly overlap with the ones in the habits section, two of the four participants said FL helped them to prove their high level of English in writing, sound more academic and make the text more readable and concise: *“I always feel like that person who said them or wrote them has very developed language skills. And I'd like to be like that one day”*. However, according to the interview, one student was unaware of using FL elements in the short tasks because of a lack of knowledge in this field. Finally, the observations show that most participants comprehended the utility of multiword terms when writing the short assignments for the Media class. The phrases helped the students to make a better impression on a reader, which is one of its functions (see 2.2).

Unexpected findings concern the students' attention to lexical phrases when reading an article or preparing for other classes. Three out of four participants mentioned that they do not collect useful FL expressions from the sources because they learn these phrases unconsciously and, therefore, do not need extra effort towards vocabulary. On the other hand, one person said that she found the university classes, including ones in Media specialization, eye-opening regarding new expressions, as professors provide students with high-level content. Lastly, the results prove that some advanced ESL learners unknowingly expand their vocabulary by picking words from the context.

In conclusion, most participants are familiar with FL and its benefits in language learning and precisely writing; however, it cannot be said that the participants are aware of them. Also, there is a proven connection between the unconscious habits of advanced ESL learners and their FL acquisition.

Table 9

Findings about students' awareness of FL

Category	Code	Definition	Examples
Awareness	AWA	Awareness about FL in general	P3: <i>“I never sit down with the intent of paying attention to the language in English. It [FL] just reaches me unconsciously because I speak the language well.”</i> P2: <i>“I think if it's done [a phrase] properly and it's done in a very nice manner, it just sticks out, and you</i>

		<i>don't think about it consciously."</i>
		P4: <i>"I'm recalling syntax, and syntax is something that I hate. So I have mixed feelings when I read this sentence [about FL]"</i> .

4.2.3 Results of students' attitudes towards FL

The last question section logically follows the previous ones. After learning about FL, the participants were asked about their plans and strategies. Therefore, the set includes questions about their desire to proceed with the FL acquisition, the means of its achievement and its benefits (see Table 10 for critical comments).

To start with, each participant wanted to study formulaic phrases more. As they said, English and other languages constantly change, and there is always room for improvement, especially in terms of FL, which is hard to acquire. A significant impact on total language skills with its help was also mentioned because using native constructions presents an advanced level of English: *"It also changes your thinking if you know how to think in English"*. After the study, the participants considered multiword sequences an essential part of the learning process.

There are several ways to develop the FL use suggested by the participants. The most efficient one, as they said, is attending classes in English. Two people stated that teachers are partly responsible for students' improvement, motivating them and distributing relevant handouts. Moreover, special language classes (Language practice, for example) might also be beneficial in learning FL phrases. An interesting point was made regarding studying methods: if a person is visual, for example, watching films or videos is helpful when learning new information. As the participant is auditory, reading out loud is the most impactful way to memorize FL. Lastly, often consuming content produced by native speakers is another way to acquire multiword expressions efficiently. To sum up, attending university, taking language development subjects, and consuming authentic materials are ways to develop FL knowledge.

Regarding the benefits of formulaic sequences in general and writing, the most demonstrative point is that they give a sense of belonging to a person, which refers to its nativity concept. Besides, they boost language skills, change the way of thinking in English, and, as a result, a learner becomes a better speaker of the language. Finally, academic style and more words in writing are an excellent addition to FL's listed benefits. As is seen, lexical phrases have many significant features in English; therefore, it is vital to be conscious of them.

Overall, the participants desired to develop their FL knowledge, and they considered institutions, English language classes, applying effective studying methods and authentic content consumption as helpful tools. In terms of FL advantage, it helped the students to improve their language proficiency and academic writing skills.

Table 10

Findings about students' attitudes towards FL

Category	Code	Definition	Examples
Attitude	ATT	Attitude towards FL development	P1: <i>"In writing, word count, sounds more academic. In spoken language, you sound more native."</i> P2: <i>"I think it makes the language a bit more real. It also gives a sense of belonging."</i> P4: <i>"Using lexical phrases will help to boost our language level."</i>

The most important findings regard the ubiquity of FL in writing: it accounts for a large percentage of advanced ESL learners in writing. However, it is seen that they are not conscious of it, which is inappropriate considering the importance of formulaic expressions and their functions. Nevertheless, the provided suggestions about education and their effort may help them to become aware of FL in the future.

5 Conclusion

This chapter summarizes the findings and answers the research questions about the characteristics of FL use of four advanced ESL learners and shows their habits, attitudes and awareness. Based on the findings, limitations, pedagogical implications and suggestions for the future are also listed.

5.1 Finding and discussions

Referring to the statistics of 52.3% of using FL in writing by English native speakers, the advanced ESL learners showed the results of approximately 30% in their short tasks. It is also evident that the most frequent formulaic elements in writing the short essays are collocations and technical lexical phrases. As a result, the findings highlight the importance of multiword elements, as they constitute a large part of the texts, and learning the proper constructions. Their frequency may fluctuate from essay to essay for numerous reasons, and it still makes up a large share of vocabulary.

One of the research tasks was to investigate the current situation regarding FL use and estimate the progress in this field. Based on the interview, dramatic changes occurred. Firstly, the participants utilize many digital tools in writing: dictionaries, translation apps and websites for correcting errors. In addition, it provides them access to authentic materials and native phrases, which is crucial in FL acquisition. For example, if there was an issue with source distribution two decades ago, now authentic materials are entirely available. Secondly, the participants have at least an idea about FL but are not aware of them still. It can be concluded from this that emphasis on multiword expressions is done, but not enough, resulting in unconscious habits in writing. Considering how useful the phrases can be in writing, more effort should be made on this in the subjects.

With regard to education, the subjects of the study mentioned its significance when learning FL. Ideally, professors should provide them with handouts and sources, hold language practice classes, where they would improve their performance, and motivate them. Nevertheless, other techniques were listed: immersion into the native environment by consuming authentic content and finding the most efficient way of studying. After the research, the participants were inspired to improve their FL knowledge through these methods; it might change their perception.

5.2 Limitations of the study

The limitations are the number of people and the vague concept of FL. Firstly, the research might be considered exploratory, as it was carried out with only four participants. As a result, it is

impossible to characterize the FL use of advanced ESL learners , monitor the changes on a general level, and draw any generalizations. Only advanced ESL learners were considered, and there is no data about other language levels at this point. The second limitation is the obscurity of multiword phrases. It is still hard to categorize them and provide clear frequency statistics. The results are preliminary, based on one institution.

5.3 Pedagogical implications

The results of the research may be applied to the learning process. Firstly, it is evident that the participants are primarily unaware of formulaic phrases; therefore, it should be emphasized more during the classes, told about its functions and ubiquity. Another implication regards the sources: the participants highlighted the role of teachers in FL acquisition. It follows that they should pay attention to the quality of handouts, use authentic sources when teaching and motivate the students to learn multiword elements.

5.4 Suggestion for further research

Based on the limitations, a more detailed study should be conducted, including more participants and level groups. Since the concept of FL is vague, it is worth conducting new research after clarifying its categories and providing a more explicit definition. As the research results are exciting, it is worth exploring FL in writing more.

6 References

- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 101-122). *Oxford: Oxford University Press*.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10(3), pp. 245–26.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72-89.
- Creswell, J.W. (1998). *Qualitative inquiry and research design*. Thousand Oaks: Sage Publications.
- Egamova, J. & Sharofova, N. (2022). The importance of vocabulary in learning English language. *Central Asian Research Journal For Interdisciplinary Studies (CARJIS)*, 2(11).
- Ellis, N. C. (2008). Phraseology: The periphery and the heart of language. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 1–13). *Amsterdam, the Netherlands: John Benjamins*.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29-62.
- Meunier, F. (2012). Formulaic Language and Language Teaching. *Annual Review of Applied Linguistics*, 32.
- Johnson, R.B., & Onwuegbuzie, A.J. (2004). Mixed method research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Ohlrogge, A. (2009). Formulaic expressions in intermediate EFL writing assessment. In R. Corrigan (Ed.), *Formulaic language* (Vol. 2, pp. 375–386).
- Kilich, S. (2015). The use of formulaic language by English as a foreign language (EFL) learners in writing proficiency exams. *The Graduate School of Education of İhsan Doğramacı Bilkent University*.
- Lundell, F. (2021). Formulaicity and corpora. *Routledge, London*, pp. 370-381.
- Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics*, 32(2), 129-148.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-225). London: Longman.

Schmitt, N. (2010). Formulaic language is as important as individual words. *Researching vocabulary: A vocabulary research manual*, 8-12.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (Vol. 9, pp.1-23). *Amsterdam: John Benjamins*.

Schmitt, N., & Martinez, R. (2004). A Phrasal Expressions List. *Oxford University Press* 33(3), pp. 299-320.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. *Oxford: Oxford University Press*.

Wood, J. (2020). Classifying and Identifying Formulaic Language. *The Routledge Handbook of Vocabulary Studies*, 30-46.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, 20, 1–28.

7 Appendices

Appendix A: The interview questions

Habits:

- How do you write a short piece?
- How do you choose which vocabulary to include?
- How do you edit your text to check and correct vocabulary?
- Do you use any (digital) help to include vocabulary? If so, what?
- To what extent do you use translator apps or dictionaries?
- To what extent/When do you use phrases rather than single words?

Awareness:

- What comes to your mind when you hear the phrase: formulaic language/sequences, lexical phrases/chunks, lexical bundles, idioms, collocations and proverbs?
- How much do you know about formulaic language use in English?
- Do you consciously use lexical phrases instead of single words?
- Do you ... (can you give an example)
- Do you sometimes pay attention to collocations or other lexical phrases when you read an article for our media class or other classes?

Attitude:

- Would you like to develop your formulaic language use?
- How do you think it can be achieved?
- What do you think are the benefits of using lexical phrases when developing your English and writing skills?

Appendix B: The text analysis sample

Collocations

Phrasal verbs

Technical lexical phrases

Lexical bundles

Linking phrases

Dan Buettner in his presentation titled "How to Live to be 100+" talks not only about the common myths people have **in connection** with longevity, **but also** offers an insight into 3 special places where people **get to live** to be more than 100 years old. What these societies have **in common** are **the following**: **First of all**, they exercise regularly, but not in the **original way**. Moving their bodies happens when they **take the stairs**, during cooking or baking, or while gardening. These are all activities they enjoy and engage in on **a daily basis**. Members of these communities are also **taking their time to slow down**, connect to **a higher power**. They have their own spiritual motivation for **waking up** in the morning and **finding inner meaning**. **Of course**, **healthy eating** is an aspect too, but it is not really a diet. They consume **plant-based food** while still **keeping meat on their tables**. **And last but not least**, they have many **social connections** and **they are surrounded by** others **who live a healthy lifestyle**, both physically and spiritually, **like they do**. The research of Buettner is groundbreaking. **It turns longevity into** an accessible goal for everyone, **as opposed to** radical lifestyle and diet trends. All of these observations reflect a simple yet active **way of life**, which can be easily achieved by **dedicating your time to it**. **After all**, we **put our energies into** such irrelevant and toxic segments—what would happen if we focused on the good?