

**Ministry of Education and Science of Ukraine
Taras Shevchenko National University of Kyiv
Scientific and Educational Institute of Philology
Department of English Philology and Intercultural Communication**

Master's thesis

**COMPARATIVE STUDY OF LEXICAL DENSITY AND LEXICAL
DIVERSITY OF CREATIVE AND ACADEMIC WRITING OF UKRAINIAN
EFL STUDENTS**

Samusenko Maryna

2nd year student of the Education Program
'English Communication Studies and Translation
and Two Western European Languages'
Field of science: 03 "Humanities"
Specialty: 035 "Philology"

Supervised by:

Karaban Anna, PhD

«Допущено до захисту»
Протокол засідання кафедри англійської філології
та міжкультурної комунікації
Протокол № 8 від 18.04.2023
Зав. кафедри _____ д. філол. н., проф. Алла БЄЛОВА

KYIV – 2023

CONTENTS

INTRODUCTION.....	4
CHAPTER I. LEXICAL DENSITY AND LEXICAL DIVERSITY IN WRITTEN ENGLISH	7
1. Lexical density of written English.....	7
1.1. Lexical density of EFL creative writing.....	10
1.2 Lexical density of EFL academic writing	13
2. Lexical diversity of written English	17
2.1 Lexical diversity of EFL creative writing.....	21
2.2. Lexical diversity of EFL academic writing	25
Conclusion.....	28
CHAPTER II. DATA SELECTION AND PROCESSING FOR RESEARCH PURPOSES.....	31
2.1. Data for lexical density and lexical diversity study and selection criteria for creative and academic writing produced by Ukrainian EFL students.....	31
2.1.1. Participants and research setting.....	31
2.1.2. Data selection process.....	32
2.2. Quantitative analysis instruments and procedures	32
2.2.1. Lexical density and lexical diversity calculation.....	33
2.2.2. Frequency analysis.....	33
Conclusion.....	34
CHAPTER III. COMPARATIVE ANALYSIS OF UKRAINIAN EFL WRITING	36
3.1 Lexical density Ukrainian students' English writing	36
3.1.1. Lexical density Ukrainian EFL students' academic writing.....	36
3.1.1.1. Lexical density of Ukrainian students' course papers.....	36
3.1.1.2. Lexical density of Ukrainian students' Bachelor's papers	40
3.1.2. Lexical density Ukrainian EFL students' creative writing.....	48
3.1.2.1. Lexical density of second year undergraduate Ukrainian students' creative writing.....	48
3.1.2.2. Lexical density of fourth year undergraduate Ukrainian students' creative writing.....	53

3.2. Lexical diversity of Ukrainian students' English writing	60
3.2.1. Lexical diversity Ukrainian EFL students' academic writing	60
3.2.1.1. Lexical diversity of Ukrainian students' course papers.....	60
3.2.1.2. Lexical diversity of Ukrainian students' Bachelor's papers.....	67
3.2.2. Lexical diversity Ukrainian EFL students' creative writing.....	76
3.2.2.1. Lexical diversity of second year undergraduate Ukrainian students' creative writing	77
3.2.2.2. Lexical diversity of fourth year undergraduate Ukrainian students' creative writing	83
Conclusion.....	92
CONCLUSION.....	97
SUMMARY.....	100
REFERENCES	102
APPENDICES	109
Appendix 1	109
Appendix 2.....	110
Appendix 3	111
Appendix 4	112
Appendix 5	113
Appendix 6	115
Appendix 7	116
Appendix 8.....	117
Appendix 9.....	119
Appendix 10	120

INTRODUCTION

For many years, the study of second language learners' writing and its features have been a subject of particular interest for scholars and educators. Therefore, various studies have been conducted in order to examine the influence of different aspects, such as syntax, grammar, and discourse organization on the quality of writing. However, there has been a paucity of attention regarding such crucial concepts as lexical density and lexical diversity in EFL writing. The pivotal role of these factors in determining the effectiveness of writing and proficiency of its author cannot be overlooked, since based on them crucial insights concerning the learners' ability to convey their ideas coherently and clearly while utilizing a wide vocabulary range may be obtained. Additionally, lexical density and lexical diversity measurement may provide researchers and language instructors with valuable insights concerning the areas where students encounter difficulties, thereby allowing them to effectuate pedagogical interventions targeted at promoting academic growth and development.

The topicality of the research is predicated by the fact, that with the increasing impact of globalization on both public and private life in the world in general and in Ukraine in particular, an increasing number of Ukrainians become EFL learners, which makes it crucial to identify their unique language learning challenges as well as to determine their relative position in the global EFL community.

The purpose of our paper is to examine the relationship between lexical density, lexical diversity, and language proficiency.

The purpose of the paper entails solving the following **tasks**:

- To explore the concepts of lexical density and lexical diversity in academic and creative writing;
- To investigate lexical density and lexical diversity values of the writing produced by international EFL students;
- To determine the existing patterns regarding lexical density and lexical diversity of academic and creative writing of Ukrainian EFL students;

- To create Lexical Frequency Profiles of Ukrainian EFL students;
- To draw connections between the language proficiency of the subjects and the lexical density and lexical diversity of their writing;
- To explore the relationship between lexical density and other linguistic features such as sentence length and word length in creative and academic writing of Ukrainian EFL students;
- To examine the potential impact of such factors as text genre and language proficiency on lexical density and lexical diversity;
- To position Ukrainian EFL students in the community of international EFL learners in terms of their lexical richness and active vocabulary.

The object of our research is the language of creative and academic writing produced by Ukrainian EFL students.

The subject of this research is lexical density, lexical diversity, and language command of Ukrainian English learners and their correlation.

The material of the research is academic and creative writing produced by the English philology students at Taras Shevchenko University of Kyiv, Ukraine, during the second and fourth years of study:

- Academic writing:
 - course paper (second year of study) - 11 course papers, on average containing 28 pages, 7680 words and 40,444 symbols without spaces. In total 294 pages, 82,456 words and 450,164 symbols without spaces;
 - Bachelor's paper (fourth year of study) - 11 Bachelor's papers, each containing on average 57 pages, 16,716 words, and 86,545 symbols without spaces. In total 629 pages, 183,874 words, and 951,992 symbols without spaces.

- Creative writing:
 - a review on a service or an essay on the topic “What freedom means to me” (second year of study) - 11 one-page long texts consisting of on average 275 words and 1,320 symbols without spaces. In total 11 pages, 3,025 words, and 14,522 symbols without spaces;
 - an essay titled “Ukrainian Childhood” (fourth year of study) - 11 two-page essays, each comprised of 558 words and 4073 symbols without spaces, which in total amounted to 23 pages, 6142 words and 44, 798 symbols without spaces.

The methodology of the research is a combination of frequency analysis, lexical density, and lexical diversity calculation, as well as other elements of quantitative analysis, which allows to determine lexical density and lexical diversity values in Ukrainian students’ writing and convey their analysis.

The scientific novelty of the paper lies in the exhaustive and systematic study and analysis of the lexical peculiarities of Ukrainian EFL students, which allows to assess their linguistic competence.

The practical value of the research lies in the fact that its results may be used in creative and academic writing seminars, as well as by EFL educators and instructors.

The Master’s paper is divided into three parts. The first part deals with the theoretical foundations of the study. The second part provides a detailed overview of the research methodology. The third part provides systematized and analyzed findings of the study. The Master’s paper also includes an introduction, conclusions, appendices, and a list of references.

CHAPTER I. LEXICAL DENSITY AND LEXICAL DIVERSITY IN WRITTEN ENGLISH

The assessment of students' writing skills has been a pivotal subject of research in the field of linguistics. Various studies investigated the means of evaluating students' linguistic proficiency as well as their progress in language acquisition. It has been argued that such measures as lexical density and lexical diversity are credible indicators to be relied upon in the process of lexical development exploration [83, p. 143]. It is these measures and their significance in the written mode of linguistic operation that will be explored in depth in the following sections of this chapter.

1. Lexical density of written English

The concept 'lexical density' was first used by Ure in 1971. The scholar used the term to denote the ratio of lexical or content words to the total number of words in a text. Ure stated that a word is only orthographic, therefore phrasal verbs such as 'speak up' are treated as two separate items: speak as a lexical word and 'up' as a non-lexical word. Additionally, when calculating lexical density using Ure's method, content words are counted in spite of their recurrence throughout the text. As for the words endowed with lexical meaning, according to the researcher these are verbs, nouns, adjectives, and some adverbs, whereas other parts of speech are viewed as carriers of only grammatical values. Ure also made an observation according to which the typical lexical density for a spoken text is under 40 percent, while a lexical density of 40 percent and more is characteristic of a written language [83, p. 96-104].

Thus, a graphic representation of Ure's formula for lexical density measurement is provided below:

$$Ld = \frac{\textit{The number of lexical items}}{\textit{The total number of words}} \times 100$$

Ure's idea was further developed by Halliday, who viewed lexical density as an indicator of proximity with which lexical items are located in a grammatical structure. The scholar stated that one lexical item could be represented by more than one word. Thus, unlike Ure who would state that word combination 'speak up' consists

of two items: lexical (speak) and grammatical (up), the scholar would consider the aforementioned phrasal verb to be one lexical item. Hence lexical items are viewed by Halliday as items carrying one lexical meaning, lexical density being calculated as the ratio of lexical items to the number of clauses [25; p. 655]. Therefore, Halliday's formula of lexical density is as follows:

$$Ld = \frac{\textit{The number of lexical items}}{\textit{The total number of clauses}} \times 100$$

This point of view is supported by Saragih and Paltridge, who stated that lexical density represents a number of content words, which are nouns, verbs, adjectives, and adverbs, encountered in a clause. Therefore, lexical density may be estimated as the number of content words in a text or a sentence in proportion to all the words in a text or a sentence [69, p. 7; 62, p. 136].

Additionally, Nunan, Johansson, Stubbs as well as Richard, and Schmidt share an opinion that lexical density signifies the ratio of content words to the total number of words, and is calculated as a percentage. The researchers agree that content words are characterized by a high information load while function words serve to create relationships and dependencies within a sentence, not being carriers of information [60, p. 11; 34, p. 65-67; 73, p. 117; 67, p. 336]. Therefore, text accessibility is inherently dependent on its lexical density, i.e. more lexically dense texts are more difficult to perceive, while texts with lower lexical density are easier to understand. This opinion is supported by Rahmansyah who agrees that higher lexical density directly corresponds to the amount of information in a text, thus determining the degree of difficulty it has [63, p. 8].

In line with the viewpoints above, Nunan and Kondal postulate that lexical density measures informational load in a sentence or text and is calculated as a proportion of content words in textual material. Nunan also suggests that lexical density is calculated as a ratio of lexical words to grammatical words [60; 37].

Moreover, Malvern, Richards, Chipere, and Duran adhere to an opinion that taken into consideration lexical density being token-token ratio (TTR), its index is not

dependent on the sample size [42, p. 192]. Therefore, lexical density measurement could be conducted in textual material of any size. Thus, if there are many lexical words in relation to the number of all the words in a text, it is considered to be 'dense'.

Additionally, there are various software programs that are able to assist the researchers and compute lexical density based on the Token-Token ratio [51]. For example, the Systematic Analysis of Language Transcripts (SALT) [14], the Oxford Concordance Programme [27], and the Computerized Language Analysis (CLAN) programs [46; 47].

In addition, Laufer and Nation state that lexical density is not limited to the measurement of lexis since it also reflects cohesive and syntactic elements of text composition. Thus, the scarcity of function words may correspond to the abundance of ellipsis or participle clauses, which belong to the structural rather than lexical characteristics of a text [39; p. 309]. Hence it may be argued that lexical density is not a one-dimensional phenomenon but a complex metric for text complexity evaluation, from which both lexical and structural characteristics of textual material may be inferred.

It should be noted that various researchers pointed out the crucial role of lexical density in terms of the differentiation between spoken and written guises of language. To be more precise, higher lexical density is an inherent feature of a written text [73; 34; 79; 82; 15; 1]. This is due to the fact that the concentration of ideas and information as rendered through lexical items is higher in written mode [65, p. 87]. It is also connected with the nature of written and spoken languages – while written language is more reflective and informative, spoken is more active and conversational, therefore information is packed more densely in written language [44, p. 4; 23, p. 63; 43, p. 158].

Taking into consideration the information discussed above, it may be noted that the critical role of lexical density in terms of text quality determination cannot be underestimated. There is uniformity within the scientific community in relation to the variables used in the process of lexical density measurement, i.e. the distinction between lexical and functional words.

It is of paramount importance to be mentioned that the subsequent sections of this chapter provide an in-depth overview of the previous research findings concerning lexical density in both creative and academic ESL/EFL writing.

1.1. Lexical density of EFL creative writing

Creative writing is a type of writing activity that focuses on expressiveness, imagination, and imagination. It encompasses a variety of genres, such as poetry, fiction, creative nonfiction, and drama, being often used as a means of artistic and personal exploration.

When approached in the context of EFL/ESL instruction, creative writing can be used as a valuable tool for honing language skills and promoting intercultural competence [30; 48]. It can also help encourage students to express themselves in a non-threatening environment and explore their identities while activating their vocabulary knowledge in the process of creative activity [66].

As it was already mentioned, lexical density is a widely used criterion for writing assessment. Thus, various researchers applied this measure to analyze ESL/EFL students' creative writing. The findings provided by some of these studies are in the focus of this section.

For instance, the research conducted by Carmen Gregori-Signes and Begoña Clavel-Arroitia analyzed creative and academic writings of first-year undergraduate EFL students at the Universitat de València with B1 level of English, at the beginning and at the end of an academic year [21]. The creative writing assignment given to the students was to write a letter to a friend describing a visit to another country. Thus, in the first semester, the average lexical density of the students' creative writings was 36.5%, and in the second - 36.6%, demonstrating a consistent index. Thus the average lexical density of the creative writings of the Universitat de València first-year undergraduate students measured in the first and second semesters is 36.55%.

Another study on lexical density in creative writings of ESL students belongs to Abdalla Elkheir Elgobshawi and Mohammed Aldawsari. The scholars evaluated

lexical density in the creative writing samples produced by Saudi undergraduates whose major was English as a Foreign Language at Prince Sattam Bin Abdulaziz University in Saudi Arabia [17]. The participants of the study, i.e. 60 undergraduate college students, divided into two groups of thirty, were assessed on the basis of short essays produced by them. Thus, the average lexical density scores obtained in the process of the analysis are 49.82% and 53.56% for groups 1 and 2 respectively, which makes an average of 51.69% for the whole group of participants.

A gender-based approach to lexical density measurement in EFL creative writing was implemented by Siti Aisah Ginting [20]. The subject group in this study consisted of 20 male and 20 female EFL students of the English Department Universitas Negeri Medan, Indonesia, whose ages were from 18 to 20 years old. The assignment on which their lexical density was evaluated, was descriptive writing about the students' favorite places. The findings of the study suggest that female writing was more lexically dense than male. As a result, the lexical density of the texts written by females was 52.74%, while that of males amounted to 48.68%, making an average of 50.71% for the whole subject group.

In addition, Daw Thida studied the lexical density of writing produced by 19 postgraduate students from the Department of English, Yadanabon University, Myanmar [78]. The researcher discovered that the minimal lexical density of these writings was 36.52% and the maximal – 57.55%. Thus, the average lexical density of the subject group's creative writings was 45.19%. The scholar also concluded that the texts written with lexical density lower than 40 may be considered to be closer to spoken mode.

The researcher who also studied lexical density in ESL students' creative writings was Josephine Uma. The participants of the research were 15 undergraduate students from the Department of Foundation in Science of Windfield International College in Malaysia, for whom English was one of the core academic subjects [81]. The nationalities of the participants were Chinese, Malaysian, and Indian, English being their second language of communication. The written assignment, which was

taken for the analysis, was an essay on the topic “The person I most admire”. The minimal lexical density result for the students in the first semester was 39% and the maximal - 53%, the average being 43.9%. The lexical density of the written materials produced in the final semester ranged from 38% to 54%, making an average of 45.4%. Thus, the lexical density indices in this research were consistent, and it may be calculated that the average lexical density in the texts produced by the students in this research is 44.65%.

Therefore, it may be observed that in the studies described above, the average lexical density of creative writings produced by EFL students from different countries and universities ranged from 36.55% to 51.69%. These statistics may be observed in the chart presented below, where the studied are marked as follows:

- A. Carmen Gregori-Signes, Begoña Clavel-Arroitia, *Analysing lexical density and lexical diversity in university students' written discourse*, Spain;
- B. Abdalla Elkheir Elgobshawi, Mohammed Aldawsari, *Lexical density as improvement indicator in the written performance of EFL majors*, Saudi Arabia;
- C. Siti Aisah Ginting, *Lexical Complexity on Descriptive Writing of Indonesian Male and Female EFL Learners*, Indonesia
- D. Daw Thida, *Lexical Density and Readability of Students' Writing*, Myanmar;
- E. Josephine Uma, *A Study of Lexical Richness in ESL Students' Writing*, Malaysia.

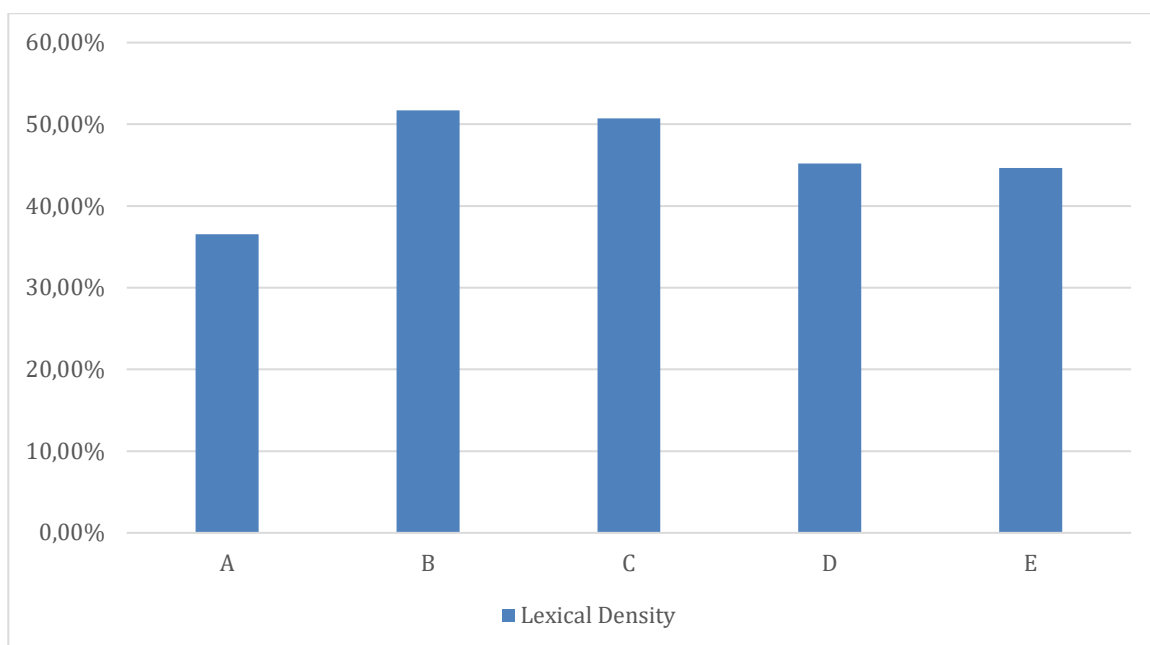


Chart 1. Lexical density of ESL students' creative writing

As it is visible on Chart 1, the results of the previous studies are fairly consistent, with the exception of the writings produced by the students of the Universidad de Valencia, whose lexical density was significantly lower than that estimated in the writings produced by the participants of other analyses. However, it may be attributed to the fact that their level of English was not as advanced as that of the other students participating in the studies discussed above. Overall, it may be calculated that the average lexical density in EFL students' creative writings considered in this section of the research is 45.76%.

It is important to mention that these indices will be further used in the course of this study to evaluate Ukrainian EFL students' writings as compared to the ones produced by their peers in the international discourse community.

1.2 Lexical density of EFL academic writing

Academic writing is a formal style of writing usually used in scholarly and professional settings. Its inherent features are formal tone, precise language, and logical structure [74, p. 8-30]. Academic writing is factual and coherent, ideas in it being conveyed clearly and accurately, which in turn entails a high level of lexical density.

It is desirable in academic writing that it has a high level of lexical density. This is due to the fact that the complex ideas in academic texts should be expressed precisely and accurately, which enhances their understanding by the audience [31, p. 167-170]. Furthermore, high lexical density may increase the credibility of the textual material, since in this case it is assumed that the writer is knowledgeable about the subject matter and is able to skillfully manipulate the language [3, p. 230-234].

Lexical diversity in academic writing has been carefully explored by various scholars. In this section of the thesis, several studies dedicated to the examination of lexical density in ESL/EFL students' academic writings and their findings will be discussed.

One such study was conducted by Carmen Gregori-Signesa and Begoña Clavel-Arroitia, whose work was also described in the previous section of the thesis since the researchers analyzed lexical density and diversity in both academic and creative writings of first-year EFL students at the Universitat de València [21]. The academic writing task assigned to the students involved producing a personal statement to apply for an Erasmus program. As a result, the average lexical density calculated in the texts written in the first semester amounted to 30.7%, while in the second semester, it amounted to 32.2%, indicating a slight increase. Therefore, the average lexical density of academic writings produced by the first-year EFL students of the Universitat de València is 31.45%.

Another research focused on lexical density in ESL academic writing was conducted by Rim Kouachi [38, p. 40-48]. The participants of the study were 22 first-year Master students of Larbi Ben M'Hidi University of Oum El Bouaghi, who learned English as a foreign language. The written assignments the subject group was asked to complete were an expository essay about a particular topic (Corpus A) and a comparison and contrast essay (Corpus B). As for the Corpus A, lexical density in it ranged from 51% to 59%, the mean being 54.76%, whereas in the Corpus B lexical density amounted to 46%-55%, with an average of 49.63%. Consequently, it can be

estimated that the average lexical density of all the texts produced for this research was 52.2%.

In addition, the scholars who investigated lexical density in EFL Indonesian students' academic writings were Hermawati Syarif and Rahmi Eka Putri [75, p. 89]. The subject group of this research consisted of English graduate students of UIN Walisongo, who produced 30 pieces of academic writing, them being their introduction proposals, to be further taken for analysis, The findings of the study demonstrate that the range of lexical density calculated in these texts was between 45.71% and 56.38%, an average being 51.19%.

The issue of lexical density in academic writing was also explored in the study conducted by Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong [59]. The researchers analyzed textual material produced by 64, of whom 55 were female, Vietnamese third-year English-major students aged 20 who attended 3 writing classes in central Vietnam. The two academic writing assignments which were taken for analysis were persuasive essays: the first one dealing with the impact of Facebook (Facebook task) and the second one being concerned with the role of school or family in personal growth (the school task). The results obtained in the course of the study demonstrate that the lexical density measured in the Facebook task ranged from 44% to 62%, while that in the school task amounted to a minimum of 47% and a maximum of 60%, the average of both being 52%.

Hence it may be estimated that the average lexical density in the academic writings discussed above ranged from 31.45% to 52.2%. The chart represented below demonstrates these statistics, where the studies described are marked in the following way:

- A. Carmen Gregori-Signes, Begoña Clavel-Arroitia, *Analysing lexical density and lexical diversity in university students' written discourse*, Spain;

- B. Rim Kouachi, *A Corpus-Based Study of the Lexical Density and Readability of Student's Academic Writings The Case of Master One EFL Students at Larbi Ben M'hidi University, Algeria*;
- C. Hermawati Syarif and Rahmi Eka Putri, *How lexical density reveals students' ability in writing academic text, Indonesia*;
- D. Bao Trang Thi Nguyen, Jonathan Newton and Tran Ngoc Quynh Phuong, *The Effect of Topic on EFL Writing by Vietnamese Tertiary Students: Insights from Combining a Lexical Richness Analysis with Student Self-Reports, Vietnam*.

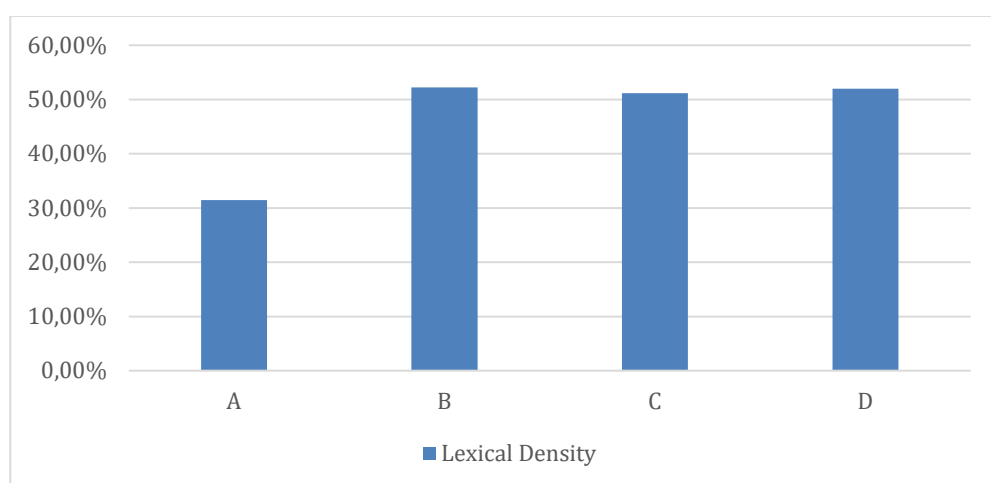


Chart 2. Lexical density of ESL students' academic writing

On Chart 2 it may be observed, lexical density indices obtained in the course of the studies discussed were almost equal, except for the ones attributed to the students of the Universitat de València. As in the case of lexical density in ESL students' creative writing, this discrepancy may be attributed to the lower education stage and consequently lower language proficiency level of these students. In general, the average lexical density of the textual material analyzed in the studies discussed in this section of the paper is 46.71. As was already mentioned in the previous section of the research, these data will be further used in order to place the writings of EFL Ukrainian students into the framework of the global EFL academic community.

2. Lexical diversity of written English

As numerous researchers have acknowledged throughout the past three decades, vocabulary plays a pivotal role in terms of gauging students' linguistic competences and second language learning progress [6; 39; 61]. One of the options for conducting such an assessment, widely applied in the research of ESL/EFL writing quality, is lexical diversity.

It is of paramount importance to be mentioned that previous research has demonstrated a direct correspondence between the students' vocabulary range and their lexical diversity, which has been assessed based on their text output in a situation where students were required to employ their lexical competence [54, p. 47; 39, p. 308-310]. What underscores the significance of vocabulary in language learning as such and the linguistic system as a whole is, according to Halliday and Hasan, its critical role in terms of meaning creation in the process of comprehending discourse [24, p. 5]. Similarly, the importance of vocabulary viewed from the linguistic perspective was highlighted by Celce and Olshtain [4].

As was noted by Nation, demonstrating a level of writing which is lexically diverse requires that students know a wide range of lexical units and can use them productively, which includes the knowledge of lexical form and lexical meaning [57, p. 262-276]. Developing a productive mastery of the lexicon is considered to be a more laborious task than the advancement of receptive vocabulary knowledge [70, p. 329–355]. This may be attributed to the fact that in the first case language learner takes an active position to retract words from memory and use them in an appropriate context.

Therefore, the implementation of lower-frequency words into writing, which contributes to the growth of a text's lexical richness, is indicative of students' abilities to communicate their ideas coherently and precisely [80, 36-62]. From this follows that the measurement of lexical richness may reveal the extent of a learner's skillful language production.

Taking into consideration the information discussed above, it is not unexpected that scholars have given the measurement and evaluation of lexical richness in written

and spoken language a considerable amount of attention. Thus, a closer look should be taken at the concept of lexical diversity per se, for it to be further practically applied in Chapter 3 of this research.

Lexical diversity or lexical variation refers to the statistical measures used to assess the quality and quantity of lexical units used in a text [10, p. 4]. It is a crucial component of language competency and is frequently used to gauge how sophisticated and complicated written or spoken linguistic material is. Researchers have extensively explored the idea of lexical diversity, examining a variety of influencing elements, including the writer's degree of education, cognitive capacity, and writing experience.

There is a general unanimity within the academic community regarding the definition of the term lexical variation, which is viewed as the amount of different words encountered in a text [34; 32, p. 279; 42, p. 3]. It is also worth mentioning that, according to Jarvis, 'different words' are words not repeated in a text sample, which may be viewed as the opposite of redundancy [33, p. 91]. However, this idea appears to be quite simplistic, overlooking a significant amount of lexically sophisticated texts, in which a low-frequency advanced vocabulary unit may be encountered more than once.

Therefore, the notion of lexical variation and its definition was further elaborated by Laufer and Nation, in accordance with whom lexical diversity accounts for the proportion of advanced words in a text [39, p. 309]. Furthermore, Housen et al. that it is phonologically and orthographically different words that represent one's vocabulary range [32, p. 3]. Thus, a text may be considered lexically diverse if it has a high percentage of advanced vocabulary units, which are not repeated frequently throughout the textual material.

One of the first studies focused on lexical diversity was conducted by Francis and Kucera in 1982. What the researchers discovered after studying a sizable written and spoken corpus of English was that an average individual uses between 2,000 and 3,000 distinct terms in daily conversation. Additionally, Francis and Kucera determined that people with higher levels of education typically use a wider variety of

words and have a bigger vocabulary size. Furthermore, the scholars discovered that lexical variation is influenced by a person's cognitive capacity and exposure to different types of language in addition to education level [19].

This idea was supported by Coxhead's research, who, based on the analysis of a corpus of academic texts discovered that students with higher levels of English proficiency tended to utilize more sophisticated and academic terminology [8]. Similar findings were made by Durrant and Doherty, who demonstrated that students who had studied English for longer periods of time had a wider lexical range and were more prone to enrich their writings with more complex and diverse syntactic structures [13, p. 87-98].

Furthermore, the existing studies demonstrate a strong correlation between academic writing of high quality, lexical diversity, and utilization of low-frequency vocabulary. This observation was valid for the writings of both native speakers of English and ESL students of different educational attainments [7; 77; 12; 52; 53, 61].

Taking into consideration the viewpoints and findings examined above, lexical diversity may be considered a reliable indicator of language proficiency. Furthermore, it can be used for the assessment of general linguistic competence, due to the fact that it is able to produce credible and consistent results.

As for the aim of the lexical variation measurement, it lies in discovering the richness of the lexicon in a given text [28, p. 121-122]. The most frequently used lexical diversity measurement method is Type-Token Ratio (TTR) [11, p. 30]. TTR corresponds to the ratio of unique words (types) to total words (tokens) in textual material [42, p. 19; 35, p. 273; 36, p. 24; 71, p. 291]. Biber et al., in turn, suggest that Type-Token Ratio demonstrates the vocabulary development throughout the text, thus indicating whether the writer frequently reiterated the same words or used non-repeated unique vocabulary units in their text [45, p. 53].

However, lexical diversity measurement based on Type-Token Ratio has been criticized by some scholars due to its dependence on the length of the discourse. This means that as the text length increases, the probability of encountering new words in it

decreases, which can be attributed to the necessity of frequently using the same function words, whose ratio will inevitably be higher than that of content words, which have a considerably higher probability of not being repeated throughout the text [55, p. 3; 50, p. 382]. This idea may be corroborated by the fact that the overwhelming majority of written texts contain 95% of high-frequency vocabulary [56; 58, p. 7]. Therefore, it is extremely improbable for a text to display an abundance in the low-frequency lexicon, which would be reflected in its lexical richness index.

As for the Type-Token lexical diversity measurement method, it is possible for a text to have a high TTR despite using mostly high-frequency words, as long as they are not constantly repeated throughout the text. Contrarily, a TTR of a text made up of many rare and specialized words, which are, however, frequently repeated in it, will have a low TTR.

Moreover, using the type-token ratio in the process of lexical diversity measurement was criticized by Johansson. The researcher stated that since one lexical word can consist of several function words, a longer text would demonstrate a lower TTR index as compared to a shorter one [34, p. 62-63]. Nevertheless, this method remains the most widely used in terms of lexical diversity measurement, not yet having a more accurate alternative.

Additionally, lexical variation tightly correlates with students' writing experience. This idea is empirically demonstrated in the research conducted by Hu and Nation who investigated the lexical diversity in texts produced by English language learners from China, taking into consideration their writing experiences. The significant finding of the study was that students with more writing experience display a wider vocabulary range and the ability to build more complicated sentences [49, p. 422-423]. This discovery highlights the importance of taking into account the writing experience of a subject group chosen for lexical variation assessment, which will enable researchers to obtain a clearer picture of the group's linguistic proficiency.

Another factor whose influence on lexical diversity cannot be underestimated is text genre. This is due to the fact that lexical richness requirements vary according to

the type of textual material. Hence studies determined that news articles tend to have a higher TTR than personal narratives or conversational speech [68, p. 333]. For this reason, in this research lexical diversity, as well as lexical density analyses will be conducted separately for academic and creative writings.

2.1 Lexical diversity of EFL creative writing

Lexical variation plays a pivotal role in enhancing the quality and impact of English as a foreign or second language writing. In addition to demonstrating the writer's command of the language, vast and varied vocabulary aid in the nuanced expression of complicated concepts and feelings.

Lexical variety helps to maintain the reader's attention in the text in addition to improving literary quality since the material is more engaging and entertaining to read when a writer utilizes a diverse vocabulary. This can be especially crucial in creative writing, where the author's capacity to hold the reader's interest is essential to the success of the work.

Nonetheless, many students may find it difficult to achieve lexical diversity in their EFL creative writing. It necessitates not just having a strong language command but also knowing how to use words in ways that are successful and appropriate for the required context.

Various research projects have been investigating lexical diversity in ESL/EFL students' creative writing. Some of the prominent findings of these studies will be discussed in this section of the thesis.

For instance, Failasofah and Hazim Alkhrisheh focused their research on the assessment of lexical diversity present in the creative writings of ten EFL Indonesian students of a public university in Jambi [18]. The lexical diversity indices obtained in the course of this analysis ranged from 23,1% to 40,5%, the mean being 31,23%. The study also revealed that the students' English language proficiency levels did not have a positive correlation with the lexical diversity of their writings.

Another research that concentrates, among other criteria, on the lexical diversity of EFL creative writing was conducted by Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong [59]. In this study, the focus group consisted of 64 Vietnamese students, whose creative writings, namely descriptive essays about a university (the University task) and a supermarket in the participant's hometown (the Supermarket task), were analyzed. In the course of the research it was discovered that the lexical diversity in the University task ranged from 44% to 65%, the average amounting to 54%, whereas in the Supermarket task, it covered the spectrum from 36% to 64% with an average of 53%, which is almost equal to that of the first writing assignment. Thus, it can be calculated that an average lexical variation for both tasks constituted 53,5%.

Lexical diversity in EFL students' short stories was a subject of the research produced by Nata Permata and Rosyida Ekawati [16]. The material of the study consisted of 48 short stories written by the fifth-semester students of the English Study Program of the University of Trunojoyo Madura. The short stories under scrutiny were divided into eight groups according to the lexical diversity results they produced. Thus, the average lexical variation indices for each group are as follows:

- Group 1 - 40.57%;
- Group 2 - 34.63%;
- Group 3 - 40.71%;
- Group 4 - 38.71%;
- Group 5 - 32.91%;
- Group 6 - 31.61%;
- Group 7 - 39.55%;
- Group 8 - 34.04%.

It may be calculated that the average lexical diversity for all 8 groups is 36.375%. The study also determined that the majority of the short stories under scrutiny

had quite low lexical richness which may be attributed to the style of writing, since in creative writing the language is usually figurative and connotative, not requiring the author to use an extremely sophisticated lexicon.

Furthermore, Siti Aisah Ginting conducted a study with the aim of investigating the influence of gender factor on the properties of the students' creative writings [20]. The group of participants consisted of 20 male and 20 female representatives aged 18-22, who were EFL students of the English Language and Literature Department, Faculty of Language and Arts, State University of Medan. The writing assignment based on which lexical diversity was estimated was a descriptive essay "My Favorite Place". In terms of lexical diversity, the research demonstrated that the descriptive writing produced by male students had an average lexical variety of 50.68%, while that of female participants amounted to 52.74%, from which it may be inferred that the texts written by female students were more lexically diverse. It can also be deduced that the average lexical diversity of creative writings produced by the whole subject group is 51.71%.

In addition, Khaled Layali and Ahmed Alshlowiy thoroughly examined the lexical diversity of creative writing produced by the students of a private university in Bahrain [41]. The participants of the study were 10 students at the beginner level and 10 students at the intermediate level aged 19-21, who have been studying English from primary to high school before entering university. Since the present study is focused on the writing quality of undergraduate Ukrainian students of English philology whose English language proficiency is intermediate and higher, only the results obtained from the group of students with intermediate English level at Bahrain University will be taken into consideration. Thus, the task for intermediate-level students was to write a narrative essay consisting of 150-200 words dealing with the topic of the Covid-19 pandemic. It was discovered that the lexical variation results obtained from the writings of these participants ranged from 43.8% to 70%, an average index being 56.9%. An important finding of this research was also the presence of a positive correlation

between lexical variation and writing proficiency regardless of either the gender or language proficiency level of the participants.

It is important to mention that the findings of the studies discussed above may be compared with the help of the chart presented below, where the research projects are marked as follows:

- A. Failasofah, Hazim Alkhrisheh, *Measuring Indonesian Students' Lexical Diversity and Lexical Sophistication*, Indonesia;
- B. Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong, *The Effect of Topic on EFL Writing by Vietnamese Tertiary Students: Insights from Combining a Lexical Richness Analysis with Student Self-Reports*, Vietnam;
- C. Nata Permata, Rosyida Ekawati, *Lexical Richness of Short Stories Written by EFL Students*, Indonesia;
- D. Siti Aisah Ginting, *Lexical Complexity on Descriptive Writing of Indonesian Male and Female EFL Learners*, Indonesia;
- E. Khaled Layali, Ahmed Alshlowiy, *Online writing of Covid-19 in Bahrain: An analysis of lexical diversity and writing proficiency*, Bahrain.

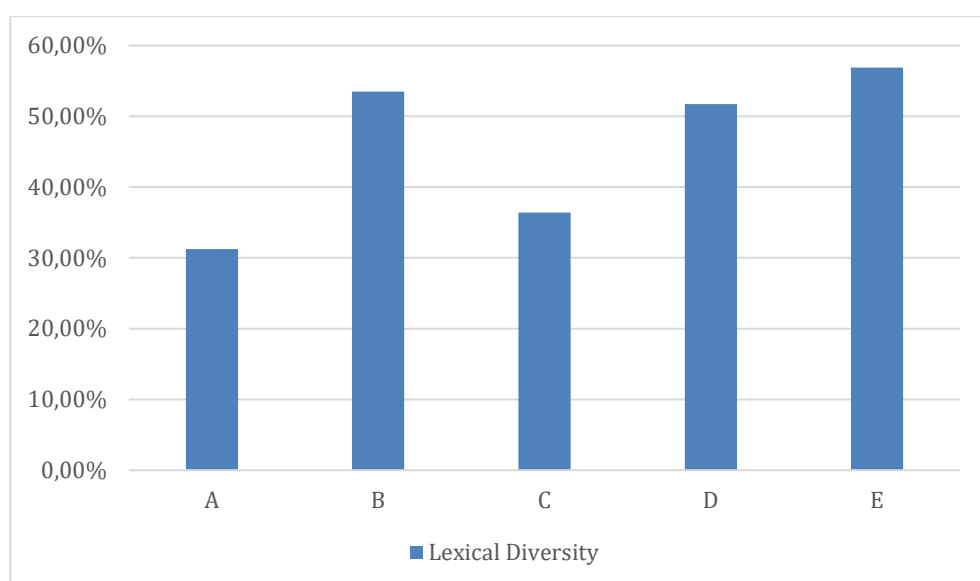


Chart 3. Lexical diversity of EFL students' creative writing

It can be inferred from Chart 3 that lexical diversity indices are not uniform in the studies discussed in this section of the thesis, varying from 31.23% to 56.9%. Furthermore, the average lexical diversity of the writings analyzed in all the studies examined above is 45.143%. With these statistics at our disposal, we will be able to analyze the lexical variation found in the Ukrainian EFL students' creative writings, which will enable us to compare them to the ones demonstrated by other EFL learners and further situate the Ukrainian EFL students' creative writings within the larger global EFL community.

2.2. Lexical diversity of EFL academic writing

The efficient communication of complicated ideas and arguments depends greatly on the use of accurate and varied terminology in academic writing. To satisfy the specific demands of academic discourse, academic writing in EFL requires a high level of lexical variety. A lack of lexical variety can degrade writing quality and make it more difficult to communicate effectively with readers. Making a positive impression on the reader and communicating the text's intended meaning both depend on the choice of proper terminology. Furthermore, a text's degree of sophistication and academic quality may be determined in large part by the appropriate vocabulary it uses.

The lexical variety of EFL academic writing is influenced by a number of variables, one of the most important of which is the lexical understanding of the students. Greater vocabulary knowledge allows for the employment of a larger variety of terms, which leads to greater lexical diversity in written work [40, p. 20].

The writing proficiency of the students is another criterion. Higher degrees of lexical variety are the outcome of learners' ability to utilize words more properly and effectively in their writing [5, p. 43].

Moreover, the discipline or area of study of the learners may have an impact on their lexical variety. For example, learners of such disciplines as linguistics, literature, and philosophy may employ more sophisticated and specialized language, leading to higher levels of lexical variety [29, p. 5].

Lexical variation in the academic writing of EFL students has been a focal point in various studies. Various researchers examined its connection to the quality of writing, linguistic proficiency, and its interrelation with other criteria of linguistic expertise.

For example, Mohammad Zare, Mohammad Sadegh Bagheri, Firooz Sadighi, and Ehsan Rassaei analyzed lexical diversity in argumentative essays of 5 male and 5 female Iranian IELTS candidates, aged 19-24, who were taking English courses at English institute in Shiraz [2]. Ten writing topics from IELTS task 2 were chosen as an assignment for the subject group, with the volume of each essay being required to comprise no fewer than 250 words. The research has revealed that the average lexical diversity found in the writings of the subject group is 58.44%.

Additionally, the lexical variety of EFL creative writing is one of the main focuses of another study by Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong [59]. The academic writings investigated in this research were persuasive essays produced by 64 Vietnamese students aged 20, the topics being the impact of Facebook (Facebook task) and the role of school or family in personal development (School task). The scholars have estimated that lexical diversity in the Facebook task ranged from 40% to 65% with the mean of 53%, while that of the School task amounted to 36%-62%, with an average of 50%. Thus, it can be calculated that the average lexical diversity for all the academic writings analyzed in this research is 51.5%.

Furthermore, lexical variation in non-native English speakers was analyzed by Hassanzadeh, Saffari, and Rezaei [26]. In the course of the research, which continued for seven weeks, 53 students at Vali-e-Asr University of Rafsanjan in Iran studying English as a foreign language (EFL) aged 19-24 were divided into two groups: the CACM condition (computer-aided concept mapping) and the standard outlining condition. The native language of the participants was Farsi, the level of English proficiency – intermediate and higher, and the major – English literature.

Each group was assigned a pre-test and post-test academic writing task. The pre-test task was to write a 300-word argumentative essay in 45 minutes. Subsequently, the students were divided into two groups - those who underwent CACM treatment and learned how to use a concept-mapping program for their writing, and those who used the strategy of outlining for their academic writing. Finally, the participants were given a post-test assignment, identical to the pre-test in topic and genre.

Thus, the average pre-test result of lexical diversity for the CACM group was 71.77%, while the post-test result amounted to 83.61%. At the same time, the mean for pre-test lexical diversity for the outlining group was 68%, with the post-test average index of 70,45%. It may be deduced that the average lexical diversity pre- and post-test index for both groups is 73.96%. Additionally, the research demonstrated the positive effect of computer-aided concept mapping and familiarity with the topic on lexical diversity in academic writing.

Another research focused on analyzing lexical diversity in undergraduate EFL students' writing was conducted by Hye Seung Ha [22]. The group of participants consisted of 35 undergraduate students at a Korean women's university in Seoul aged 20-27, Korea, who had different majors and years of study. The writing assignment based on which lexical variation was estimated was an argumentative essay in which a student had to express an opinion in relation to the issue of whether it is better to start learning English at an early age. In the course of the research, it was discovered that the average lexical variation of the participants' writings was 68.45%.

Thus, lexical diversity statistics discussed in this section of the thesis may be demonstrated on the chart provided below, where the studies mentioned are marked in the following way:

- A. Mohammad Zare, Mohammad Sadegh Bagheri, Firooz Sadighi and Ehsan Rassaei, *An Investigation of the Linguistic Complexity of IELTS Writing Topics Based on the Levels of Discourse Representation and the Degree of Meaning Coding*, *Cogent Education*, Iran;

- B. Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong, *The Effect of Topic on EFL Writing by Vietnamese Tertiary Students: Insights from Combining a Lexical Richness Analysis with Student Self-Reports*, Vietnam;
- C. Mohammad Hassanzadeh, Elahe Saffari, and Saeed Rezaei, *The impact of computer-aided concept mapping on EFL learners' lexical diversity: A process writing experiment*, Iran;
- D. Hye Seung Ha, *Lexical Richness in EFL Undergraduate Students' Academic Writing*, the Republic of Korea.

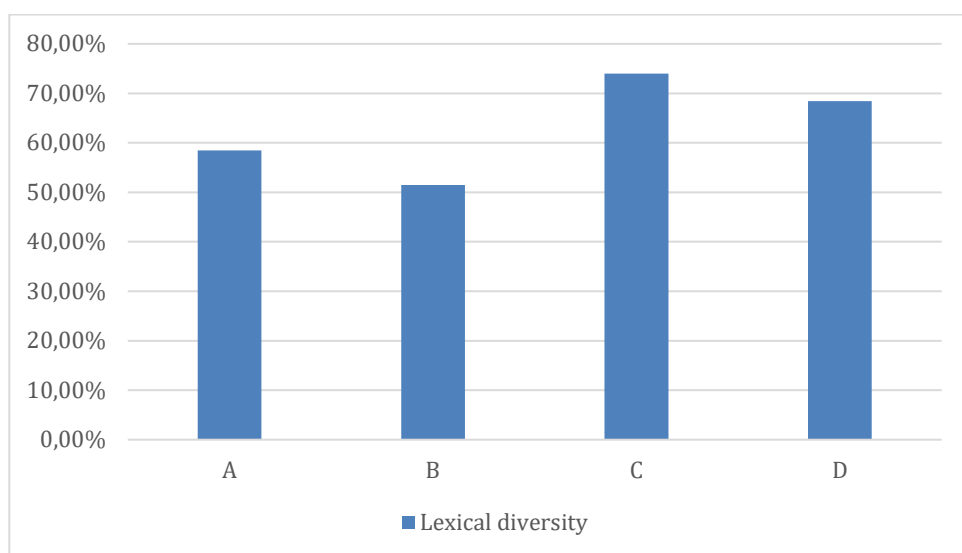


Chart 4. Lexical diversity of ESL students' academic writing

It is visible on Chart 4 that lexical density in the studies discussed in this section of the paper ranged from 51.5% to 73.96%, their average being estimated as 63.59%. As it was already mentioned, these data will be further used in the course of the analysis conducted in Chapter 3 of this paper in order to contrast the statistics obtained for the Ukrainian ESL academic writing with the lexical variety displayed by other ESL students and place the Ukrainian ESL students' creative writings within the larger, international context of EFL academic writing.

Conclusion

Lexical density and lexical diversity are considered to be reliable criteria frequently used to gauge students' writing skills and linguistic proficiency.

Therefore, their thorough examination was of paramount importance for this research.

Hence, lexical density is an indicator of information load in a sentence or in a text and is calculated in percentage as a ratio of content words to a total number of words in a sentence or in a text. Additionally, there is a correlation between lexical density and language guise, to wit it tends to be higher in written and lower in spoken modes.

In the context of EFL/ESL teaching, creative writing can be effectively utilized to encourage students to express themselves and explore their identities as well as to help them develop language proficiency and intercultural competence. The examination of various studies of lexical density in creative writing of international EFL/ESL students, enabled us to make an observation that lexical density in the texts examined in these studies varied from 36.55% to 51.69% with an average of 45.76%.

As for academic writing, it is characterized by a high level of formality, logical structure and precise language. From the information gathered from multifarious research papers concerning lexical density in EFL/ESL academic writings produced by undergraduate students from different countries it has been inferred that lexical density in the textual material analyzed in these studies varied from 31.45% to 52.2% with an average of 46.71%.

Another concept widely applied in the process of students' language proficiency assessment is lexical diversity, which refers to the amount of different words in a text. It is noteworthy that this concept is directly connected to a student's writing experience as well as to text genre.

Based on the lexical diversity calculations conducted in the previous studies, it was observed that lexical diversity in the creative writing of EFL/ESL undergraduate students with different linguistic and cultural backgrounds ranged from 31.23% to 56.9%, an average being 45.143%, while in academic writing it amounted to 51.5% - 73.96% with an average of 63.59%.

Overall, it may be stated that lexical density and lexical diversity are reliable and widely applied indicators of text quality. It may also be observed that academic writing is characterized by higher lexical density and lexical diversity. Additionally, both metrics are intertwined with various criteria, such as the formality level, students' linguistic proficiency and text genre.

CHAPTER II. DATA SELECTION AND PROCESSING FOR RESEARCH PURPOSES

2.1. Data for lexical density and lexical diversity study and selection criteria for creative and academic writing produced by Ukrainian EFL students

Data selection and collection as well as the choice of participants is a pivotal phase in any research, which determines the prospective reliability of its results. Thus, this chapter of the research provides a detailed and coherent description of the sources used and procedures applied in the course of data processing.

2.1.1. Participants and research setting

The participants of this study were eleven students of Taras Shevchenko National University of Kyiv, Ukraine, with the study program *English Philology and Translation, Western European Language*. They were informed about the research and provided consent for the collection and processing of the data. When entering the university at the age of 17 in 2017, all the participants had C1 level of English proficiency on the Common European Framework of Reference (CEFR), which was measured by a standardized test conducted by the Department of English Philology and Intercultural Communication of Taras Shevchenko National University of Kyiv in the first week of the first semester. All the students were citizens of Ukraine whose first language was Ukrainian.

The language of instruction in this program is English. This program does not include a separate course for academic writing. However, the total number of 60 study hours is allocated to creative writing in the first semester of the fourth year of study, which may have influenced the results obtained in Chapter 3 of this research.

Thus, by the end of the second year of study, the students have received 960 hours of instruction in general English. At the end of the second semester of the second and the fourth years of study each student is required to submit their own research – a course paper in the second year of study and a Bachelor's paper in the fourth. Additionally, it should be mentioned that the aspect of general English studied by the

students during the first semester of the fourth year of study was creative writing, which is particularly relevant for this research since the participants were expected to be familiar with different styles of writing and techniques of conveying their ideas precisely and coherently.

2.1.2. Data selection process

In order to analyze lexical density and lexical diversity in Ukrainian EFL students' writing, we compiled a corpus of their academic and creative writing produced in the second and fourth years of study. The choice of these years of study, in particular, was deliberate, since it would enable us to observe a dynamic within the group in the course of their studies. Thus, by the second year of study the students were expected to have gained a solid knowledge of the language and have developed essential writing skills, while by the fourth year of study, it is anticipated that they have advanced their linguistic competence, which may have an impact on their writing style and vocabulary use.

Therefore, the following writings produced by the participants during the second year of study were selected for the purpose of this study:

- *Academic writing* – course paper;
- *Creative writing* – essays produced for the assignment to write either a review on a service (opted for by Student 3 and Student 6) or to write an essay on the topic “What freedom means to me” (opted for by all the other students).

As for the texts produced during the fourth year of study, the ones listed below were chosen for the analysis:

- *Academic writing* – Bachelor's paper;
- *Creative writing* – an essay titled “Ukrainian Childhood” written for the creative writing course.

2.2. Quantitative analysis instruments and procedures

Following the selection stage, data has to be properly prepared in order to be used in the course of the research. This phase is crucial since the reliability of the results depends on it. Since statistical data was fundamental for the credibility of this research, it was quantitative analysis that we performed in the stage of data collection and processing. Thus, the tools and procedures utilized in this study will be examined in detail in the sections that follow.

2.2.1. Lexical density and lexical diversity calculation

Lexical diversity being dependent on the sample size, its values were calculated in the 250-word excerpts from both academic and creative writing produced by the students. The length of the material was specifically 250 words due to the fact that in order to obtain reliable data regarding vocabulary use it is suggested to analyze the texts which consist of at least 200 words [39, p. 310].

In order to obtain lexical density and lexical diversity indices, we resorted to the assistance of an online tool called *Textalyzer* [76]. It also calculated the number of characters per word and words per sentence, which enabled us to gain further insights concerning the participants' writing style and typical sentence structure.

Additionally, this tool was also applied in the course of lexical diversity analysis. In particular, it provided the lexical diversity percentage in the 250-word samples from academic and creative writings produced by Ukrainian undergraduate EFL students.

2.2.2. Frequency analysis

Frequency-based word lists are an inalienable part of linguistic research methodology. One of the scholars that emphasized the significance of frequency analysis was Sinclair, according to whom a researcher has a proclivity to be curious about the frequency with which a separate word occurs within the text under study [72, p. 30]. Frequency lists may be arranged according to different criteria, e.g. frequency, first appearance, or alphabetical order [64, p. 3].

Moreover, frequency lists are used to create Lexical Frequency Profiles. This idea was developed by Laufer and Nation in order to distribute the vocabulary utilized

by language learners into different categories in terms of frequency levels, such as the first thousand of most frequent words, the second thousand of most frequent words, and 570 most frequent words belonging to academic vocabulary. The scholars also stated that when creating a profile for more advanced students, apart from the first and second thousands of most frequent words, it should include information concerning the words belonging to academic vocabulary and other words, not belonging to the previously mentioned categories [39, p. 311].

In the process of lexical diversity analysis Lexical Frequency Profiles were generated with the help of an online vocabulary profiler VP Classic [84]. This tool provided the quantity of types, type-token ratio, and the number of tokens per type in the texts under scrutiny.

Moreover, it distributed the vocabulary used by the participants of the research into four categories: the first one thousand of the most frequent English words (K1), the second thousand of the most frequent English words (K2), lexical units from Academic Word List (AWL) [8], and off-list words (OLW), not encountered in the previously mentioned three categories. Thus, we obtained information on the percentage of vocabulary in the creative and academic writing of Ukrainian EFL students belonging to these categories, which enabled us to make assumptions about their vocabulary range and language command.

Conclusion

The participants of this research were 11 undergraduate students of Taras Shevchenko National University of Kyiv, Ukraine, with the study program English Philology and Translation, Western European Language. The subjects' native language is Ukrainian, while their level of English proficiency was determined as C1 in the Common European Framework of Reference (CEFR). All the students had the same curriculum and attended the same courses.

As for academic writings analyzed in the course of this study, these were course papers produced by the students in the second year of study and Bachelor's papers produced in the fourth. The participants' creative writing was represented by an

assignment to either write a review on a service or on a product or compose an essay on the topic “What freedom means to me” completed by the students in the second year of study; and by an essay called “Ukrainian Childhood”.

In this study, we resorted to quantitative research methods since it is statistical data that lies at its core. Thus, lexical diversity indices, an average number of words per sentence, and syllables per word were obtained with the help of *Textalyzer*. Moreover, it is this tool that calculated lexical diversity values in the 250-word samples of the students’ creative and academic writings.

Thus, frequency analysis, which is concerned with the number of single-word recurrences in a text [72, p. 30], was used in the course of lexical diversity analysis when creating Lexical Frequency Profiles. The tool which enabled us to do so was an online vocabulary profiler VP Classic. It provided crucial data regarding the number of types, type-token ratio, and the number of tokens per type in the textual material under analysis.

Additionally, this tool distributed the vocabulary utilized by the students into four categories according to the percentage of a given text they occupy: the first one thousand of the most frequent English words (K1), the second thousand of the most frequent English words (K2), lexical units from Academic Word List (AWL) [8], and off-list words (OLW), not encountered in the previous three categories.

CHAPTER III. COMPARATIVE ANALYSIS OF UKRAINIAN EFL WRITING

3.1 Lexical density Ukrainian students' English writing

3.1.1. Lexical density Ukrainian EFL students' academic writing

To conduct the analysis of lexical density in academic writings of Ukrainian ESL students we compiled a corpus of 11 Bachelor's papers, written at the end of the fourth year of study, and 11 course papers, written in the second semester of the second year of study, by the same group of participants. The subject group was represented by 11 undergraduate students of Taras Shevchenko University of Kyiv, whose study program was *English Philology and Translation, Western European Language*. The students' level of English was determined as C1, both during the second and fourth year of study. The course papers under analysis were produced by the students at the end of the second year of study, students being 18-19 years old, while Bachelor's papers were written by the participants of the research in the last semester of the fourth year of study when they were aged 20-21.

In the course of the research, academic writings were analyzed with the assistance of *Textalyzer*, which provided the data concerning lexical density, the average number of words per sentence, and an average number of syllables per word encountered in the textual material under scrutiny.

3.1.1.1. Lexical density of Ukrainian students' course papers

As it was already mentioned, the lexical density of course papers produced by the participants of the study was analyzed with the assistance of *Textalyzer*. This tool enabled us to calculate lexical density, as well as the average number of words per sentence and an average number of syllables per word in the academic writings of second-year undergraduate Ukrainian EFL students. Hence, Table 1 presents an opportunity to examine the indices obtained during the process of the analysis.

Student	Average Words/Sentence	Average Syllables/Word
Student 1	13.9	1.7
Student 2	12.2	1.7
Student 3	14.1	1.7
Student 4	12.4	1.7
Student 5	11.8	1.7
Student 6	17.8	1.7
Student 7	15.9	1.7
Student 8	18.9	1.7
Student 9	15.3	1.7
Student 10	19.9	1.7
Student 11	13.7	1.8
Mean	15	1.7
Standard deviation	2.68	0.0426

Table 1. Average sentence and word length in Ukrainian EFL students' course papers

After careful consideration of the indices provided in Table 1. it may be noticed that the average number of words per sentence in the course papers of the subject group varies between 11.8 and 19.9 with a mean of 15 and a standard deviation equaling 2.68. Since the length of an average English sentence is 15-20 words [9, p. 1-2], based on

the mean provided in the table above, the sentences of the course papers produced by the subject group are of average length. However, 54.55% of students utilized shorter than average sentences.

As for the average number of syllables per word, it is the same for all the writings – 1.7, except for the one produced by Student 11 – 1.8. This uniformity may serve to prove the homogeneousness of the group. Additionally, the indices may indicate that the participants of the research were inclined to implement mono- and disyllabic words in their academic writings, both during the second and fourth years of study.

The correlation between the average number of words per sentence and the number of syllables per sentence in the academic writings produced by the subject group in the second semester of the second year of study is displayed in the chart below:

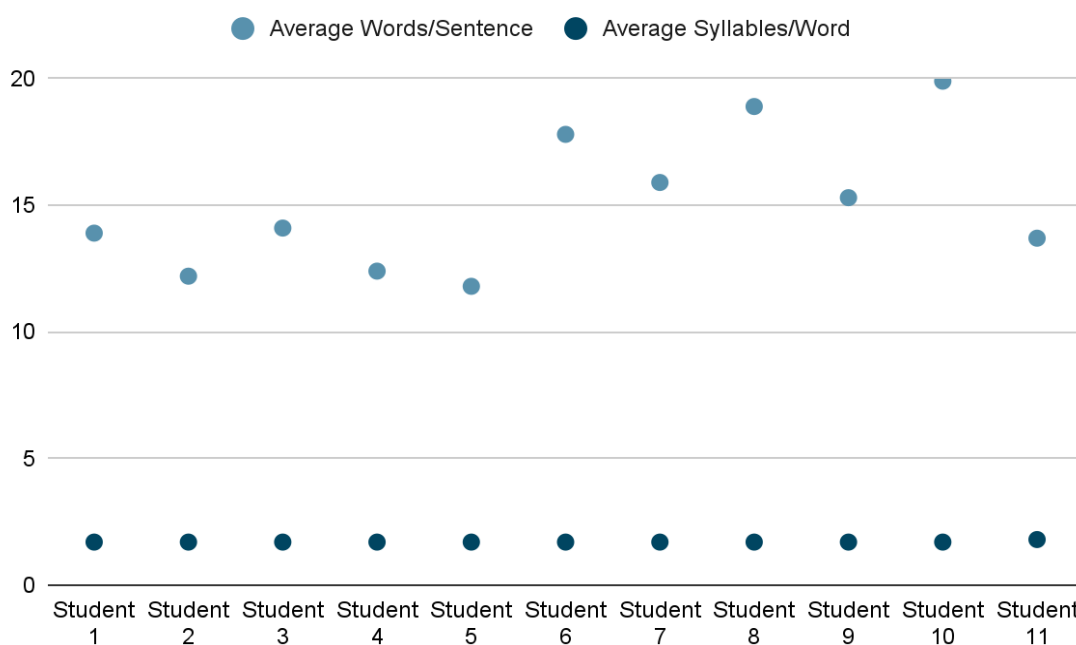


Chart 5. Average word and sentence length in the course papers of Ukrainian EFL students

Furthermore, the lexical density of the course papers written by the participants was calculated. The indices obtained with the help of *Textalyzer* are demonstrated in the table below:

Student	Lexical density
Student 1	46%
Student 2	50%
Student 3	47%
Student 4	51%
Student 5	53%
Student 6	48%
Student 7	47%
Student 8	49%
Student 9	50%
Student 10	48%
Student 11	51%
Mean	49%
Standard deviation	2.89

Table 2. Lexical density of Ukrainian EFL students' course papers

It may be inferred from the table provided above that the lexical density of Ukrainian EFL students' course papers ranged from 46% to 51%, the mean being 49% with a standard deviation of 2.89. This data indicates that the subject group is quite homogeneous in relation to the lexical density of their academic writing, the standard deviation of 2.89 suggesting that there are slight deviations from the average. This variability may be attributed to the differences in the topics of course papers, as well as to the personal writing style of the participants. One of the significant observations to be made is the fact that 5 students out of 11 scored slightly below the mean in terms of lexical density (Student 1, Student 3, Student 6, Student 7, Student 10).

It is vital to highlight that the statistical data concerning the lexical density of course papers written by Ukrainian EFL students are represented in the chart below:

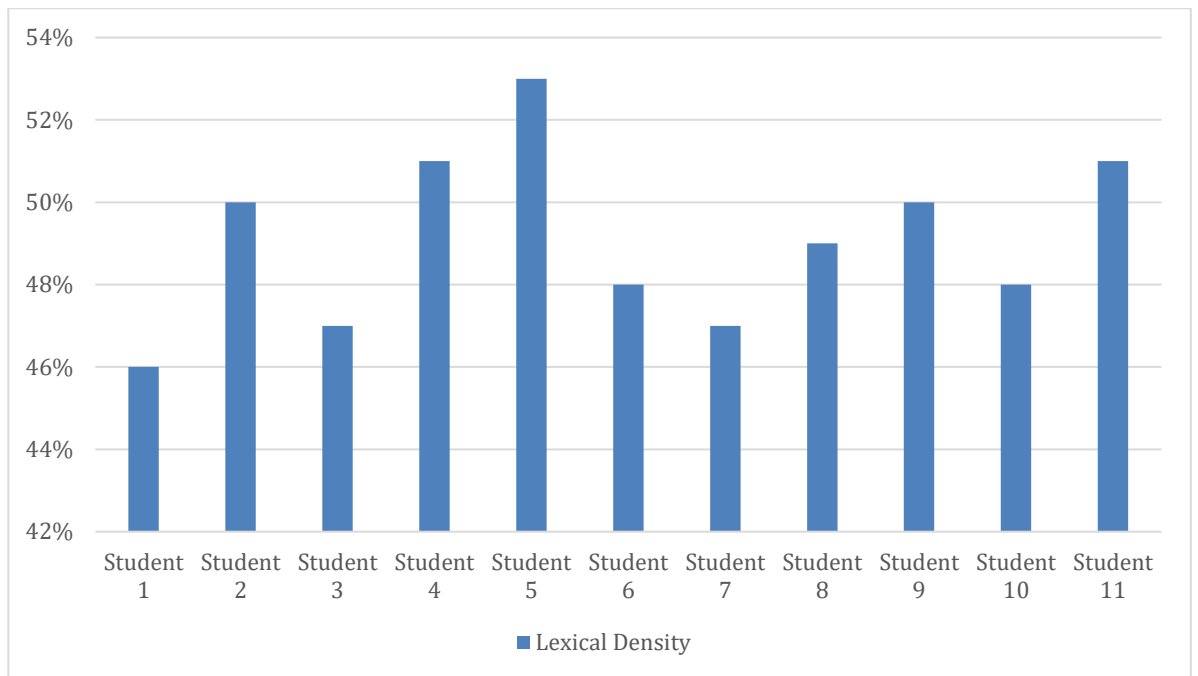


Chart 6. Lexical density of Ukrainian EFL students' course papers

It is visible on the chart provided above that the deviations of the lexical density in the course papers produced by the subject group are not particularly consequential. This may be attributed to participants being of the same age, having the same level of English proficiency and study program.

3.1.1.2. Lexical density of Ukrainian students' Bachelor's papers

In terms of the average number of words per sentence as well as an average number of syllables per word, these statistical data is provided in Table 3, which is represented below:

Student	Average Words/Sentence	Average Syllables/Word
Student 1	13.8	1.6
Student 2	12.9	1.7
Student 3	14.9	1.7
Student 4	16.3	1.7

Student 5	15.1	1.6
Student 6	14.3	1.6
Student 7	15.4	1.6
Student 8	14.1	1.6
Student 9	14.2	1.7
Student 10	22.4	1.6
Student 11	16.2	1.7
Mean	15.4	1.63
Standard deviation	2.93	0.02

Table 3. Average sentence and word length in Ukrainian EFL students' Bachelor's papers

As may be seen in Table 3 above, the average number of words per sentence as calculated for the writings under analysis ranges from 12.9 to 22.4. the mean being 15.4 with a standard deviation of 2.93. Taken into consideration that an average English sentence consists of 15-20 words [9, p. 1-2], based on the mean index, the sentences written by the subject group in their Bachelor's papers may be considered of average length. However, the papers produced by Student 1, Student 2, Student 3, Student 6, Student 8, and Student 9 (54.55% of the students) consisted of sentences that were slightly shorter than average. Additionally, academic writing produced by the Student 10 had generally longer sentences. These deviations may be connected to the writing style of the authors or the subjects of their studies. Therefore, other statistical data need to be analyzed in order to make a judgment concerning the linguistic competences of the subject group.

In comparison, it may be deduced that 50% of the students who had the proclivity to write shorter sentences (Student 1, Student 2, and Student 3) maintained this tendency in their further academic writings. The same conclusion may be drawn about Student 10, who produced writings with the longest sentences at the end of both the second and fourth years of study.

Interestingly, Student 5, whose course paper consisted of the shortest sentences among the subject group – 11.8 words per sentence – managed to significantly improve their writing in two years and utilize sentences of an average length. It is noteworthy that the same development is evident for Students 4, 5, and 11, whose academic writing was composed of shorter than average sentences during the second year of study but reached the average in the fourth. Additionally, although still slightly lower than average, the academic writings of Students 2 and 3 display an increment in average sentence length. Thus, it may be assumed that the average sentence length depends on the personal style of the writer. However, this presupposition needs to be further investigated.

Additionally, it is peculiar that the average number of syllables per word is consistent among the subject group. Varying between 1.6 (5 students) and 1.7 (6 students) with a mean of 1.63 and a standard deviation of 0.02. This statistic indicates that the students utilized predominantly shorter mono- and disyllabic words, which could be conducive to a higher accessibility of the textual material and, potentially lower lexical density, which will be further investigated in this section of the paper.

It should be mentioned that *Chart 7* is provided below in order to demonstrate sentence and word length statistics calculated for academic writing of fourth year undergraduate Ukrainian EFL students, who participated in this research:

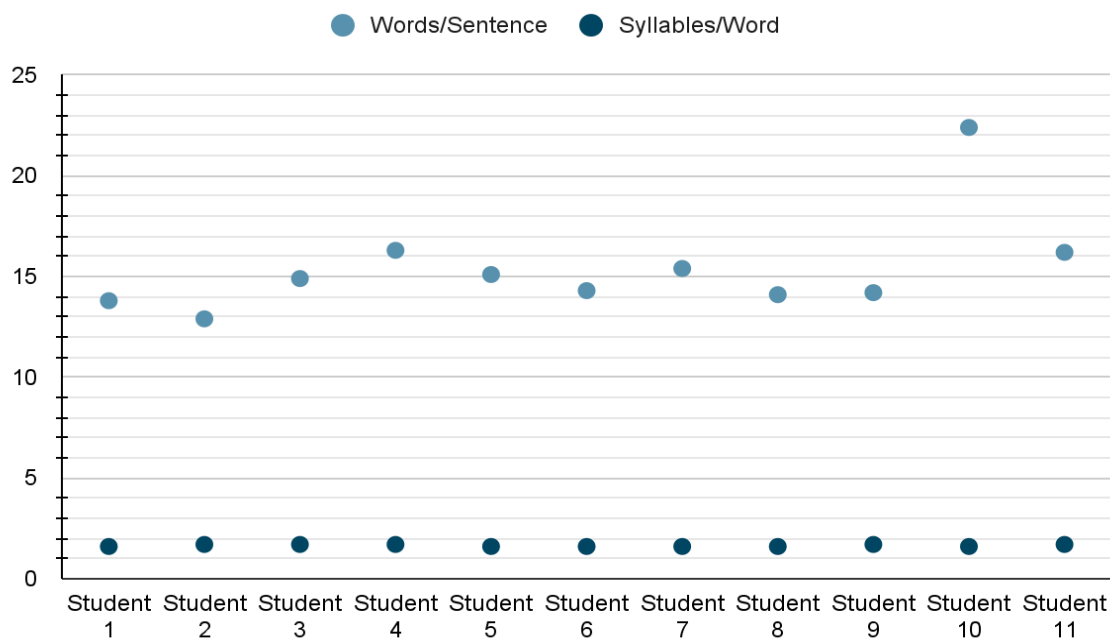


Chart 7. Average word and sentence length in the Bachelor's papers of Ukrainian EFL students

With the help of *Chart 7*, the deviation regarding the amount of words per sentence may be noticed. However, it is also visible that word length is consistent among the subject group.

Additionally, the lexical density results obtained in the course of the research are represented in Table 4 below:

Student	Lexical density
Student 1	45%
Student 2	51%
Student 3	51%
Student 4	49%
Student 5	49%
Student 6	49%
Student 7	48%

Student 8	50%
Student 9	53%
Student 10	46%
Student 11	49%
Mean	49%
Standard deviation	3.35

Table 4. Lexical density of Ukrainian EFL students' Bachelor's papers

It may be inferred from Table 4 that the lexical density of fourth-year undergraduate EFL students from Ukraine ranges from 45% to 53%. indicating a general consistency in the lexical density calculated in the corpus. Additionally, the mean score is 49% with a standard deviation of 3.35, which suggests that, on average, the students participating in the research produced written material with a moderate level of lexical density.

As for Student 2, Student 3, Student 8, and Student 9, the lexical density obtained from their Bachelor's papers is slightly above the mean, based on which it may be assumed that they demonstrate a more proficient level of academic language and can use it effectively in their own research. What is more, the suggestion which might be made based on this data is that these students can utilize a more sophisticated vocabulary. This presupposition will be further either disproved or substantiated in the next section of this study, where there will be provided an analysis of the lexical diversity of these writings.

However, there are students, who produced Bachelor's papers with lexical density results lower than average. For example, Student 1 and Student 10 scored the lowest, which may provide the ground to conjecture that their writing is less lexically loaded, hence is less effective in communicating ideas precisely and coherently.

It is of paramount importance to be mentioned that the data provided above may be compared with the help of the following chart, on which it is visible that

the deviation between the scores of lexical density demonstrated by the writings produced by the participants is minimal:

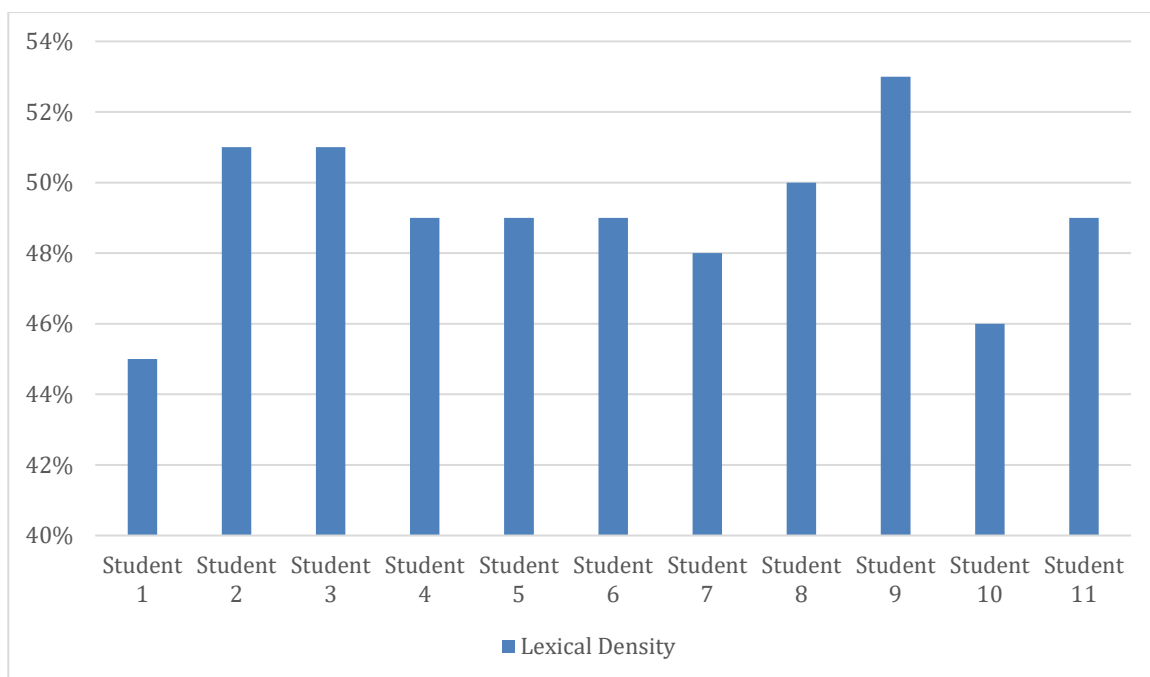


Chart 8. Lexical density of Ukrainian EFL students' Bachelor's papers

As for the fact that the data discussed in this section of the paper does not demonstrate drastic changes in terms of lexical density, the average number of words per sentence, and the average number of syllables per sentence, a potential reason for this may be that the participants already had a C1 level of English during the second year of study, therefore their linguistic competence was not improved as significantly as it could have if the students experienced a transition from one level of linguistic proficiency to another. Additionally, strict requirements for academic writing may be a contributing factor since they prevent the students from engaging in linguistic experimentation in their writing. However, it is still possible to witness the development of academic writing experienced by certain participants.

Interestingly, when compared to their earlier academic writings, only Student 3 and Student 7 managed to increase the lexical density of their textual material, while Student 1, Student 7, and Student 10 maintained a proclivity for academic writing with lower lexical density, Student 1 and Student 10 even experiencing a slight decrease in lexical density of their academic writing over time. However, this tendency is not an

isolated case, since the writings analyzed, produced by Student 4, Student 5, and Student 11, although staying within the average and above, demonstrate a decrease in lexical density over the period between the second and the fourth levels of study (see Appendix 1). These fluctuations may have various causes, such as the change of the area of research. lack of sufficient practice with academic writing etc.

Additionally, in order to compare the lexical density of Ukrainian EFL students' academic writing with the results of international EFL students discussed in Chapter 1 of this paper, an illustrative chart was compiled, where the obtained data is marked as follows:

- A. Carmen Gregori-Signes, Begoña Clavel-Arroitia, *Analysing lexical density and lexical diversity in university students' written discourse, Spain, first-year undergraduate students;*
- B. Rim Kouachi, *A Corpus-Based Study of the Lexical Density and Readability of Student's Academic Writings The Case of Master One EFL Students at Larbi Ben M'hidi University, Algeria, graduate students;*
- C. Hermawati Syarif and Rahmi Eka Putri, *How lexical density reveals students' ability in writing academic text, Indonesia, graduate students;*
- D. Bao Trang Thi Nguyen, Jonathan Newton and Tran Ngoc Quynh Phuong, *The Effect of Topic on EFL Writing by Vietnamese Tertiary Students: Insights from Combining a Lexical Richness Analysis with Student Self-Reports, Vietnam, third-year undergraduate students aged 20;*
- E. Lexical density of Ukrainian EFL students' academic writing, whose index is constant both in the second and the fourth undergraduate years of study.

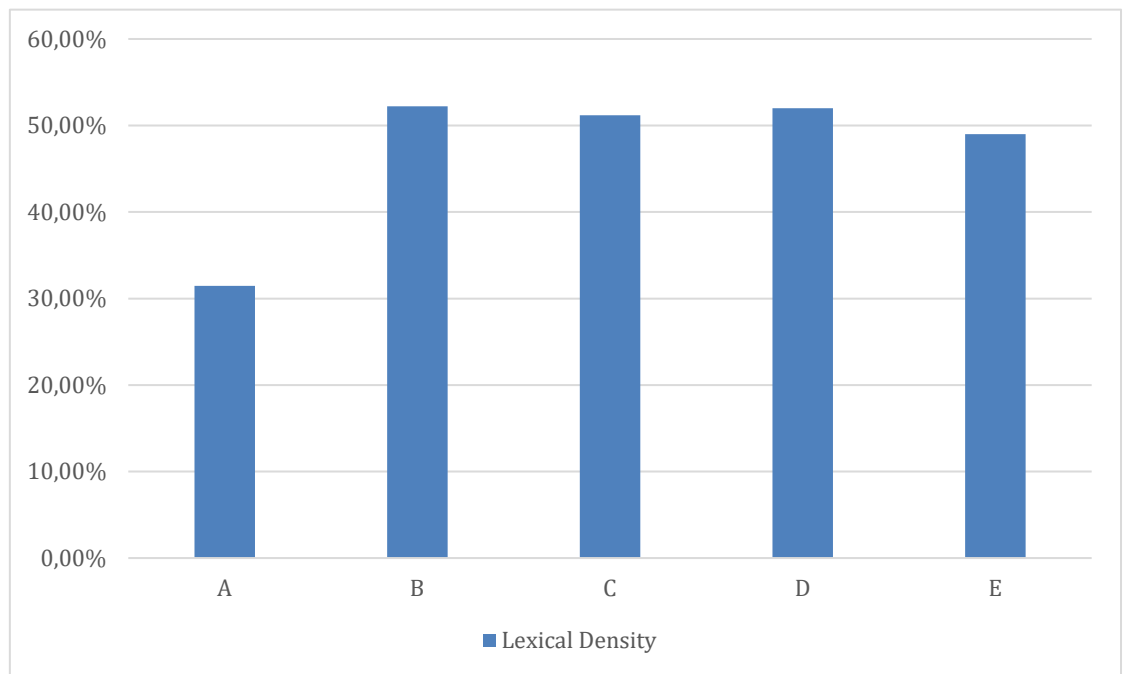


Chart 9. Lexical density of international and Ukrainian EFL students' academic writing

As it may be deduced from the chart, the results of Ukrainian students are positioned second to last in comparison to the ones demonstrated by international EFL students. However, they are still scored very closely to the participants from other countries. Moreover, given that the average lexical density in the studies reviewed in the first chapter of this study is 46.71, it may be stated that lexical density of Ukrainian EFL students' academic writing is comparable to that obtained in the course of other researchers' studies. Despite the fact that lexical density is only one of the multifarious facets of writing evaluation, the results suggest that Ukrainian EFL students are not disadvantaged in comparison to their peers from other countries in this regard. Additionally, the Algerian and Indonesian research participants are graduate students, which provides them with age- and experience-related advantages, as well as presumably greater exposure to academic and professional literature which is naturally reflected in their richer vocabulary.

It is critical to mention that further analysis of the lexical diversity of the corpus of academic writing produced by the subject group will enable us to make more

conclusive judgments concerning the quality and development of academic writing skills demonstrated by Ukrainian EFL students.

3.1.2. Lexical density Ukrainian EFL students' creative writing

The lexical density measurement of the creative writing produced by Ukrainian EFL students during the second and the fourth years of undergraduate studies was conducted in the same manner as that of academic writing. Specifically, we collected one essay per student in the same subject group consisting of 11 people written during the second and the fourth years of study, i.e. 22 essays in total. The subjects, as it was mentioned in the previous section of the paper, were 11 undergraduate students of the program *English Philology and Translation, Two Western European Languages* at Taras Shevchenko University of Kyiv, with C1 level of English proficiency. It is noteworthy that during the second year of study the subjects' age was 18-19 years old while during the fourth year of study they were 20-21 years old.

As in the process of academic writing analysis in terms of lexical density, we opted for the assistance of *Textalyzer*, which enabled us to obtain such valuable information as lexical density, the average number of words in a sentence as well as an average number of syllables in a word as estimated in the creative writings under scrutiny.

3.1.2.1. Lexical density of second year undergraduate Ukrainian students' creative writing

The creative writing samples analyzed in this section of the thesis were essays produced by the participants during the second year of study for the assignment which involved writing either about the concept of freedom and its meaning (chosen by 9 students) or writing a review on a service (chosen by two students). The results of average word and sentence lengths are provided in Table 5.

Student	Average Words/Sentence	Average Syllables/Word
----------------	-------------------------------	-------------------------------

Student 1	13.4	1.4
Student 2	17.2	1.7
Student 3	14.2	1.3
Student 4	16.6	1.6
Student 5	13.0	1.4
Student 6	13.6	1.4
Student 7	16.2	1.6
Student 8	15.5	1.5
Student 9	18.5	1.6
Student 10	19.5	1.6
Student 11	20.9	1.4
Mean	16.0	1.5
Standard deviation	3.0199	0.122

Table 5. Average sentence and word length in Ukrainian second-year EFL students' creative writing

As demonstrated in the table above, the average number of words per sentence in the writings analyzed varied from 13.0 to 20.9 with a mean of 16.0 and a standard deviation of 3.0199. Given that the average English sentence length is 15-20 words, it may be stated that the subject group in general produced sentences of average length, with the exception of those written by Student 1, Student 3, Student 5, and Student 6, whose sentences were shorter than average; and Student 11, who slightly surpassed the average [9, p. 1-2].

Peculiarly, an average number of syllables per word in the creative writing of the subjects demonstrates a greater variability than that of academic writing. Thus, this index ranges from 1.3 to 1.7, the mean being 1.5 and the standard deviation 0.122. This disparity with academic writing may indicate that the students were experimenting with

language using more varied vocabulary in their creative essays in contrast to academic writing, which entails the use of specific terminology and formulaic language. The fact that the words utilized by the subjects in creative writing are generally shorter than those encountered in academic text written by them may be stipulated by the informal style of creative writing contributing to its higher accessibility among the general audience.

It is noteworthy that the statistical data discussed above is demonstrated in the chart below:

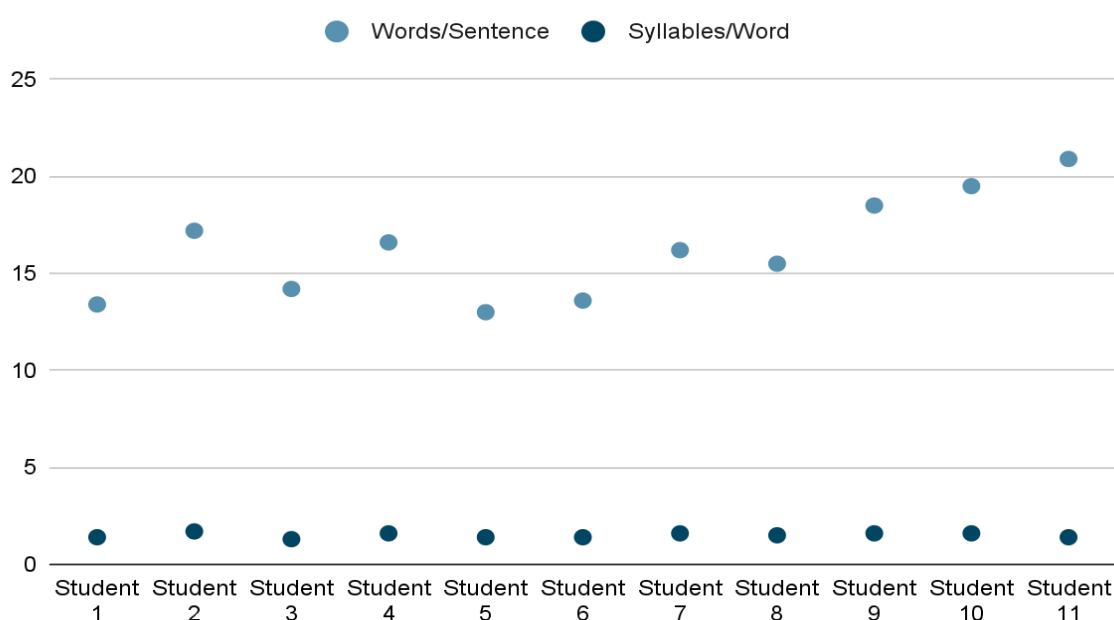


Chart 10. Average word and sentence length in Ukrainian second-year EFL students' creative writing

Chart 10 provides a graphic representation of the data discussed above, which enables us to see the fluctuations in the amount of words per sentence among the subject group as well as some consistency in terms of the number of syllables per word. The divergence in the sentence length criterion may be attributed to a less regimented style of writing, which allows more freedom to the writers, enabling them to experiment with the content and form of their texts.

Furthermore, *Textalyzer* provided us with the lexical density indices measured in the creative writings produced by the subjects during the second year of undergraduate studies. It is in the table below where these figures are represented.

Student	Lexical density
Student 1	38%
Student 2	42%
Student 3	34%
Student 4	33%
Student 5	40%
Student 6	40%
Student 7	38%
Student 8	37%
Student 9	36%
Student 10	42%
Student 11	36%
Mean	38%
Standard deviation	5.0453

Table 6. Lexical density in Ukrainian second-year EFL students' creative writing

According to the data presented in the table above, the lexical density is the creative writing of the subjects during the second year of Bachelor's degree studies ranges from 33% to 42%, with a mean of 38% and a standard deviation of 5.0453. This data suggests that the students produced writing with lexical density which is lower than average. Additionally, the standard deviation of 5.0453 suggests that the spread of the data points is relatively large, indicating that the students were less restricted in the use of specific vocabulary and wrote texts closer in style to the spoken mode.

In addition, the writings produced by Student 2, Student 5, Student 6, and Student 10 are characterized by lexical density which is higher than the mean. This means that there is a higher proportion of content words to function words in their essays. Thus, these writings may be seen as more precise, detailed, and informative, since content words are carriers of the meaning in the text.

Contrarily, the lexical density of the creative writing samples written by Student 3, Student 4, Student 8, Student 9, and Student 11 is lower than the mean. This may indicate that these subjects were less focused on providing precise and detailed information and opted for a more accessible style of writing, with emphasis on conveying the atmosphere, mood, and tone with the use of figurative language. Importantly, the assumptions about the lexical richness and diversity of the texts analyzed based on their lexical density will be further investigated in the following sections of this thesis.

The data concerning lexical density in the textual material under scrutiny is visually represented in the following chart:

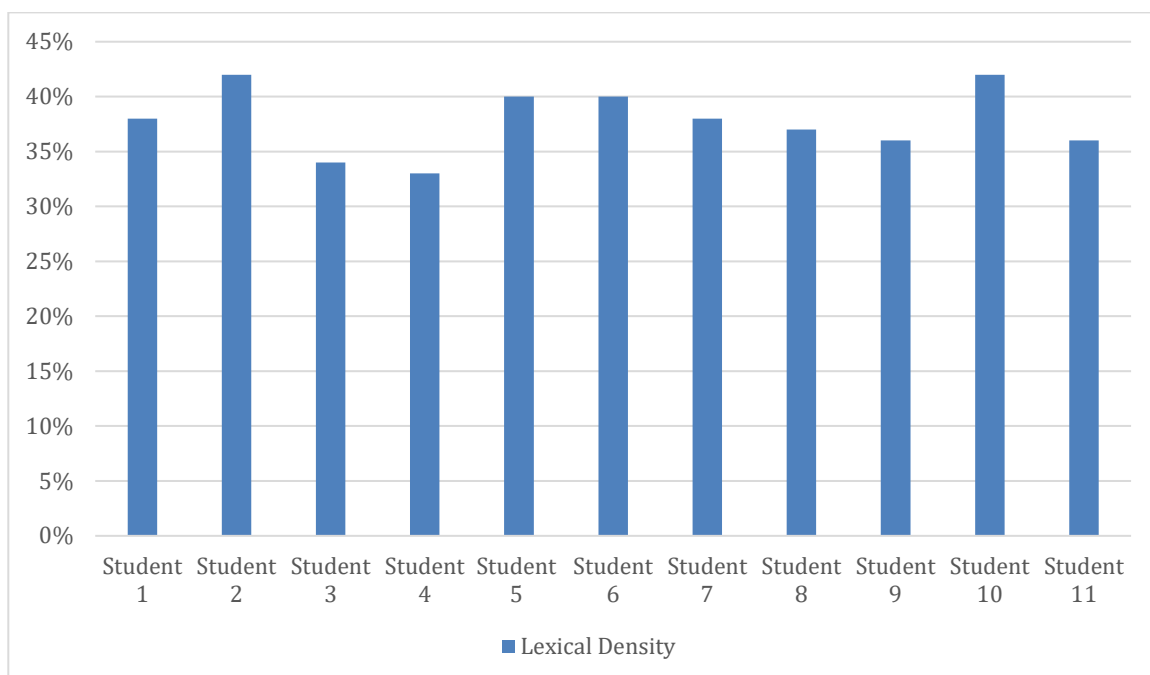


Chart 11. Lexical density in Ukrainian second-year EFL students' creative writing

As it is visible on the chart, lexical density in the writings under analysis is not uniform and varies considerably. This may be attributed to the fact that creative

writing allows greater freedom of artistic expression, encouraging students to develop their personal writing style and convey their ideas in more informal and creative ways than it is possible in academic writing.

3.1.2.2. Lexical density of fourth year undergraduate Ukrainian students' creative writing

During the fourth year of study, undergraduate students who participated in the research attended a course of creative writing. It is the essays produced for this course that were selected to represent the writing of the subject group. Specifically, each student's essay (11 essays in total) that recalls an aspect of their childhood which helped shape their personality was taken for the analysis.

As in the previous sections of the research, average sentence and word length were obtained with the help of *Textalyzer*. This data is illustrated in Table 7:

Student	Average Words/Sentence	Average Syllables/Word
Student 1	19.1	1.4
Student 2	19.7	1.5
Student 3	33.3	1.4
Student 4	17.0	1.5
Student 5	10.8	1.5
Student 6	10.2	1.2
Student 7	20.2	1.4
Student 8	16.8	1.4
Student 9	28.8	1.3
Student 10	15.6	1.4
Student 11	25.9	1.4

Mean	20.36	1.4
Standard deviation	8.711	0.0775

Table 7. Average sentence and word length in Ukrainian fourth-year EFL students' creative writing

The data provided in the table above presents a heterogeneous picture. Interestingly, the average sentence length varies extraordinarily widely among the subjects – from 10.2 to 33.3 words per sentence with a standard deviation of 8.711. Borne in mind that an average English sentence length is 15-20 words, with a mean of 20.36, the sentences comprising the writings under analysis may be considered generally longer than average [9, p. 1-2]. The subjects who exceeded this standard are Student 3, Student 7, Student 9, and Student 11, while those who utilized considerably shorter than average sentences in their writing are Student 6 and Student 7. The outstanding disparity which may be observed in the data may ensue from the development of a personal style which arose from the freedom to experiment with creative writing.

Surprisingly, the words used by the participants of the research are rather short, an average number of syllables per word in the writings under scrutiny varying from 1.2 to 1.5. with a mean of 1.4 and a standard deviation of 0.0775. This signifies that the students favored monosyllabic words in their writings, regardless of sentence length. Compared with the average word length in academic writing produced in the same year, a considerable difference in the word length indices. While average word length, as estimated in the number of syllables, amounts to 1.6333 in academic writing produced by the subjects during the fourth year of study, in creative writing produced in the same year it equals only 1.4. Such heterogeneity may be attributed to the specific features of academic and creative writing on the lexical level - viz. the abundance of terminology and low-frequency vocabulary in academic writing, which are usually longer than mid- and high-frequency words. It is noteworthy that the data concerning

average word and sentence length calculated in the fourth year undergraduate Ukrainian EFL students' creative writing is illustrated by Chart 12 below.

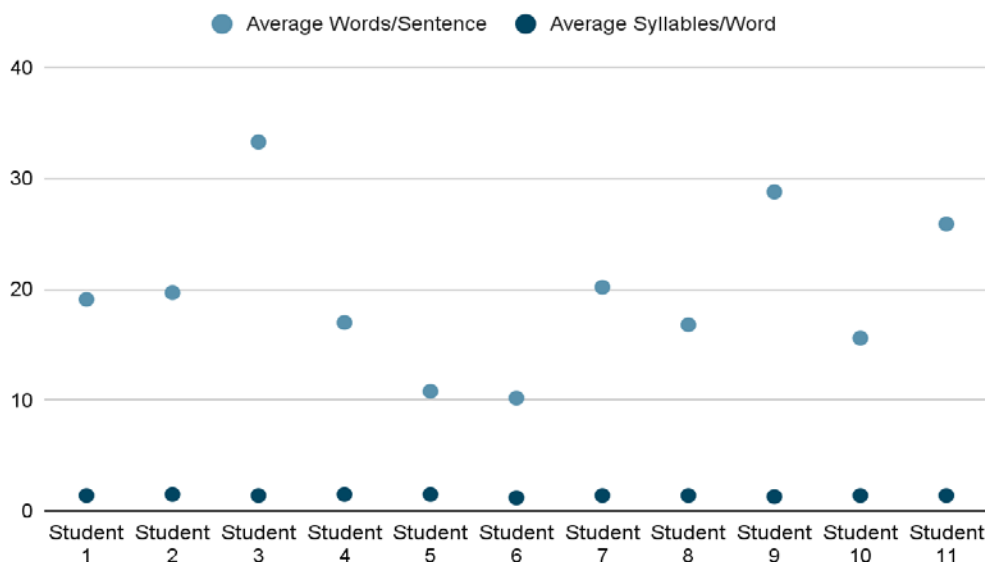


Chart 12. Average word and sentence length in Ukrainian fourth-year EFL students' creative writing

Additionally, Table 8 provides the lexical density values of the creative writing produced by the fourth year Ukrainian EFL students, which were obtained with the help of *Textalyzer*.

Student	Lexical density
Student 1	44%
Student 2	38%
Student 3	34%
Student 4	40%
Student 5	40%
Student 6	41%
Student 7	36%
Student 8	39%

Student 9	37%
Student 10	45%
Student 11	40%
Mean	39.45%
Standard deviation	3.29

Table 8. Lexical density in Ukrainian fourth-year EFL students' creative writing

As provided in the table above, the lexical density of the writings under analysis varies from 36% to 45% with the mean of 39.45% and a standard deviation of 3.29. Overall, these data suggest that the essays of the subject group had a moderate lexical density. The standard deviation index of 3.29 demonstrates a certain variation across the subjects' writings, with some students producing textual material with a higher proportion of content words, e.g. Student 1, Student 10, and some including more function words into their writings, e.g. Student 3, Student 7.

Interestingly, 54.54% of the subjects' writings had lexical density higher than the mean, which indicates that the vocabulary utilized in the textual material was more precise and complex. The lower lexical density of other writings may be attributed to the authors' desire to pursue a more effortless writing style since the writing assignment was quite simple – to describe a formative childhood experience.

It is worth mentioning that the chart presented below provides a graphical depiction of the lexical density indices of creative writings discussed in this section of the thesis.

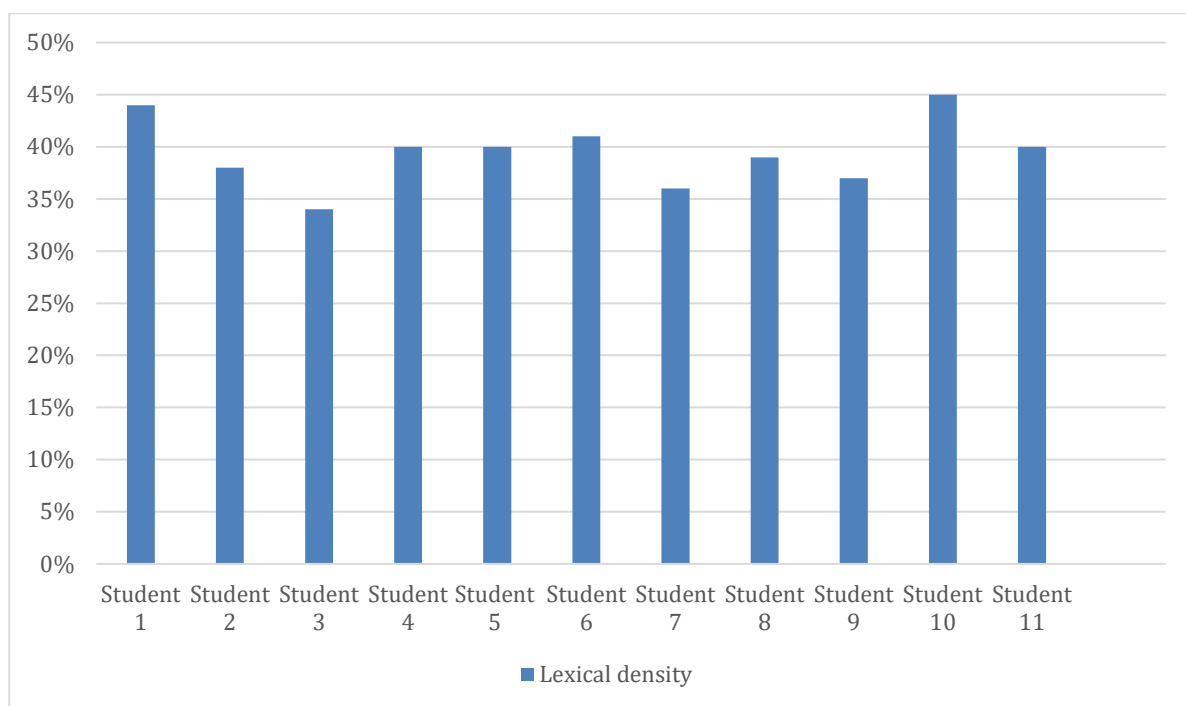


Chart 13. Lexical density in Ukrainian fourth-year EFL students' creative writing

Having compared the lexical density indices calculated in the second- and fourth-year Ukrainian EFL students' creative writings, an upward trend may be noticed (see Appendix 2). In particular, the mean has grown from 38% in the second year to 39.45% in the fourth year demonstrating that the students' use of content words has improved, which indicates an increase in the depth and sophistication of their writing. Significantly, 7 out of 11 students (63.64% of the participants) have produced writings with higher lexical density during their fourth year of study, to wit Student 1 (from 38% to 44%), Student 4 (from 33% to 40%), Student 6 (from 40% to 41%), Student 8 (from 37 to 39%), Student 9 (from 36% to 37%), Student 10 (from 42% to 45%), Student 11 (from 36% to 40%). Additionally, lexical density in the creative writing of 2 students (18.18% of the participants) stayed consistent during both the fourth and second years of study, these subjects being Student 3 (34%) and Student 5 (40%). However, the writings of two other students (18.18% of the participants), namely Student 2 and Student 7 demonstrated a slight decrease in lexical density from 42% to 38% and from 38% to 36% respectively.

Furthermore, it may be observed that the standard deviation slightly decreased from 5.0453 in the second year to 3.29 in the fourth year. This suggests that the variance among the group of participants decreased with the progression of their studies. There is a possibility that this decrease is a result of the students receiving more consistent instruction on writing as they advanced through their studies. Additionally, it may be linked to the participants attending the same block of courses throughout their studies, thus adopting a similar writing approach based on the unified academic curriculum that they had.

Thus, it may be stated that in the creative writings analyzed a negative correlation between lexical density and variance among the students may be observed, since with the increase in lexical density over time there is a visible trend for the decrease in lexical density variance among the participants. These findings reveal that in the period between the second and fourth years of study, the participants develop their linguistic proficiency which is indicated by their more frequent use of content words, which in turn may demonstrate a growth in terms of complexity, sophistication, and unambiguity of their writing.

Additionally, using the information gathered in Chapter 1 of this research, we created a chart that demonstrates the position of the lexical density of Ukrainian EFL students' creative writing in relation to that of international students. In this chart, the following markings were used:

- A. Carmen Gregori-Signes, Begoña Clavel-Arroitia, *Analysing lexical density and lexical diversity in university students' written discourse*, Spain, first year undergraduate students;
- B. Abdalla Elkheir Elgobshawi, Mohammed Aldawsari, *Lexical density as improvement indicator in the written performance of EFL majors*, Saudi Arabia, undergraduate college students;
- C. Siti Aisah Ginting, *Lexical Complexity on Descriptive Writing of Indonesian Male and Female EFL Learners*, Indonesia, age of the participants – 18-22 years old;

- D. Daw Thida, *Lexical Density and Readability of Students' Writing*, Myanmar, postgraduate students;
- E. Josephine Uma, *A Study of Lexical Richness in ESL Students' Writing*, Malaysia, undergraduate students.
- F. Lexical density of Ukrainian second-year undergraduate EFL students' creative writing, age of the participants – 17-18 years old;
- G. Lexical density of Ukrainian fourth-year undergraduate EFL students' creative writing, age of the participants – 20-21 years old.

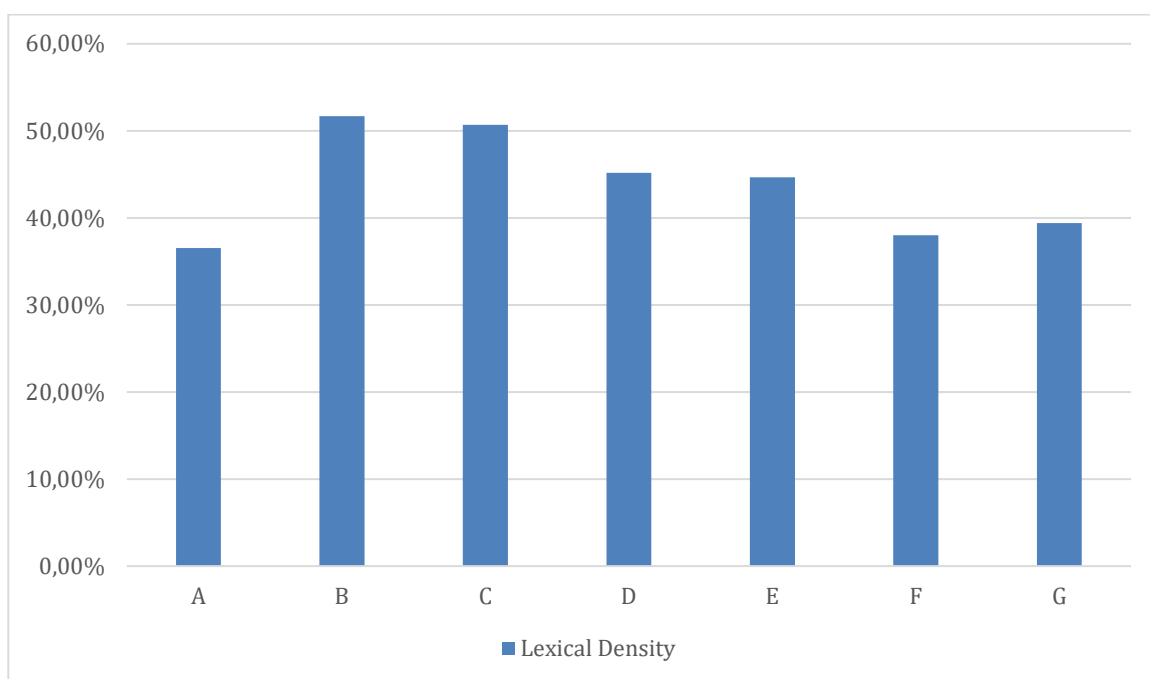


Chart 14. Lexical density of international and Ukrainian EFL students' creative writing

It is visible on the chart above that Ukrainian EFL students outperformed only the participants of the study conducted in Spain by Carmen Gregori-Signes and Begoña Clavel-Arroitia. Borne in mind that the average lexical density of the creative writings analyzed in the studies to which the indices obtained in the course of this research are compared is 45.76%, it may be stated that the performance of Ukrainian EFL students in terms of lexical density in creative writing is exhibited a minor decrement as compared to that of their international peers. Nevertheless, the lexical density indices

of Ukrainian EFL students are not significantly lower than those of other international students.

3.2. Lexical diversity of Ukrainian students' English writing

3.2.1. Lexical diversity Ukrainian EFL students' academic writing

In the course of analyzing lexical diversity in Ukrainian EFL students' academic writing, the corpus of 11 course papers and 11 Bachelor's papers, whose lexical density was measured in the previous sections of this chapter, was used. Additionally, no alterations were made to the group of participants, them being 11 undergraduate students of English Philology and Translation, Two Western European Languages at Taras Shevchenko University of Kyiv, whose level of the English language proficiency was C1.

In addition, the lexical diversity data analyzed in the following sections of the study were obtained with the help of *Textalyzer*, which calculated the lexical diversity percentages in the academic writings under analysis. Additionally, crucial information concerning the lexical variation of the texts under scrutiny, i.e. the percentage of lexical units from Academic Word List, the list of the first thousand most frequent English words, the list of the second thousand of most frequent English words, the number of tokens, types, type-token ratio and the number of tokens per type, was obtained by resorting to the VP-Classic vocabulary profiler.

It is of paramount importance to be taken into consideration that since the lexical diversity index is dependent on the sample size, 250-word excerpts from the conclusions of both course and Bachelor's papers were chosen as representative samples for the analysis.

3.2.1.1. Lexical diversity of Ukrainian students' course papers

As it was already mentioned, the lexical diversity of the second-year Ukrainian EFL undergraduate students' academic writing was measured with the help of *Textalyzer*. The samples extracted from the participants' course papers for the analysis were 250-word extracts from the conclusions of the research projects. As a result,

lexical diversity percentages calculated with the assistance of the aforementioned tool are represented in the table below.

Student	Lexical diversity
Student 1	56%
Student 2	55%
Student 3	56%
Student 4	54%
Student 5	52%
Student 6	51%
Student 7	57%
Student 8	62%
Student 9	50%
Student 10	57%
Student 11	53%
Mean	54.82%
Standard deviation	5.44

Table 9. Lexical diversity of Ukrainian EFL students' course papers

Having examined the data provided in the table above, it may be inferred that lexical diversity in the textual material analyzed varies from 50% to 62% with a mean of 54.82%. Consequently, it may be suggested that the students produced academic writing with a moderate number of unique words. Additionally, the standard deviation index of 5.44 indicates that there is a fairly wide range of lexical diversity within the group of subjects. Moreover, six out of eleven students scored within one standard deviation of the mean, which indicates the relative heterogeneity within the group with the majority of the results being clustered around the mean.

Such a discrepancy may be connected to the topic of each student's research. Thus, the participants who were less familiar with the area of their research could utilize a narrower range of vocabulary. Additionally, it is usual for scientific writing to frequently reiterate terms and fixed expressions, which in turn reduces the lexical diversity of textual material.

In order to illustrate the lexical diversity data provided in Table 9, the following chart was compiled:

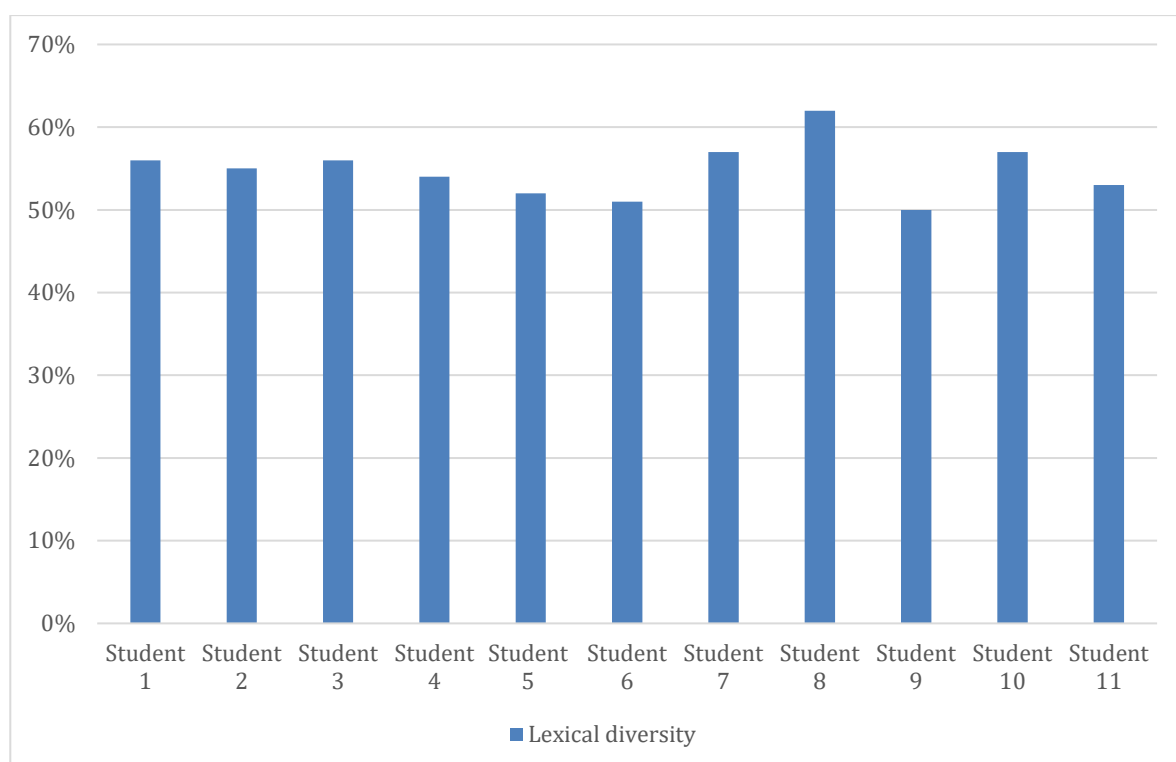


Chart 15. Lexical diversity of Ukrainian EFL students' course papers

Furthermore, VP-Classic vocabulary profiler software enabled us to develop Lexical Frequency Profiles of each student's writing. The table presented below demonstrates Lexical Frequency Profiles in terms of the number of tokens and types and their ratios. Thus, in this table, the following markings were used:

- TTR refers to the type-token ratio;
- TPT refers to the number of tokens per type.

Additionally, attention must be drawn to the fact that each sample has the same number of types (250) since 250-word excerpts from the conclusions of the participants' course papers were taken for the analysis.

Student	Types	TTR	TPT
Student 1	137	0.55	1.82
Student 2	135	0.54	1.85
Student 3	141	0.56	1.77
Student 4	134	0.54	1.87
Student 5	123	0.49	2.03
Student 6	127	0.51	1.97
Student 7	139	0.56	1.80
Student 8	147	0.59	1.70
Student 9	123	0.49	2.03
Student 10	139	0.56	1.80
Student 11	129	0.52	1.94
Mean	134.00	0.54	1.87
Standard deviation	7.7	0.03	0.10

Table 10. Lexical Frequency Profile of Ukrainian EFL students' course papers in terms of the number of tokens and types

From the analysis results presented in the table, it may be inferred that the number of types in the academic writings under scrutiny ranges from 123 to 147 with a mean of 134, suggesting that the textual material produced by the subjects has a moderate level of lexical diversity. Additionally, the fact that the standard deviation in this case is 7.35 enables us to assume that there was a general heterogeneity in terms

of vocabulary range demonstrated by the subjects in the academic writing under analysis.

As for the TTR indices, these vary from 0.49 to 0.59 with a mean of 0.54 and a standard deviation of 0.03. Since TTR indicates the ratio of unique words to the total number of words in the writing, there being the positive correlation between TTR and the variety of the vocabulary, the results obtained suggest that lexical diversity in the academic writing under scrutiny was moderate.

Furthermore, the range of TPT is from 1.70 to 2.03. the mean being 1.86 and the standard deviation 0.19, the latter indicating that the results were quite homogeneous among the participants. Additionally, since TPT demonstrates the frequency with which each type recurs within the textual material, i.e., the higher the TPT index, the more frequently each word is used in a text. Thus, the average TPT in academic writing under analysis being 1.86 demonstrates that on average each word is repeated fewer than two times. Thus, the students utilize a rather diverse vocabulary and do not excessively reiterate certain terms in their writing.

In addition, in the course of the analysis, a table displaying Lexical Frequency Profiles in terms of word frequency lists was compiled, in which:

- K1 refers to the first 1000 most frequent English words;
- K2 refers to the second 1000 most frequent English words;
- AWL refers to words from the Academic Word List [8];
- OLW refers to off-list words, i.e., words not included into any of the previous categories.

Student	K1	K2	AWL	OLW
Student 1	73.60%	4.00%	11.60%	10.80%
Student 2	76.00%	3.20%	8.40%	12.40%
Student 3	69.20%	4.00%	19.20%	7.60%
Student 4	69.60%	0.80%	11.60%	18.00%

Student 5	70.80%	8.00%	8.00%	13.20%
Student 6	72.40%	6.40%	10.00%	11.20%
Student 7	70.80%	3.60%	14.40%	11.20%
Student 8	70.00%	3.60%	5.20%	14.00%
Student 9	68.00%	6.40%	14.40%	11.20%
Student 10	74.40%	2.00%	8.80%	14.80%
Student 11	73.20%	4.00%	14.40%	8.40%
Mean	71.64%	4.18%	11.45%	12.07%
Standard deviation	2.46	2.06	3.94	2.92

Table 11. Lexical Frequency Profile of Ukrainian EFL students' course papers in terms of word frequency lists

As it is visible in the table provided above, the vocabulary in the participants' academic writings was analyzed in terms of its correlation with different word lists. In order to illustrate the results of this analysis, the following chart was created:

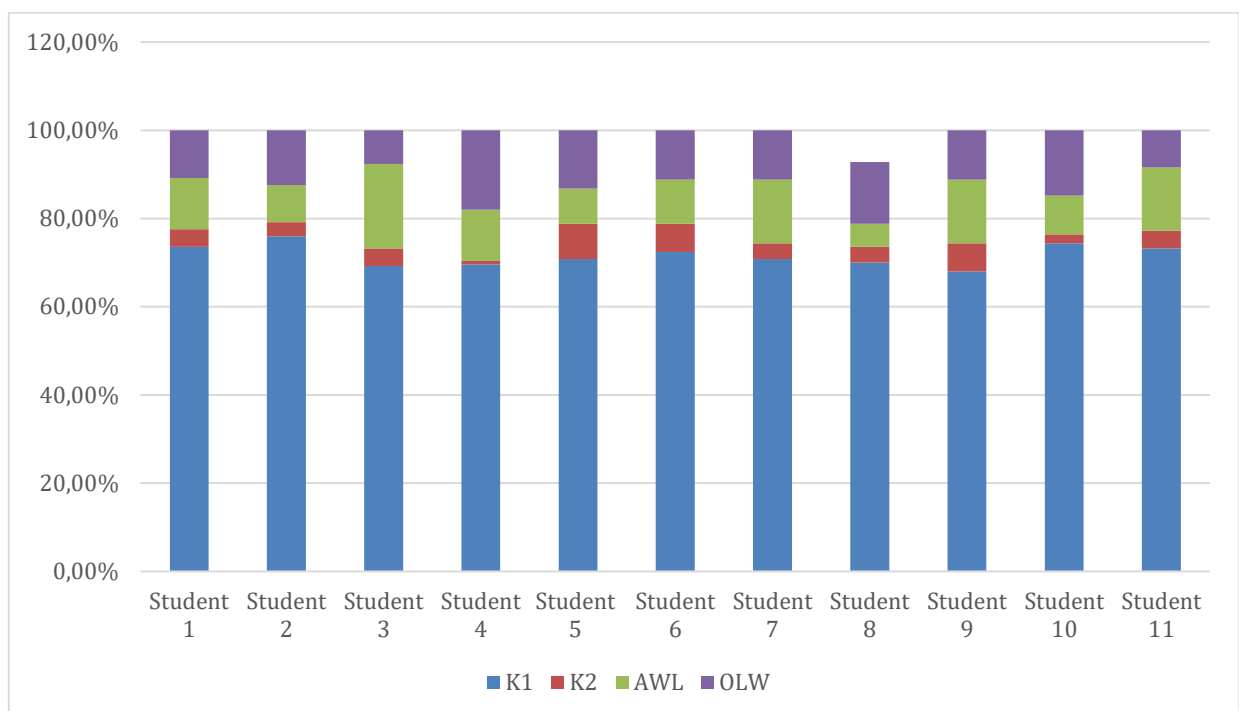


Chart 16. Lexical Frequency Profile of Ukrainian EFL students' course papers according to word frequency lists

From the data illustrated above it may be deduced that the vocabulary category which the participants used the most frequently is K1 – the first one thousand of the most frequent English words, which, on average, constitute 71.64% of the texts analyzed. Moreover, the standard deviation of 2.46 indicates that the results were fairly similar among the subjects. Therefore, it may be stated that all the students demonstrate a solid knowledge of the most frequent English words and can use them adeptly.

However, it may be observed that the scores for the second thousand of most frequent English words are the lowest among the group of participants, ranging from 0.8% to 8% with a mean of 4.18% and a standard deviation of 2.06. This suggests that the participants lack knowledge of lower-frequency words and operate predominantly with commonly used lexical units.

As for the Academic Word List vocabulary, its usage ranges from 8% to 19.20% among the participants, its average being 11.45% with a standard deviation of 3.94. Expectedly, this indicates that the students utilized a moderate number of academic words in their writing.

Additionally, the participants used from 7.6% to 18% of the words not belonging to the abovementioned categories. In general, the subjects' writings consisted on average of 12.07% of OLW, which demonstrates the students' knowledge of diverse vocabulary.

Overall, it may be noticed that the second-year undergraduate Ukrainian EFL students resorted to predominantly K1 words in their writings, and used K2 words least frequently, which demonstrates their insufficient familiarity with low-frequency vocabulary. Additionally, the participants' moderate usage of AWL vocabulary indicates that they adhered to the requirements of academic writing and endeavored to implement appropriate lexical units into their writing, which would correspond to its style. Furthermore, a relatively high frequency of OLW indicates the students' ability to operate with diverse vocabulary in academic writing.

3.2.1.2. Lexical diversity of Ukrainian students' Bachelor's papers

Following the procedure applied during lexical diversity analysis of Ukrainian EFL students' course papers, i.e. using *Textalyzer*, our investigation regarding lexical diversity in these students' Bachelor's papers yielded the results presented in the table below. It is important to mention that, as in the case of examining the students' course papers, 250-word samples from the studies' conclusions were selected for the analysis.

Student	Lexical diversity
Student 1	62%
Student 2	69%
Student 3	61%
Student 4	59%
Student 5	51%
Student 6	51%
Student 7	62%
Student 8	61%
Student 9	51%
Student 10	55%
Student 11	49%
Mean	57.36%
Standard deviation	6.36

Table 12. Lexical diversity of Ukrainian EFL students' Bachelor's papers

It is evident that the lexical diversity of the subjects' academic writing ranged from 49% to 69%, the mean being 57.36% and the standard deviation 6.36. These indices indicate that the participants' academic writing included a considerable number

of content words. In addition, the fact that the standard deviation equaled 6.36 demonstrates that there was a significant variation regarding lexical diversity across the group of participants. Additionally, the majority of students (six out of eleven) scored above the mean, indicating their higher proficiency in utilizing a wider vocabulary range in their writing. Furthermore, the chart represented below provides a visualization of the discussed data.

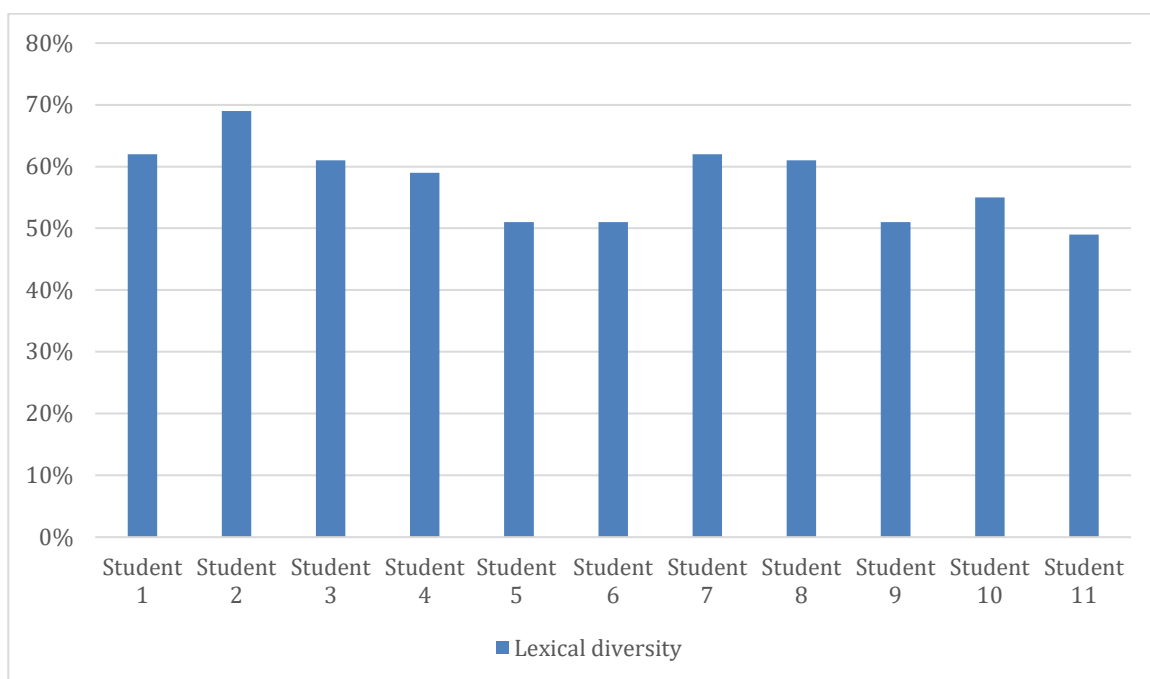


Chart 17. Lexical diversity of Ukrainian EFL students' course papers

When compared to lexical diversity indices obtained as a result of the analysis of the student's academic writing produced during the second year of study, certain interesting patterns may be brought to attention (see Appendix 3).

Thus, there is a general upward trend that may be observed with an average lexical diversity rising from 54.82% to 57.36% indicating an improvement in the student's vocabulary skills. Moreover, the standard deviation increasing from 5.44 to 6.36 demonstrates the growing heterogeneity of the group. Additionally, while during the second year of study only one student produced academic writing with lexical diversity of over 60% (Student 8), in the fourth year this number grew to two students (Student 2 and Student 8), also indicating the Student 8's sustained proclivity to write texts with higher lexical diversity. Furthermore, the trend to produce writing with

higher than the mean lexical diversity was maintained by the majority of students (4 out of 6).

Interestingly, the majority of the subjects increased their lexical variation indices between their second and fourth years of study, namely Student 1, Student 2, Student 3, Student 4, Student 7, and Student 9, while four students experienced a decrease in lexical diversity, to wit Student 5, Student 8, Student 10, Student 11, the result of Student 6 remaining constant.

Additionally, it is worth mentioning that significant standard deviation indices both in the second and the fourth years of study indicate the correlation of lexical diversity and such factors as personal writing style, individual language proficiency, and even personal interest in the topic of the writing, etc.

In order to further analyze the lexical diversity of the participants' academic writing produced during the fourth year of study, we resorted to the assistance of the VP Classic vocabulary profiler. It is this software that enabled us to develop Lexical Frequency Profiles of the students' writing. Consequently, the obtained data regarding the number of tokens and types and their ratios in the textual material under analysis is provided in the table below, where:

- TTR refers to the type-token ratio;
- TPT refers to the number of tokens per type.

It is of paramount importance to be mentioned that the number of tokens is equal for each student (250) due to the fact that it is 250-word samples from the conclusions of the subjects' Bachelor's papers that were selected for the analysis.

Student	Types	TTR	TPT
Student 1	149	0.60	1.68
Student 2	170	0.68	1.48
Student 3	148	0.59	1.69
Student 4	145	0.58	1.72

Student 5	124	0.50	2.02
Student 6	124	0.50	2.02
Student 7	150	0.60	1.67
Student 8	150	0.60	1.67
Student 9	125	0.50	2.00
Student 10	134	0.54	1.87
Student 11	123	0.49	2.03
Mean	140.18	0.56	1.80
Standard deviation	15.2	0.06	0.19

Table 13. Lexical Frequency Profile of Ukrainian EFL students' Bachelor's papers in terms of the number of tokens and types

Based on the information provided in the table, it may be assumed that the academic writings under analysis are characterized by a moderate level of lexical diversity, with the number of types varying from 123 to 170, the mean being 140.18. Moreover, the standard deviation equaling 15.2 demonstrates a considerable heterogeneity regarding the vocabulary range utilized by Ukrainian EFL students of Taras Shevchenko University of Kyiv in their academic writing.

After comparing these results with the ones obtained during the analysis of the academic writing produced in the second year of study, it can be noticed that the average number of types substantially increased over the period of two years from 134 to 140.18, which demonstrates that the students developed their ability to implement a more diverse vocabulary into their writing (detailed information is provided in Appendix 4).

Additionally, the standard deviation index doubled with the progression of time, it growing from 7.7 during the second year of study to 15.2 during the fourth. This indicates an increase in discrepancy among the group, regarding their use of different

terms in their textual material. Thus, it may be assumed that the homogeneity of the group decreased as they progressed with their studies and developed their individual writing styles.

As for the type-token ratio values obtained during the analysis of the fourth-year Ukrainian EFL students' academic writing, they ranged from 0.49 to 0.68, the mean being 0.56 and the standard deviation 0.06. This data provides a basis to assume that the writing produced by the subjects is characterized by an average lexical diversity.

Additionally, the number of tokens per type in the academic writing of the subjects varied from 1.48 to 2.03 with a mean of 1.80 and a standard deviation of 0.19. This signifies that, since, on average, a separate word was repeated fewer than two times in a text, the students demonstrated their ability to utilize a rather wide range of vocabulary in writing.

In comparison with the results of the analysis of academic writing produced by the subjects during the second year of study, an upward trend in terms of lexical diversity may be noticed. The growth in the number of types was accompanied by the increase in type-token ratio in the writings of 8 students out of 11, them being Student 1, Student 2, Student 3, Student 4, Student 5, Student 7, Student 8, and Student 9. In addition, only 3 students experienced a slight decrease in the number of types and TTR, namely, Student 6, Student 10, and Student 11.

Furthermore, the growth in the number of types and TTR over the period of two years was complemented in all cases by the decrease regarding the number of tokens per type. This tendency signifies the students' proclivity to reduce the single-word repetition frequency as they progress through their studies, gain linguistic expertise and develop language proficiency.

Another criterion in lexical diversity analysis was creating Lexical Frequency Profiles based on word frequency lists with the help of the VP Classic vocabulary profiler. The following table displays the data that was obtained in the course of the research, where:

- K1 refers to the first 1000 most frequent English words;
- K2 refers to the second 1000 most frequent English words;
- AWL refers to words from the Academic Word List [8];
- OLW refers to off-list words, i.e., words not included into any of the previous categories.

Student	K1	K2	AWL	OLW
Student 1	68.40%	4.00%	14.00%	13.60%
Student 2	66.53%	3.98%	14.74%	14.74%
Student 3	62.40%	10.00%	11.60%	16.00%
Student 4	66.40%	3.60%	14.80%	15.20%
Student 5	73.60%	7.20%	6.40%	12.80%
Student 6	75.20%	1.60%	11.20%	12.00%
Student 7	69.20%	6.00%	14.80%	10.00%
Student 8	74.40%	4.00%	15.20%	6.40%
Student 9	74.00%	3.60%	14.00%	8.40%
Student 10	66.00%	5.60%	9.20%	19.20%
Student 11	65.60%	3.60%	11.20%	19.60%
Mean	69.25%	4.83%	12.47%	13.45%
Standard deviation	4.36	2.27	2.82	4.14

Table 14. Lexical Frequency Profile of Ukrainian EFL students' Bachelor's papers in terms of word frequency lists

It may be observed that the overwhelming majority of words used by the subjects in their academic writing belong to the first thousand of the most frequent English words. The percentage of vocabulary belonging to this category ranges from 62.40%

to 64.40%, an average being 69.25% with a standard deviation of 4.36. As expected, it is the most frequently used vocabulary category, and the participants demonstrate an ability to confidently operate with the lexical units belonging to it. The visualization of the discussed statistics is provided in the chart below.

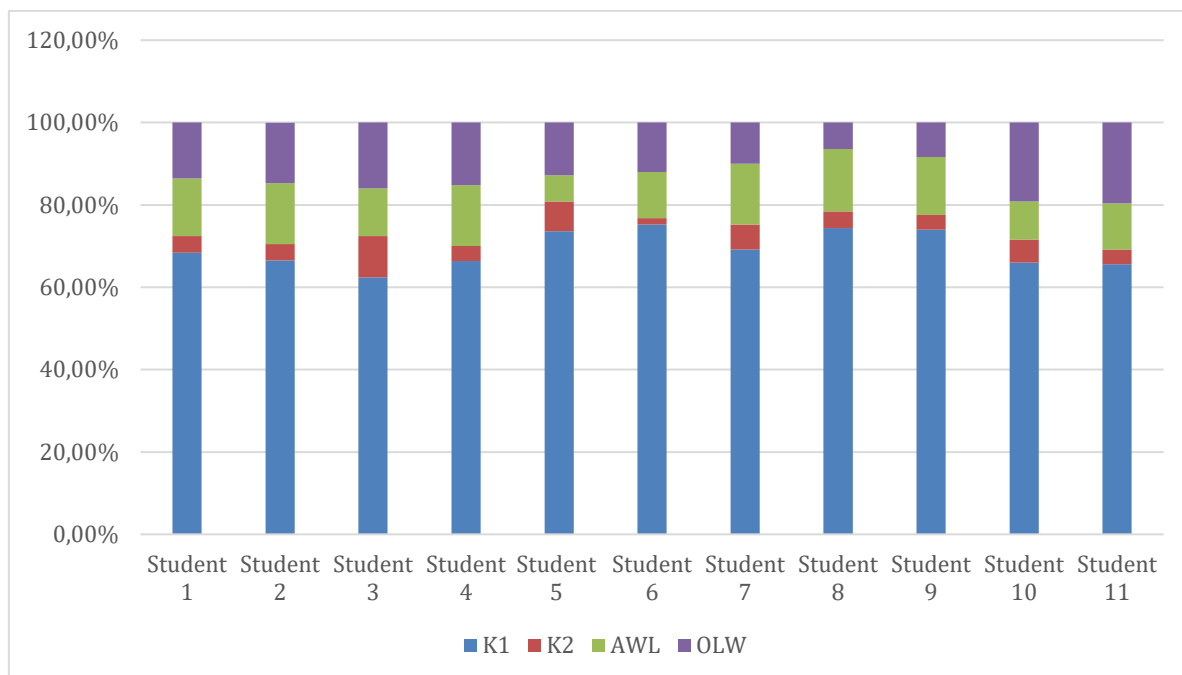


Chart 18. Lexical Frequency Profile of Ukrainian EFL students' Bachelor's papers according to word frequency lists

It should be mentioned that, as in the case of the academic writing produced by the subjects during the second year of study, K2 vocabulary is opted for the least frequently out of the four categories. The percentage of lexical units belonging to this list, which were used by the participants of the research in the fourth year of undergraduate studies varies from 1.60% to 10%, with a mean of 4.83% and a standard deviation of 2.27. As it was already mentioned in relation to the course paper excerpts analyzed, this statistic demonstrates that the students favor high-frequency vocabulary in their academic writing.

Furthermore, the subjects used a moderate number of lexical units belonging to the Academic Word List, specifically, its range amounted to 6.40% - 15.20%, the mean being 12.47%, while the standard deviation index equaled 2.82. As expected, in

general, the students adhered to the specifics of academic writing, which is reflected in the percentage of academic vocabulary they implemented in their writing.

Additionally, off-list words constituted from 6.40% to 19.60% of all the tokens, an average being 13.45% and a standard deviation of 4.14. Based on this data it may be assumed that the students demonstrated an ability to operate with a wide range of vocabulary.

Interestingly, when comparing the results discussed above with the ones obtained in the course of the analysis of samples extracted from the academic writing produced by the group of participants during the second year of study, it may be noticed that the average usage of K1 words decreased from 71.64% in the academic writing produced by the subjects during the second year of study to 69.25% during the fourth. This reduction was accompanied by the general growth in the usage of K2 words from 4.18% to 4.83% in the course of two years. Moreover, during this time period, both the students' implementation of Academic Word List vocabulary, as well as off-list vocabulary increased from 11.45% to 12.47% and from 12.07 to 13.45% respectively.

More specifically, seven students out of eleven reduced their usage of K1 words, to wit Student 1, Student 2, Student 3, Student 4, Student 7, Student 10, and Student 11. Furthermore, the number of K2 words increased in the writing of six students out of eleven, them being Student 2, Student 3, Student 4, Student 7, Student 8, and Student 10. Seven subjects out of eleven improved their usage of AWL vocabulary, namely Student 1, Student 2, Student 4, Student 6, Student 7, Student 8, and Student 10. In addition, six students experienced an increase in OLW implementation in their writing, these are Student 1, Student 2, Student 3, Student 6, Student 10, and Student 11. The complete data discussed above is provided in Appendix 5.

Therefore, the dynamic discussed in this section substantiates the assumption that the participants of the research developed their academic writing skills as well as language proficiency in the period from the second to fourth years of study. This improvement is reflected in the increase of lexical diversity in the majority of students' writings over time.

Moreover, borne in mind that the average lexical diversity of Ukrainian EFL students' academic writing, as estimated in this research, amounted to 54.82% in the second year of study and 57.36% in the fourth, it can be compared to the results obtained in the previous countries, the participants of which were the students residing in and originating from other countries. For this purpose, the chart which is represented below was created, where the following markings were used:

- A. Failasofah, Hazim Alkhrisheh, *Measuring Indonesian Students' Lexical Diversity and Lexical Sophistication*, Indonesia, undergraduate students;
- B. Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong, *The Effect of Topic on EFL Writing by Vietnamese Tertiary Students: Insights from Combining a Lexical Richness Analysis with Student Self-Reports*, Vietnam, age of the participants – 20 years old;
- C. Nata Permata, Rosyida Ekawati, *Lexical Richness of Short Stories Written by EFL Students*, Indonesia, fifth semester undergraduate students, age - 19-20 years old;
- D. Siti Aisah Ginting, *Lexical Complexity on Descriptive Writing of Indonesian Male and Female EFL Learners*, Indonesia, age of the participants – 18-22 years old;
- E. Khaled Layali, Ahmed Alshlowiy, *Online writing of Covid-19 in Bahrain: An analysis of lexical diversity and writing proficiency*, Bahrain, age of the participants – 19-21 years old.
- F. Lexical diversity of Ukrainian second-year undergraduate EFL students' creative writing, age of the participants – 17-18 years old;
- G. Lexical diversity of Ukrainian fourth-year undergraduate EFL students' creative writing, age of the participants – 20-21 years old.

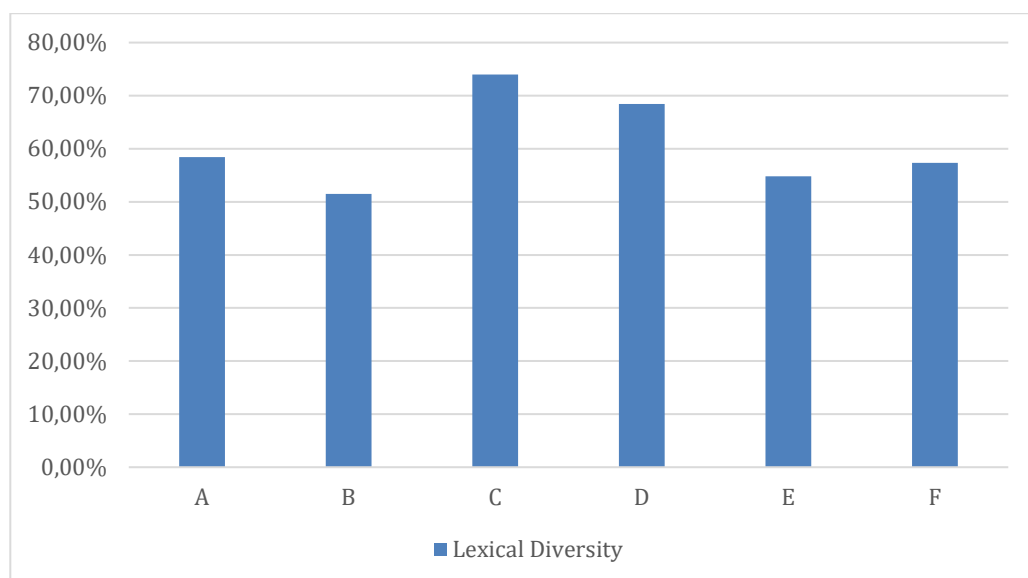


Chart 19. Lexical diversity of international and Ukrainian EFL students' academic writing

It may be inferred from the chart that the results of Ukrainian EFL students surpassed only those demonstrated by the Vietnamese students, who participated in the research conducted by Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong. This may be due to the fact that Vietnamese students were the closest in age to the Ukrainian EFL students, other research participants being generally older. Nonetheless, the lexical diversity of the Ukrainian students falls within the range of lexical diversities between 51.5% and 73.96% which were measured in the studies of lexical variance of academic writing produced by their international peers. Taken this information into consideration, it may be stated that in terms of lexical diversity, academic writing produced by Ukrainian EFL students is not inferior to that of their colleagues from other countries. However, there are still areas for improvement, since the average lexical diversity in the studies discussed in Chapter 1 of this research amounts to 63.59%.

3.2.2. Lexical diversity Ukrainian EFL students' creative writing

The analysis procedure of lexical diversity in Ukrainian undergraduate EFL students' creative writing was conducted in the same manner as that of lexical diversity in their academic writing. In particular, a total of 22 essays, 11 written during the second year of study and 11 during the fourth, were collected from the group consisting

of 11 Ukrainian undergraduate EFL students of Taras Shevchenko University of Kyiv, whose study program was *English Philology and Translation, Two Western European Languages*. Additionally, the subjects demonstrated a level of English proficiency, equivalent to the C1 level on the Common European Framework of Reference for Languages (CEFR), both during the second and the fourth year of study, them being 18-19 and 20-21 years old respectively.

Subsequently, since lexical diversity is sample size dependent, 250-word samples were extracted from the students' essays. Further, with the help of *Textalyzer* calculations of lexical diversity values in the texts analyzed were performed. Additionally, the utilization of the VP Classic vocabulary profiler enabled us to receive important information concerning the percentage of vocabulary belonging to Academic Word List, the list of the first thousand most frequent English words, the list of the second thousand of the most frequent English words, as well as the data regarding the number of tokens, types, type-token ratio and the number of tokens per type in the creative writing produced by Ukrainian undergraduate EFL students.

3.2.2.1. Lexical diversity of second year undergraduate Ukrainian students' creative writing

As previously stated, *Textalyzer* was used to assess the lexical variety of the creative writing produced by second-year Ukrainian undergraduate EFL students. For the analysis, 250-word extracts were taken from essays on the topic of either freedom as a concept and its meaning for the participants of the research or from the assignment which involved writing a review on a service (written by Student 3 and Student 6). The lexical diversity percentages determined using the aforementioned method are therefore shown in the table below.

Student	Lexical diversity
Student 1	66%
Student 2	66%

Student 3	57%
Student 4	63%
Student 5	68%
Student 6	68%
Student 7	63%
Student 8	62%
Student 9	54%
Student 10	59%
Student 11	65%
Mean	62.82%
Standard deviation	4.53

Table 15. Lexical diversity of Ukrainian second-year EFL students' creative writing

It is visible on the table that lexical diversity in the creative writing of the subjects varies from 54% to 68%, the mean being 62.82%, based on which it may be assumed that in general students utilized a relatively high proportion of unique words, their creative writing being characterized by a higher than average level of lexical diversity. Additionally, the standard deviation value of 4.53 suggests that the obtained results were relatively tightly clustered around the mean, which shows that there is a relatively even distribution of scores among the students rather than a significant variance in lexical variety.

Additionally, in order to visualize the information discussed above the following chart was created.

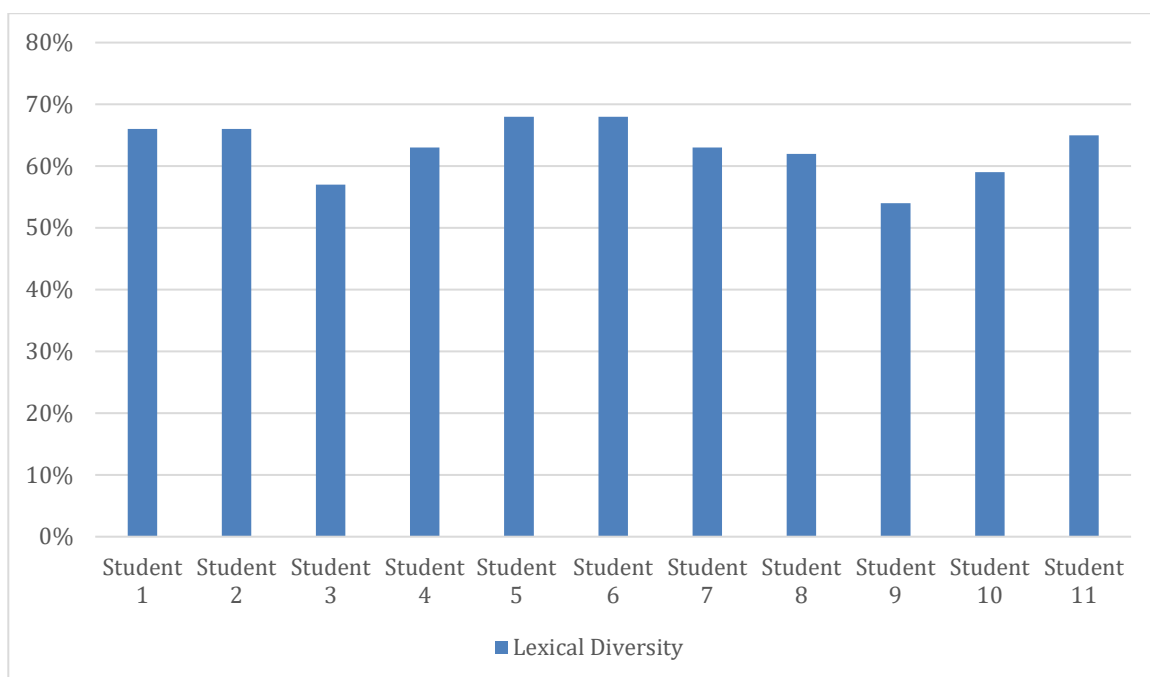


Chart 20. Lexical diversity of Ukrainian second-year EFL students' creative writing

Moreover, using the VP-Classic vocabulary profiler, we were able to create Lexical Frequency Profiles of each student's writing. Therefore, bearing in mind that the number of tokens in each textual material is 250, the table that follows shows the number of tokens and types and their ratios. In this table, the following markings are used:

- TTR refers to the type-token ratio;
- TPT refers to the number of tokens per type.

Student	Types	TTR	TPT
Student 1	155	0.62	1.61
Student 2	162	0.65	1.54
Student 3	138	0.55	1.81
Student 4	151	0.60	1.66
Student 5	161	0.64	1.55

Student 6	162	0.65	1.54
Student 7	154	0.62	1.62
Student 8	145	0.58	1.72
Student 9	132	0.53	1.89
Student 10	145	0.58	1.72
Student 11	153	0.61	1.63
Mean	150.73	0.60	1.66
Standard deviation	9.86	0.04	0.11

Table 16. Lexical Frequency Profile of Ukrainian second-year EFL students' creative writing in terms of the number of tokens and types

According to the table provided above, the number of types in the textual material under analysis varies from 132 to 162, an average being 150.73. This suggests a moderate level of lexical diversity demonstrated by the subjects, which is reflected in the number of unique words used by them. Additionally, the standard deviation value equaling 9.86 indicates a substantial level of variability among the group, which could be connected to the fact that in creative writing the students had more freedom in the expression of their ruminations and ideas.

As for TTR, which demonstrates a direct correlation to lexical diversity, its values range from 0.53 to 0.65, 0.60 being the average. This implies that the participants produced texts with a moderate level of lexical variance. Additionally, the standard deviation of 0.04 reveals a relatively low variability in the ratios across the group.

Furthermore, each separate word was used in the texts under analysis less than once, which is inferred from TPT ranging from 1.54 to 1.89, an average being 1.66. With a standard deviation of 0.11, the results of the participants have little variation in terms of this criterion. According to this data it may be stated that the

students implemented a relatively wide range of vocabulary into their creative writing, not excessively repeating lexical units.

In addition, the following table demonstrates the data concerning the percentage of words in the texts under scrutiny belonging to various word frequency lists, them being:

- K1 – the first 1000 most frequent English words;
- K2 – the second 1000 most frequent English words;
- AWL – Academic Word List [8];
- OLW – off-list words, i.e., words not included into any of the previous categories.

Student	K1	K2	AWL	OLW
Student 1	87.60%	4.80%	2.80%	4.80%
Student 2	80.00%	4.80%	10.00%	5.20%
Student 3	74.40%	14.00%	1.60%	10.00%
Student 4	86.40%	5.60%	4.80%	3.20%
Student 5	84.00%	3.60%	7.60%	4.80%
Student 6	82.00%	7.60%	2.00%	8.40%
Student 7	84.00%	3.20%	8.40%	4.40%
Student 8	87.20%	2.80%	5.20%	4.80%
Student 9	92.00%	2.40%	2.40%	3.20%
Student 10	86.00%	4.80%	5.60%	3.60%
Student 11	86.80%	2.40%	4.80%	6.00%
Mean	84.58%	5.09%	5.02%	5.31%
Standard deviation	4.62	3.34	2.76	2.13

Table 17. Lexical Frequency Profile of Ukrainian second-year EFL students' creative writing in terms of word frequency lists

In addition, the illustration of the information organized in the table above is provided by the following chart:

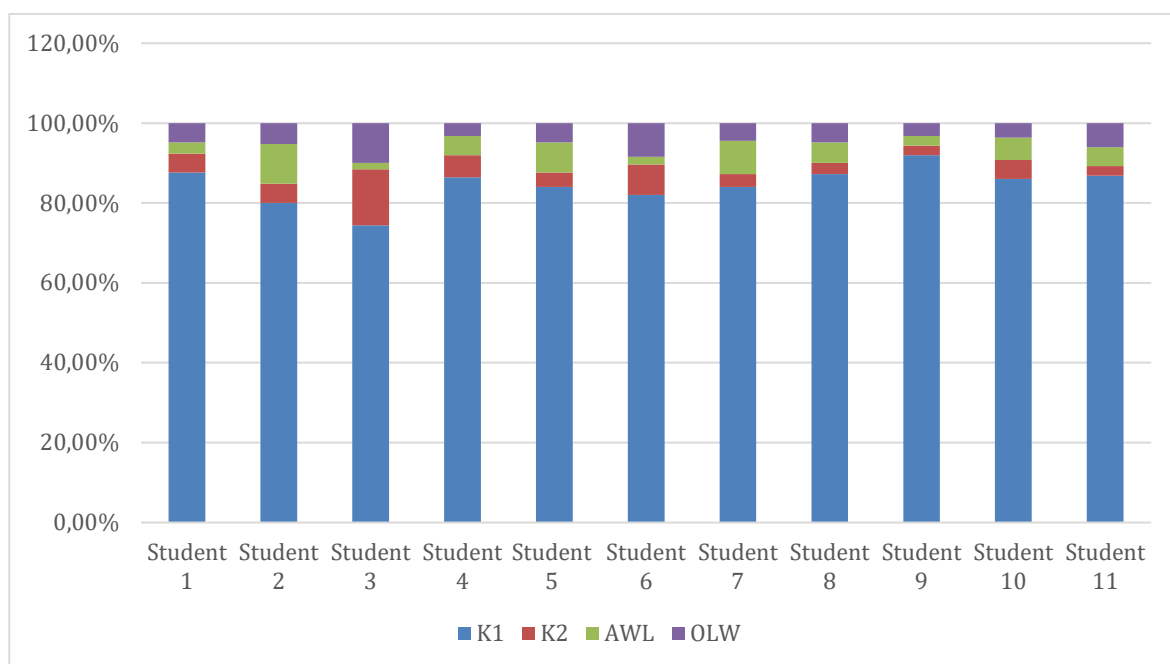


Chart 21. Lexical Frequency Profile of Ukrainian second-year EFL students' creative writing in terms of word frequency lists

Having observed the values obtained in the course of the analysis, it may be noticed that the most frequently used category of vocabulary belongs to the first one thousand of the most frequent English words. Their percentage in the subjects' creative writing varies from 74.40% to 92% with a mean of 84.58%, which demonstrates the students' strong preference for commonly used vocabulary. This may be linked to their wish to emulate informal speech in creative writing or insufficient knowledge of lower-frequency words. Additionally, the standard deviation of 4.62 indicates a relative heterogeneity among the group.

As for K2 words, their proportion in the texts ranges from 2.40% to 14.00% with a mean of 5.09% and a standard deviation of 3.34. This data substantiates the assumption that, in general, the participants of the research lacked familiarity with less common lexical units.

The least frequently implemented category of vocabulary in the texts produced by the participants was words belonging to Academic Word List, whose usage ranged from 1.60% to 10% with an average of 5.02% and standard deviation of 2.76. This result is not unexpected, since the task was to produce creative writing, which is not characterized by a high proportion of academic vocabulary.

In addition, off-list words occupied from 3.20% to 10% of the textual material under analysis, with a mean of 5.31% and a standard deviation of 2.13. Such values demonstrate the general lack of acquaintance with diverse vocabulary among the group, with only three students scoring higher than the mean.

Thus, it may be stated that the overwhelming majority of vocabulary in the second-year undergraduate Ukrainian EFL students' academic writing belonged to the K1 category. The subjects, however, demonstrated quite a limited knowledge of less frequent and academic words, which is reflected in their percentage in the texts produced by the participants of the research.

3.2.2.2. Lexical diversity of fourth year undergraduate Ukrainian students' creative writing

Having selected 250-word samples from the Ukrainian EFL students' creative writing produced in the fourth year of study, whose subject was Ukrainian childhood, we proceeded with its analysis with the help of *Textalyzer*. Thus, lexical diversity indices obtained with this tool are provided in the table below.

Student	Lexical diversity
Student 1	65%
Student 2	64%
Student 3	61%
Student 4	68%
Student 5	70%

Student 6	67%
Student 7	61%
Student 8	65%
Student 9	59%
Student 10	60%
Student 11	64%
Mean	64%
Standard deviation	3.49

Table 18. Lexical diversity of Ukrainian fourth-year EFL students' creative writing

From the table it may be inferred that lexical diversity in the writings analyzed varied from 59% to 70% with a mean of 64%, therefore, presumably, the students' essays had a rather plentiful lexical heterogeneity. In addition, the value of standard deviation equaling 3.49 implies that the subjects' lexical diversity scores did not deviate substantially from the average, indicating the general homogeneity of the group in relation to their writing skills and language proficiency. It is important to mention that the visualization of this information is provided in the chart below.

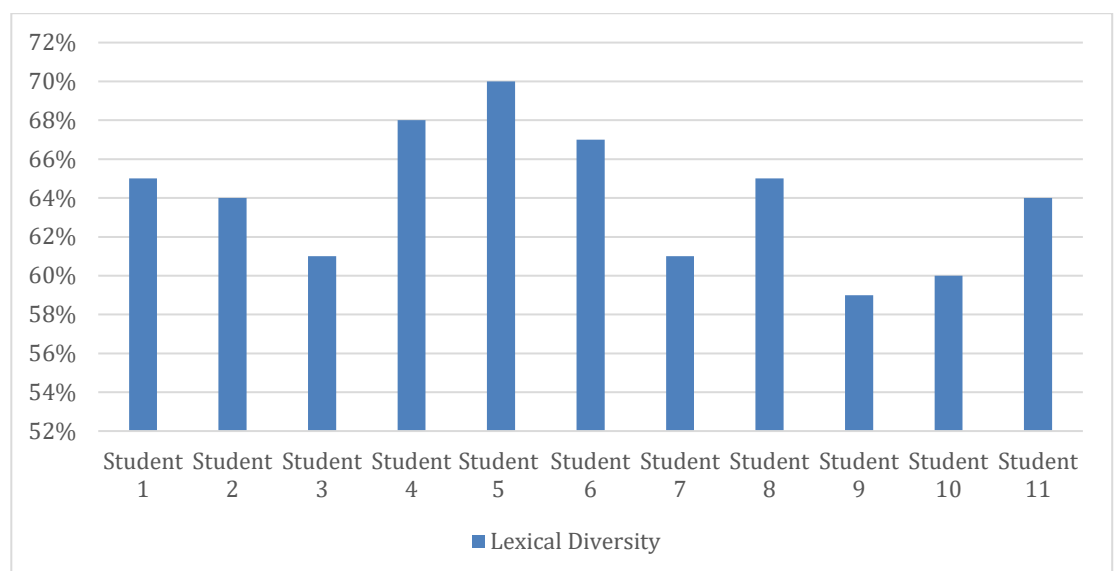


Chart 22. Lexical diversity of Ukrainian fourth-year EFL students' creative writing

If lexical diversity values measured in the Ukrainian EFL students' creative writing produced in the second and the fourth years of study are compared, a general rising tendency may be noticed with an average lexical variance index increasing from 62.82% to 64% which indicates the improvement in the students' vocabulary knowledge and their ability to implement a wider range of unique lexical units into their writing. Additionally, the fact that the standard deviation decreased in the course of two years from 4.53 to 3.49 reveals the trend for equalization within the group in terms of their linguistic proficiency, which may result from the students having the same writing assignments and receiving clear instructions as to the requirements to them. Moreover, the creative writing course that the subjects attended in the first semester of the fourth year of study, according to their curriculum, could have provided them with vocabulary knowledge and writing recommendations that may have led to a more consistent level of lexical diversity among the group.

It should be mentioned that lexical variation increased in six out of eleven students' creative writing, while five students experienced a slight decrease of a maximum of two units regarding this measure. To expand, the participants of the research whose lexical diversity grew in the course of two years are Student 3, Student 4, Student 5, Student 8, Student 9, and Student 10, while those whose lexical variance shrank are Student 1, Student 2, Student 3, Student 6, Student 7, Student 11. For the complete data discussed above refer to Appendix 6.

Further, with the assistance of the VP Classic vocabulary profiler, we have created Lexical Frequency Profiles of the subjects' creative writing produced during their fourth year of study. In relation to that, it is important to mention that due to the fact that it is 250-word samples of the students' essays whose lexical diversity was measured, the number of tokens in each writing is 250. The following table displays the results obtained in the course of the analysis regarding the number of tokens, types, and their ratios, where:

- TTR refers to the type-token ratio;
- TPT refers to the number of tokens per type.

Student	Types	TTR	TPT
Student 1	156	0.62	1.60
Student 2	160	0.64	1.56
Student 3	152	0.61	1.64
Student 4	165	0.66	1.52
Student 5	172	0.69	1.45
Student 6	162	0.64	1.56
Student 7	152	0.61	1.64
Student 8	155	0.62	1.61
Student 9	146	0.58	1.71
Student 10	145	0.58	1.72
Student 11	159	0.64	1.57
Mean	156.7	0.63	1.60
Standard deviation	8.04	0.03	0.08

Table 19. Lexical Frequency Profile of Ukrainian second-year EFL students' creative writing in terms of the number of tokens and types

As is visible in the table, the number of types in the writings under scrutiny ranged from 145 to 172 with a mean of 156.7, which suggests that on average the students implemented a moderate number of unique words into their essays. Additionally, a standard deviation equalling 8.04 implies a certain dispersion of the data around the mean, thus variability among the group in terms of the number of unique words used by the participants of the research.

An upward trend may be noticed when comparing the number of tokens in the creative writing produced by the students during the second and fourth years of study, which grew from 150.73 to 156.7. Based on this information it may be assumed that

the students improved their language proficiency over time and developed an ability to express their ideas with a wider vocabulary range in their creative writing. As for a slight decrease in standard deviation for this category from 9.86 to 8.04, it may suggest the movement towards consistency in using unique words within the group.

In addition, the TTR in the textual material analyzed ranged from 0.58 to 0.69 with an average amounting to 0.63 and a standard deviation of 0.03. This suggests that with the results tightly clustered around the mean, the students produced writings with a moderate level of lexical variance.

As for the number of tokens per type in the texts analyzed, it varied from 1.45 to 1.72. The mean in this category was 1.60, while the standard deviation equaled 0.08, based on which it may be stated that on average, the students did not tend to excessively use the same words throughout their writings.

Interestingly, when comparing the results obtained during the analysis of the creative writing produced in the second and the fourth years of study, a positive correlation between the increase in the growth in the number of types and TTR for six students out of seven, to wit, the increase in both measures was experienced by Student 3, Student 4, Student 5, Student 8, Student 9, Student 11, while with an increase in the number of types in the writing of the Student 1, the TTR value remained unchanged. Furthermore, only in the essays of two students, i.e. Student 1 and Student 7, there was a slight decrease in the number of types and TTR over time. Interestingly, Student 10 produced writing with the same number of types, TTR and TPT both in the second and the fourth years of study.

Moreover, eight students out of eleven experienced a decrease in the number of tokens per type over time (Student 1, Student 2, Student 3, Student 4, Student 5, Student 8, Student 9, Student 11). Based on this downward trend it may be assumed that there is a negative correlation between the students' progression in their studies and a separate word repetition frequency in their creative writing. It is also important to mention that an average TTR value increased from 0.60 to 0.63, while the mean for TPT shrank from 1.66 to 1.60, which corroborates the assumption of the students

developing their vocabulary knowledge and lexical diversity as demonstrated in their creative writing over the period of two years. The full statistic concerning these measures is provided in Appendix 7.

Additionally, the VP Classic vocabulary profiler enabled us to analyze the fourth year Ukrainian EFL students' creative writings in terms of the percentage of words in them belonging to different word frequency lists, which have the following markings in the table provided further in this section:

- K1 – the first 1000 most frequent English words;
- K2 – the second 1000 most frequent English words;
- AWL – Academic Word List [8];
- OLW – off-list words, i.e., words not included into any of the previous categories.

Student	K1	K2	AWL	OLW
Student 1	86.80%	6.00%	0.40%	6.80%
Student 2	82.40%	5.20%	6.00%	6.40%
Student 3	84.80%	4.80%	4.80%	5.60%
Student 4	80.80%	6.80%	1.60%	10.80%
Student 5	74.00%	8.80%	2.80%	14.40%
Student 6	84.19%	8.70%	0.40%	6.72%
Student 7	86.00%	4.00%	0.80%	9.20%
Student 8	86.00%	4.00%	1.20%	8.80%
Student 9	82.80%	9.60%	1.20%	6.40%
Student 10	86.00%	4.40%	2.40%	7.20%
Student 11	77.60%	8.80%	2.00%	11.60%
Mean	82.85%	6.46%	2.15%	8.54%

Standard deviation	4.02	2.16	1.80	2.74
---------------------------	-------------	-------------	-------------	-------------

Table 20. Lexical Frequency Profile of Ukrainian fourth-year EFL students' creative writing in terms of word frequency lists

Moreover, the visualization of the data presented in this table is provided in the chart below.

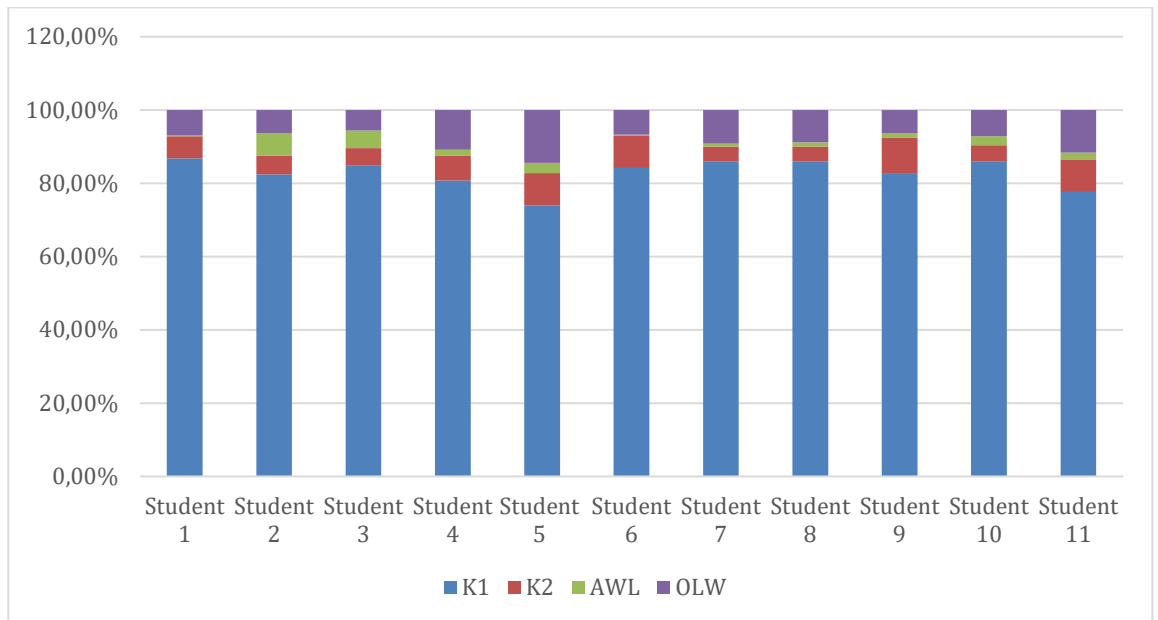


Chart 23. Lexical Frequency Profile of Ukrainian fourth-year EFL students' creative writing in terms of word frequency lists

Unsurprisingly, the most frequently used vocabulary category in the participants' creative writing belongs to the list of the first thousand most frequent English words, the percentage of which ranges from 74% to 86.80 with a mean of 82.85% and a standard deviation of 4.02. This suggests that the students acquired a substantial knowledge of the vocabulary encountered in this list and are able to efficiently implement it in their creative writing.

The second most frequently used vocabulary category in the writings under analysis is off-list words. Their values range from 5.60% to 14.40%, the mean amounting to 8.54%, and the standard deviation 2.74. Based on this data it may be assumed that the participants endeavor to utilize more sophisticated and less common lexical units in the course of the writing process.

This category is followed by the words belonging to the second thousand of the most frequent English words which occupy from 4% to 9.60% of the texts under analysis, the mean being 6.46% and standard deviation 2.16. With this in mind, it may be assumed that the participants lack knowledge of more unique low-frequency words.

The least frequently used words in the texts produced by the participants of the research are lexical units belonging to the Academic Word List, which occupy only from 0.40% to 6% of the written material, with a mean of 2.15% and a standard deviation of 1.80. This result is not unexpected, since it is creative writing that was produced by the participants, which does not require students to adhere to the requirements of more formal academic writing.

Interestingly, in the period between the second and the fourth years of study, the percentage of K1 words used by the students dropped from 84.58% to 82.52% on average, with six out of eleven students being subject to this change (Student 1, Student 4, Student 5, Student 8, Student 9, Student 11). This decrease was accompanied by an increase in the usage of K2 words from 5.09% to 6.46% on average. This growth is evident in the writings of nine students out of eleven, while the writings produced by only two students (Student 9 and Student 11) demonstrated a reduction in this category.

Additionally, almost all the participants (ten out of eleven) reduced the percentage of AWL vocabulary in their creative writing over time, which dropped from 5.02% in the second year of study to 2.15% in the fourth. This dynamic may signify the participants becoming more acquainted with the features of different styles of writing and their successful attempts to adhere to the requirements of creative writing.

Also peculiar is the fact that the amount of off-list words grew in the students' creative writing in the period of two years from 5.31% to 8.54% on average. Such progress is evident in the texts produced by all the students except for Student 3 and Student 6, which suggests that in the period of time between the second and the fourth years of study the participants expanded their vocabulary ranges and improved their language competences (see Appendix 8).

Taking into consideration the statistics discussed above, it may be stated that overall the Ukrainian EFL students, who are the participants of this research, increased their lexical diversity over time. This is substantiated by an upward trend in terms of OLW and K2 percentage, number of types, and TTR as well as by a decrease in the TPT, in the utilization of both K1 and AWL words in their creative writing.

Taken into consideration that the average lexical diversity in Ukrainian EFL students' creative writing equaled 62.82% during the second year of study and 64% during the fourth, these results can be compared to the ones obtained in the previous studies focused on the lexical variance in the creative writing of other EFL students originating from other countries. Therefore, in order to illustrate the positioning of Ukrainian EFL students among their international peers regarding the lexical diversity of their creative writing, the chart with the following markings was created:

- A. Failasofah, Hazim Alkhrisheh, *Measuring Indonesian Students' Lexical Diversity and Lexical Sophistication*, Indonesia;
- B. Bao Trang Thi Nguyen, Jonathan Newton, and Tran Ngoc Quynh Phuong, *The Effect of Topic on EFL Writing by Vietnamese Tertiary Students: Insights from Combining a Lexical Richness Analysis with Student Self-Reports*, Vietnam;
- C. Nata Permata, Rosyida Ekawati, *Lexical Richness of Short Stories Written by EFL Students*, Indonesia;
- D. Siti Aisah Ginting, *Lexical Complexity on Descriptive Writing of Indonesian Male and Female EFL Learners*, Indonesia;
- E. Khaled Layali, Ahmed Alshlowiy, *Online writing of Covid-19 in Bahrain: An analysis of lexical diversity and writing proficiency*, Bahrain.
- F. Lexical diversity of Ukrainian second-year undergraduate EFL students' creative writing;
- G. Lexical diversity of Ukrainian fourth-year undergraduate EFL students' creative writing.

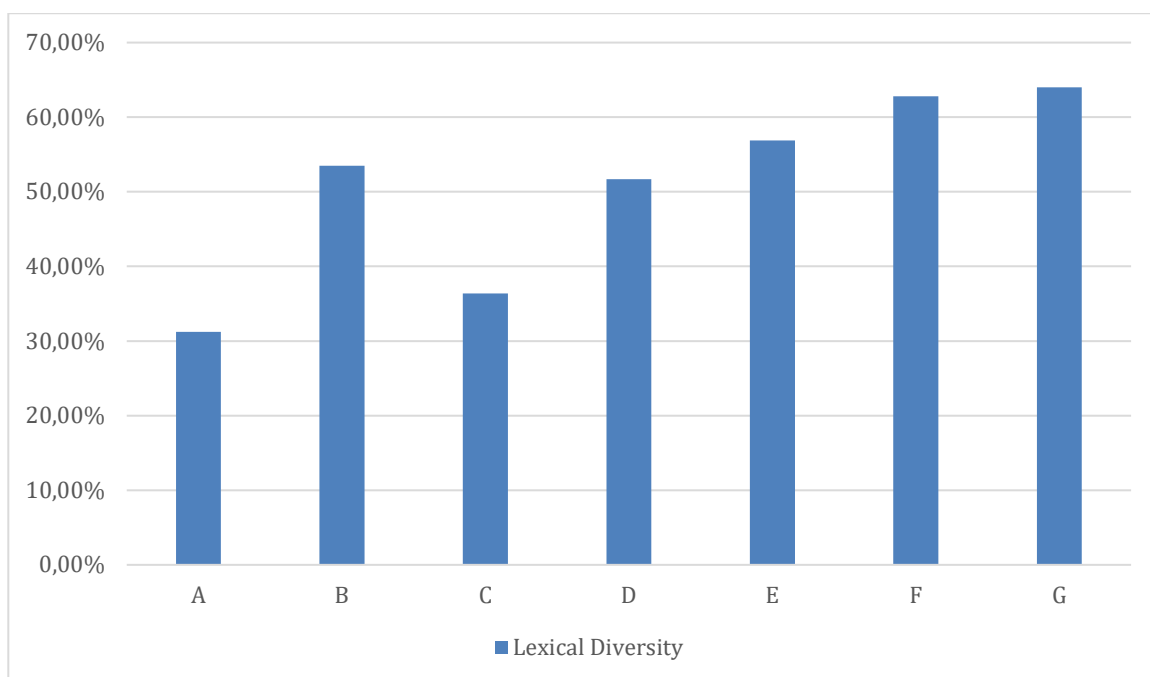


Chart 24. Lexical diversity of international and Ukrainian EFL students' creative writing

It may be inferred from the chart that Ukrainian EFL students demonstrated the highest lexical diversity scores both in the second and the fourth years of study, with the results of their peers in the global EFL community ranging from 31.23% to 56.9%, and an average of 45.143%. Thus, it may be assumed that the freedom of expression and lower formality than that required in academic writing, might have enabled the participants of this research to operate with a wider range of vocabulary and use it in a creative way. Additionally, the expressive nature of creative writing might have encouraged the students to experiment with language and literary devices. Moreover, the fact that international peers and Ukrainian EFL students were close in age minimizes the impact of age-related differences on writing proficiency and provides a fairer comparison picture. The lexical diversity values observed in the Ukrainian EFL students' creative writing indicate their strong command of the English language, which provides them with an advantage when compared to their international peers.

Conclusion

According to the data obtained in the course of lexical density analysis, the average sentence length in the course papers produced by Ukrainian EFL students is

15 words, each being comprised of 1.7 syllables, while in the Bachelor's papers, it amounted to 15.4 words consisting of 1.63 syllables. In relation to this, it may be stated that the participants of the research gave preference to sentences of average length as well as mono- and disyllabic words in their academic writing.

Additionally, an average lexical density index equaled 49% in the academic writing produced by Ukrainian EFL students both in the second and fourth years of study, which is considered to be a moderate level of lexical density. The absence of a significant development in terms of lexical density of academic writing produced by the participants of the research may be connected to the fact that the participants already had C1 level of English proficiency when entering the university, therefore, they did not experience a transition from one level to another, which may have led to a more profound change in the lexical density of their academic writing. Additionally, a possible reason for that may be the absence of a separate academic writing course in the participants' curriculum as well as rigid requirements of academic writing which did not allow the students to experiment either with form or with the style of their writing.

When compared to the lexical density of the academic writing produced by international EFL students, given that the average lexical density in the studies analyzing academic writings of EFL students from different countries is 46.71%, one might claim that in this respect the scores obtained by the subjects of this research (49%) cannot be considered inferior. The disparity in the lexical density of academic writing produced by Ukrainian and international EFL students may be connected to the higher levels of study (graduate) of two out of four groups of participants from other countries. Moreover, it may be assumed that age and maturity influence the breadth of worldview and thus the sophistication of a person's ideas, which in turn is reflected in the complexity of vocabulary. Since lexical density signifies the proportion of content words to a total number of words, the lower scores of Ukrainian EFL students may signify that they use more function words and thus are less precise in the expression of their ideas. However, the general lexical density of 49% provides a basis to assume

that the participants of this research maintained a balance between lexical and grammatical words thus presumably reducing the convolutedness of their writing and increasing its reading ease.

As for creative writing produced by the Ukrainian students during the second year of study, it contained on average 16-word sentences, the words which consisted of 1.5 syllables. In the fourth year of study sentence length grew considerably to 20.36 words per sentence with an average word length dropping to 1.4 syllables. Based on this information it may be stated that during the second year of study, the participants used sentences of average length, while in the fourth they exceeded this hallmark and utilized sentences slightly longer than average with primarily monosyllabic words.

Moreover, the lexical density of second-year students' creative writing amounted to 38%, rising to 39.45% in the fourth year of study. This demonstrates that the students started implementing more diverse and sophisticated vocabulary in their writing. The comparison with lexical density results of the students from other countries demonstrated that the scores of Ukrainian students were not significantly lower, with those obtained by the students from other countries ranging from 36.55% to 51.79%. Therefore, it may be stated that creative writings produced by Ukrainian EFL students are on par with the ones written by their international peers.

Additionally, lexical diversity measurement of the participants' course papers yielded an average result of 54.82%. Given that the average number of types in these writings was 134, the type-token ratio was 0.54 and the number of tokens per type was 1.87, the lexical variance in them may be characterized as moderate. Moreover, the students resorted to K1 words in their writing the most frequently (71.64%), them being followed by OLW (12.07%), AWL (11.45%), and K2 (4.18%).

In the fourth year of study, lexical diversity of academic writing grew to 57.36%, an average number of types to 140.18, with TTR rising to 0.56, and TPR dropping to 1.80. This data provided a foundation to assume that the participants' Bachelor's papers have an average lexical diversity. As for the vocabulary categories used by Ukrainian EFL students, the percentage of K1 words in the texts under analysis decreased to

69.25%, while OLW increased to 13.45%, AWL to 12.47% and K2 to 4.83%. This dynamic demonstrates the development of the students' vocabulary range and their ability to implement it in their academic writing.

It should be mentioned that in terms of lexical diversity, the academic writing of Ukrainian students is compatible with that of their peers, whose lexical diversity ranges from 51.5% and 73.96%. However, an average of their results being 63.59%, there exists potential for enhancement in this area.

Measured in creative writing produced by the participants of the research during the second year of study, lexical diversity amounts to 62.82%. The students demonstrated a moderate level of lexical diversity with an average number of types of 150.73, TTR 0.60, and TPT 1.66. The majority of vocabulary encountered in the writings analyzed belongs to the K1 category (84.58%), which is followed by OLW (5.31%), K2 (5.09%), and AWL (5.02%).

In the fourth year of study, the lexical diversity of Ukrainian EFL students' creative writing rose to 64%, while the number of types increased to 156.7, TTR to 0.63, and TPT decreased to 1.60. Interestingly, the amount of K1 vocabulary shrank to 82.85% and AWL to 2.15%, while OLW advanced to 8.54% and K2 to 6.46%. Therefore, it may be inferred that in the period of two years, the students managed to improve their language competencies and expand their vocabulary with lower-frequency lexical units.

It should be mentioned that lexical diversity values in creative writing of Ukrainian students produced both in the second and the fourth years of study are higher than those of their peers in the global EFL community whose results had an average of 45.143%. Thus, one may assume that it is the expressive nature and lower formality level of creative writing that enabled Ukrainian students to successfully demonstrate their high level of proficiency in the English language.

Additionally, a positive correlation between lexical density and lexical diversity values is demonstrated both in creative and academic writing by the majority of students. Thus, this tendency may be observed in the academic writing of Student 2,

Student 3, Student 5, Student 7, Student 10, and Student 11 over the period of two years. Likewise, a positive correlation between these values in creative writing is evident in the textual material produced by Student 2, Student 4, Student 7, Student 8, Student 9, and Student 10. This indicates the existing interdependence between lexical density and lexical diversity.

Furthermore, it is evident that lexical density in all the students' creative writing is lower than that of academic, whereas lexical diversity in creative writing is generally higher than in academic, which is demonstrated by all the students except Student 8 in the second year of study as well as Student 3, Student 4 and Student 7 in the fourth. Therefore, it may be assumed that creative writing is more conducive to the application of more varied vocabulary by the students, while a higher lexical density of academic writing suggests a higher level of precision and thus concentration of words in a text as required by this type of writing. For the detailed data used for the comparison, see Appendices 9 and 10.

CONCLUSION

Lexical density and lexical diversity are regarded to be valid indicators widely applied to assess language learners' linguistic competencies and proficiency. Lexical density measures information load in a text and is calculated as a ratio of content words to the total number of words in a sentence or a text. Lexical diversity, being interdependent with the text genre as well as the students' prior writing experience, indicates the amount of different words in textual material.

As inferred from the various research papers concerning lexical density in EFL writing, its values equaled on average 46.71% in academic and 45.76% in creative writing produced by international EFL students. As for lexical diversity, it amounted to on average 45.143% in creative writing, and to 63.59% in academic writing of the learners of English as a foreign language in different countries.

The data source for our study was a corpus of academic and creative writing produced in the second and the fourth years of study by 11 participants of this research, undergraduate EFL students of Taras Shevchenko National University of Kyiv, Ukraine. In this study, we applied quantitative methods of research. Lexical density and lexical diversity indices as well as an average number of words per sentence, and syllables per word in the textual material were calculated with the help of *Textalyzer*. Additionally, frequency analysis was conducted with the assistance of an online vocabulary profiler VP Classic in order to create Lexical Frequency Profiles, which distributed words from the texts under analysis into different frequency layers and provided data concerning the number of tokens and types in the texts and their ratios.

In the course of lexical density analysis, it was determined that the Ukrainian EFL students' academic writing produced both in the second and the fourth years of study generally consisted of monosyllabic and disyllabic words organized into sentences of average length. Additionally, academic writing produced by Ukrainian EFL students in the second and in fourth years of study is characterized by a moderate level of lexical density, which equals 49%. The high level of English proficiency possessed by the research participants already when entering university as well as the

absence of a specific academic writing course and latitude in writing requirements may have prevented major improvements in the lexical density of their academic writing.

In addition, the lexical density measured in the academic writing of Ukrainian EFL students may be considered comparable to the average lexical density of academic writing produced by international EFL students (46.71%). Furthermore, the observed disparities in the results of Ukrainian undergraduate and international graduate EFL students provided us with the basis to assume that since age and maturity directly correlate with a person's ability to express complex ideas, these factors have a profound influence on vocabulary richness.

While creative writing produced by the participants of this study both in the second and the fourth years of study is characterized by the preference for monosyllabic words, in the fourth year of study general sentence length exceeded the hallmark of average. Moreover, in the fourth year of study, students started utilizing a more sophisticated vocabulary which is reflected in the general growth of lexical density. When compared to the creative writing produced by EFL students from different countries, it was observed that the creative writing of Ukrainian EFL students remained within the bounds of the lexical density results of their international peers.

According to the lexical diversity scores and Lexical Frequency Profiles created during the analysis of the Ukrainian student's course papers, lexical variation in those academic writings may be characterized as moderate with the dominance of vocabulary belonging to the first thousand of the most frequent English words (K1). In the fourth year of study, the students alongside their lexical diversity scores improved the sophistication of their vocabulary which is demonstrated by the decrease in the percentage of K1 words in the texts under analysis and the increase in all the other vocabulary categories. Additionally, after a comparison of the lexical diversity results in the academic writing of Ukrainian and international EFL students, it was determined that the lexical diversity of the former is not inferior to the latter.

As for the lexical diversity scores in the creative writing of the research participants, they grew in the period between the second and fourth years of study.

Additionally, the students demonstrated lexical diversity development, which may be inferred from their Lexical Frequency Profiles that demonstrated that the percentage of K1 and Academic Word List vocabulary shrank whereas the percentage of the second thousand of the most frequent English words and off-list words increased in the texts under analysis over time. Interestingly, Ukrainian EFL students demonstrated higher lexical diversity results in creative writing as compared to those of their international peers. Thus, the lower formality level and less rigid requirements of creative writing allowed the research participants to demonstrate their advanced English proficiency and vocabulary richness.

Additionally, in the course of the research an interrelation between lexical density and lexical diversity was discovered, with these measures displaying a positive correlation both in academic and creative writing. Moreover, it was determined that lexical density tends to be lower and lexical diversity higher in creative writing than in academic. This suggests that creative writing provides a more favorable environment for the application of a wider range of vocabulary, while academic writing demands a higher information load and greater precision in word choice.

Taking everything stated above into consideration, it may be concluded that this study managed to position Ukrainian students among the international community in terms of EFL operational vocabulary, namely lexical density and lexical diversity.

SUMMARY

In order to assess the quality of foreign language learners' writing and their linguistic competencies, various measurements have been conducted by educators and language instructors. Such concepts as lexical density and lexical diversity when applied in the process of EFL acquisition progress assessment can provide pivotal information concerning language learners' writing skills and highlight prospective areas for improvement which in turn will help facilitate academic growth. Therefore, the aim of this paper was to investigate the correlation between lexical density, lexical diversity, and language proficiency.

This study explored the concepts of lexical density and lexical diversity in academic and creative writing and determined the existing patterns regarding lexical density and lexical diversity of academic and creative writing of Ukrainian EFL students. In the course of this research Lexical Frequency Profiles of Ukrainian EFL students were created, which made it feasible to draw connections between the language proficiency of the subjects and the lexical density and lexical diversity of their writing. Additionally, the relationship between lexical density and other linguistic features such as sentence length and word length in creative and academic writing of Ukrainian EFL students was explored. Furthermore, in this thesis, the potential impact of such factors as text genre and language proficiency on lexical density and lexical diversity was examined and Ukrainian EFL students were positioned in the community of international EFL learners regarding their lexical richness and active vocabulary. To do so, we resorted to the quantitative method, which consisted of frequency analysis, as well as lexical density and lexical diversity measurement.

The paper's scientific originality stems from the evaluation of the language proficiency of Ukrainian EFL students conducted through the exhaustive and systematic study and analysis of lexical peculiarities demonstrated in their writing. The findings of this research may be applied in creative and academic writing seminars, as well as used by EFL educators and instructors.

The theoretical background to the study is described in the first chapter of the thesis, whereas the methodology and data analyzed are presented in the second. It is the third chapter that provides systematized and analyzed findings of the research.

It was determined that the participants of the research produced academic writing with moderate lexical density levels both in the second and the fourth years of study, their lexical density values not being inferior to the average lexical density index of their international peers.

In the creative writing of the Ukrainian EFL students, which was characterized as moderately lexically dense, there was detected a lexical density growth in the period between the second and the fourth years of study. Additionally, their scores lie within the bounds of the lexical density results of other students in the global EFL community.

The data concerning lexical diversity indices in the academic writing of Ukrainian EFL students indicate a moderate level of lexical variance with a growth in the fourth year of study. Additionally, according to the Lexical Frequency Profiles, in the period of two years, the participants managed to widen their vocabulary range and start implementing more unique lexical units into their academic writing. When compared to those of their international peers, the results of Ukrainian students cannot be considered inferior.

Similarly, the lexical diversity of the research participants' creative writing increased in the period between the second and fourth years of study. Their developed language proficiency and enriched vocabulary are reflected by an increase of lower frequency words as demonstrated by their Lexical Frequency Profiles. Moreover, the lexical diversity scores of Ukrainian students were higher than those of EFL learners from other countries, which underscores their profound English knowledge as well as the conduciveness of creative writing to the students' implementation of a wider vocabulary range.

Thus, our analysis placed Ukrainian students among the global EFL community based on their operational vocabulary as measured by lexical density and lexical diversity.

REFERENCES

1. Al-Wahy A. S. Towards a methodology for measuring lexical density in Arabic. *International journal of arabic linguistics*. 2017. Vol. 3, no. 1. P. 1–33.
2. An Investigation of the Linguistic Complexity of IELTS Writing Topics Based on the Levels of Discourse Representation and the Degree of Meaning Coding / M. Zare et al. *Cogent Education*. 2021. Vol. 8.
3. Biber D., Conrad S. Register, Genre and Style. Cambridge : Cambridge University Press, 2009. 344 p.
4. Celce-Murcia M., Olshtain E. Discourse and context in language teaching: A guide for language teachers. Cambridge : Cambridge University Press, 2000. 286 p.
5. Chen Y.-H., Baker P. Lexical bundles in L1 and L2 academic writing. *Language learning & technology*. 2010. Vol. 14, no. 2. P. 30–49.
6. Connor U. Coherence in Writing: Research and Pedagogical Perspectives / ed. by A. M. Johns. Teachers of English to, 1990. 263 p.
7. Coxhead A. Academic vocabulary, writing and english for academic purposes: Perspectives from second language learners. *RELC Journal*. 2012. Vol. 43, no. 1. P. 137–145.
8. Coxhead A. A New Academic Word List. *TESOL Quarterly*. 2000. Vol. 34, no. 2. P. 213–238.
9. Cutts M. Oxford Guide to Plain English. Oxford University Press, 2013. 288 p.
10. Daller H., van Hout R., Treffers-Daller J. Lexical Richness in the Spontaneous Speech of Bilinguals. *Applied Linguistics*. 2003. Vol. 24, no. 2. P. 197–222.
11. Discourse Complexity of College Writers With and Without Disabilities / N. Gregg et al. *Journal of Learning Disabilities*. 2002. Vol. 35, no. 1. P. 23–38.
12. Dong Y. R. Non-native Graduate Students' Thesis/Dissertation Writing in Science: Self-reports by Students and Their Advisors from Two U.S. Institutions. *English for Specific Purposes*. 1998. Vol. 17, no. 4. P. 369–390.

13. Durrant P., Doherty A. Teachers' perceptions of corpus use in the foreign language classroom. *Language Learning & Technology*. 2010. Vol. 14, no. 2. P. 80–99.
14. Dynamic Norming for Systematic Analysis of Language Transcripts / A. Tucci et al. *Journal of Speech, Language, and Hearing Research*. 2022. Vol. 65, no. 1. P. 320–333.
15. Eggins S. An introduction to systemic functional linguistics. 2nd ed. New York : Continuum, 2004. 384 p.
16. Ekawati R., Permata N. Lexical Richness of Short Stories Written by EFL Students. *EFL Education Journal*. 2022. Vol. 9, no. 1. P. 102–116.
17. Elgobshawi A. E., Aldawsari M. Lexical Density as Improvement Indicator in the Written Performance of EFL Majors. *International Journal of English Language and Literature Studies*. 2022. Vol. 11, no. 4. P. 181–190.
18. Failasofah F., Alkhrisheh H. Measuring Indonesian Students' Lexical Diversity and Lexical Sophistication. *Indonesian Research Journal in Education*. 2018. Vol. 2, no. 2. P. 97–107.
19. Francis W. N., Kucera H. Frequency analysis of English usage: Lexicon and grammar. Boston : Houghton Mifflin, 1982. 561 p.
20. Ginting S. A. Lexical Complexity on Descriptive Writing of Indonesian Male and Female EFL Learners. *International Journal of English Linguistics*. 2018. Vol. 8, no. 3. P. 297.
21. Gregori-Signes C., Clavel-Arroitia B. Analysing Lexical Density and Lexical Diversity in University Students' Written Discourse. *Procedia - Social and Behavioral Sciences*. 2015. Vol. 198. P. 546–556.
22. Ha H. S. Lexical Richness in EFL Undergraduate Students' Academic Writing. *English Teaching*. 2019. Vol. 74, no. 3. P. 3–28.
23. Halliday M. A. K. Spoken and written language. Oxford : Oxford University Press, 1989. 109 p.
24. Halliday M., Hasan R. Cohesion in English. London : Longman, 1976. 374 p.

25. Halliday M., Matthiessen C. M. I. M. An introduction to functional grammar. 3rd ed. London : Hodder Education Publishers, 2004. 700 p.
26. Hassanzadeh M., Saffari E., Rezaei S. The impact of computer-aided concept mapping on EFL learners' lexical diversity: A process writing experiment. *ReCALL*. 2021. Vol. 33, no. 3. P. 1–15.
27. Hockey S., Martin J. The oxford concordance program version 2. *Literary and linguistic computing*. 1987. Vol. 2, no. 2. P. 125–131.
28. Hout R. v., Vermeer A. Comparing measures of lexical richness. *Modelling and Assessing Vocabulary Knowledge* / ed. by H. Daller, J. Milton, J. Treffers-Daller. Cambridge. P. 93–115.
29. Hyland K. As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*. 2008. Vol. 27, no. 1. P. 4–21.
30. Hyland K. Creative writing in EAP: Exploring an emerging field. *Journal of English for Academic Purposes*. 2015. No. 18. P. 83–86.
31. Hyland K., Tse P. Metadiscourse in Academic Writing: A Reappraisal. *Applied Linguistics*. 2004. Vol. 25, no. 2. P. 156–177.
32. Investigating lexical proficiency development over time – the case of Dutch-speaking learners of French in Brussels / B. Bulté et al. *Journal of French Language Studies*. 2008. Vol. 18, no. 3. P. 277–298.
33. Jarvis S. Capturing the Diversity in Lexical Diversity. *Language Learning*. 2013. Vol. 63. P. 87–106.
34. Johansson V. Lexical diversity and lexical density in speech and writing. *Working papers / Lund University, Department of Linguistics and Phonetics*. 2008. Vol. 53. P. 61–77.
35. Kim S., Jeon M. An analysis study of English writing of elementary school 6th grade English language learners using Coh-Metrix. *The Modern English Society*. 2016. Vol. 17, no. 3. P. 263–287.
36. Kojima M., Yamashita J. Reliability of lexical richness measures based on word lists in short second language productions. *System*. 2014. Vol. 42. P. 23–33.

37. Kondal B. Effects of lexical density and lexical variety in language performance and proficiency. *International journal of IT, engineering and applied sciences research (IJIEASR)*. 2015. Vol. 4, no. 10. P. 25–28.
38. Kouachi R. A Corpus-Based Study of the Lexical Density and Readability of Student's Academic Writings The Case of Master One EFL Students at Larbi Ben M'hidi University : Dissertation. Oum El Bouaghi, 2021. 84 p.
39. Laufer B., Nation P. Vocabulary size and use: lexical richness in L2 written production. *Applied linguistics*. 1995. Vol. 16, no. 3. P. 307–322.
40. Laufer B. The Lexical Plight in Second Language Reading: Words You Don't Know, Words You Think You Know and Words You Can't Guess. *Second Language Vocabulary Acquisition: A Rationale for Pedagogy* / ed. by J. Coady, T. Huckin. Cambridge, 1997. P. 20–34.
41. Layali K., Alshlowiy A. Online writing of Covid-19 in Bahrain: An analysis of lexical diversity and writing proficiency. 2022.
42. Lexical diversity and language development: quantification and assessment / D. Malvern et al. Palgrave Macmillan, 2004. 269 p.
43. Lexical Diversity and Lexical Sophistication in First Language Writing / D. Malvern et al. *Lexical Diversity and Language Development*. London, 2004. P. 152–176.
44. Li X., Zhang H. Developmental features of lexical richness in English writings by Chinese L3 beginner learners. *Frontiers in psychology*. 2021. Vol. 12.
45. Longman Grammar of Spoken and Written English / D. Biber et al. Longman, 1999. 1203 p.
46. MacWhinney B., Snow C. The child language data exchange system: an update. *Journal of child language*. 1990. Vol. 17, no. 2. P. 457–472.
47. MacWhinney B. The childe project: tools for analyzing talk, volume I: transcription format and programs. 3rd ed. New York : Psychology Press, 2000. 384 p.
48. Maley A. Creativity in the language classroom. *The handbook of language teaching* / ed. by M. H. Long, C. J. Doughty. 2009. P. 503–518.

49. Marcella H., Nation P. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*. 2000. Vol. 13, no. 1. P. 403–430.
50. McCarthy P. M., Jarvis S. MTL, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*. 2010. Vol. 42, no. 2. P. 381–392.
51. McKee G., Malvern D., Richards B. Measuring vocabulary diversity using dedicated software. *Literary and linguistic computing*. 2000. Vol. 15, no. 3. P. 323–335.
52. McNamara D. S., Crossley S. A., McCarthy P. M. Linguistic Features of Writing Quality. *Written Communication*. 2010. Vol. 27, no. 1. P. 57–86.
53. McNamara D. S., Crossley S. A., Roscoe R. Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*. 2013. Vol. 45, no. 2. P. 499–515.
54. Modelling and Assessing Vocabulary Knowledge / ed. by H. Daller, J. Milton, J. Treffers-Daller. Cambridge : Cambridge University Press, 2007. 292 p.
55. Nasser M., Thompson P. Lexical density and diversity in dissertation abstracts: Revisiting English L1 vs. L2 text differences. *Assessing Writing*. 2021. Vol. 47. P. 1–11.
56. Nation I. How Large a Vocabulary is Needed For Reading and Listening?. *Canadian Modern Language Review*. 2006. Vol. 63, no. 1. P. 59–82.
57. Nation I. S. P. Learning Vocabulary in Another Language. Cambridge University Press : Cambridge, 2001. 492 p.
58. Nation P., Anthony L. Mid-frequency readers. *Journal of Extensive Reading*. 2013. Vol. 1. P. 5–16.
59. Nguyen B. T. T., Newton J., Tran Q. P. N. The Effect of Topic on EFL Writing by Vietnamese Tertiary Students: Insights from Combining a Lexical Richness Analysis with Student Self-Reports. *Language Related Research*. 2022. Vol. 13, no. 5. P. 511–540.
60. Nunan D. Introducing discourse analysis. London : Penguin English, 1993. 134 p.

61. Olinghouse N. G., Leaird J. T. The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing*. 2008. Vol. 22, no. 5. P. 545–565.
62. Paltridge B. *Discourse analysis: an introduction*. 2nd ed. New York : Continuum, 2012.
63. Rahmansyah H. Grammatical intricacy and lexical density of the SMA student's textbooks : Thesis. Medan, 2012. 94 p.
64. Rayson P. E. *Matrix : a statistical method and software tool for linguistic analysis through corpus comparison* : Thesis. 2002.
65. Read J. *Assessing vocabulary*. Cambridge : Cambridge University Press, 2000. 279 p.
66. Richards J. C., Jones R. H. *Creativity in Language Teaching: Perspectives from Research and Practice*. Taylor & Francis Group, 2015. 264 p.
67. Richards J. C., Schmidt R. W. *Longman dictionary of language teaching and applied linguistics*. 4th ed. Pearson Education Limited, 2010. 644 p.
68. Sadeghi K., Dilmaghani S. K. The Relationship between Lexical Diversity and Genre in Iranian EFL Learners' Writings. *Journal of Language Teaching and Research*. 2013. Vol. 4, no. 2. P. 328–334.
69. Saragih A. Metaphorical representations and scientific texts. *Englonesia: an indonesian scientific journal on linguistics and literature*. 2006. Vol. 2, no. 1. P. 1–11.
70. Schmitt N. Review article: Instructed second language vocabulary learning. *Language Teaching Research*. 2008. Vol. 12, no. 3. P. 329–363.
71. Silverman S., Ratner N. B. Measuring lexical diversity in children who stutter: application of vocd. *Journal of Fluency Disorders*. 2002. Vol. 27, no. 4. P. 289–304.
72. Sinclair J. *Corpus, Concordance, Collocation*. Oxford University Press, 1991. 179 p.
73. Stubbs M. *Words and phrases: corpus studies of lexical semantics*. Oxford : Blackwell, 2002. 267 p.

74. Swales J. M., Feak C. *Academic Writing for Graduate Students: Essential Tasks and Skills* (Michigan Series In English For Academic & Professional Purposes). 3rd ed. University of Michigan Press, 2012. 418 p.
75. Syarif H., Putri R. E. How lexical density reveals students' ability in writing academic text. *Lingua Didaktika: Jurnal Bahasa dan Pembelajaran Bahasa*. 2018. Vol. 12, no. 2. P. 86–94.
76. Textalyzer: Keyword Density + Word Count Tool. *seoscout*. URL: <https://seoscout.com/tools/text-analyzer>.
77. The Development of Writing Proficiency as a Function of Grade Level: A Linguistic Analysis / S. A. Crossley et al. *Written Communication*. 2011. Vol. 28, no. 3. P. 282–311.
78. Thida D. Lexical Density and Readability of Students' Writing. *International Journal of Science and Research (IJSR)*. 2019. Vol. 8, no. 9. P. 893–895.
79. To V., Fan S., Thomas D. Lexical density and readability: A case study of English textbooks. *Internet journal of language, culture and society*. 2013. No. 37. P. 61–69.
80. Uccelli P., Dobbs C. L., Scott J. Mastering Academic Language: Organization and Stance in the Persuasive Writing of High School Students. *Written Communication*. 2013. Vol. 30, no. 1. P. 36–62.
81. Uma J. A Study of Lexical Richness in ESL Students' Writing : Dissertation. Kuala Lumpur, 2019. 113 p.
82. Ure J. Lexical density and register differentiation. contemporary educational psychology. *Contemporary educational psychology*. 1971. No. 5. P. 96–104.
83. Vidakovic I., Barker F. Lexical development across second language proficiency levels: a corpus-informed study. 2009. P. 143–146.
84. Web VP Classic. URL: <https://www.lextutor.ca/vp/eng/>.

APPENDICES

Appendix 1

Lexical density in Ukrainian EFL students' academic writing

	2 nd year of study			4 th year of study		
Student	Lexical density	Words/ Sentence	Syllables/ Word	Lexical density	Words/ Sentence	Syllables/ Word
Student 1	46%	13.9	1.7	45%	13.8	1.6
Student 2	50%	12.2	1.7	51%	12.9	1.7
Student 3	47%	14.1	1.7	51%	14.9	1.7
Student 4	51%	12.4	1.7	49%	16.3	1.7
Student 5	53%	11.8	1.7	49%	15.1	1.6
Student 6	48%	17.8	1.7	49%	14.3	1.6
Student 7	47%	15.9	1.7	48%	15.4	1.6
Student 8	49%	18.9	1.7	50%	14.1	1.6
Student 9	50%	15.3	1.7	53%	14.2	1.7
Student 10	48%	19.9	1.7	46%	22.4	1.6
Student 11	51%	13.7	1.8	49%	16.2	1.7
Mean	49%	15	1.7	49%	15.4	1.6333
Standard deviation	2.89	2.68	0.0426	3.35	2.93	0.0173

Appendix 2

Lexical density in Ukrainian EFL students' creative writing

	2 nd year of study			4 th year of study		
Student	Lexical density	Words/ Sentence	Syllables/ Word	Lexical density	Words/ Sentence	Syllables/ Word
Student 1	38%	13.4	1.4	44%	19.1	1.4
Student 2	42%	17.2	1.7	38%	19.7	1.5
Student 3	34%	14.2	1.3	34%	33.3	1.4
Student 4	33%	16.6	1.6	40%	17.0	1.5
Student 5	40%	13.0	1.4	40%	10.8	1.5
Student 6	40%	13.6	1.4	41%	10.2	1.2
Student 7	38%	16.2	1.6	36%	20.2	1.4
Student 8	37%	15.5	1.5	39%	16.8	1.4
Student 9	36%	18.5	1.6	37%	28.8	1.3
Student 10	42%	19.5	1.6	45%	15.6	1.4
Student 11	36%	20.9	1.4	40%	25.9	1.4
Mean	38%	16.0	1.5	39.45%	20.36	1.4
Standard deviation	5.0453	3.0199	0.122	3.29	8.711	0.0775

Appendix 3

Lexical diversity in Ukrainian EFL students' academic writing

Student	2nd year of study	4th year of study
Student 1	56%	62%
Student 2	55%	69%
Student 3	56%	61%
Student 4	54%	59%
Student 5	52%	51%
Student 6	51%	51%
Student 7	57%	62%
Student 8	62%	61%
Student 9	50%	51%
Student 10	57%	55%
Student 11	53%	49%
Mean	54.82%	57.36%
Standard deviation	5.44	6.36

Appendix 4

Lexical Frequency Profiles in Ukrainian EFL students' academic writing according to the number of tokens and types

	2 nd year of study			4 th year of study		
Student	Types	TTR	TPT	Types	TTR	TPT
Student 1	137	0.55	1.82	149	0.60	1.68
Student 2	135	0.54	1.85	170	0.68	1.48
Student 3	141	0.56	1.77	148	0.59	1.69
Student 4	134	0.54	1.87	145	0.58	1.72
Student 5	123	0.49	2.03	124	0.50	2.02
Student 6	127	0.51	1.97	124	0.50	2.02
Student 7	139	0.56	1.80	150	0.60	1.67
Student 8	147	0.59	1.70	150	0.60	1.67
Student 9	123	0.49	2.03	125	0.50	2.00
Student 10	139	0.56	1.80	134	0.54	1.87
Student 11	129	0.52	1.94	123	0.49	2.03
Mean	134.00	0.54	1.87	140.18	0.56	1.80
Standard deviation	7.7	0.03	0.10	15.2	0.06	0.19

Appendix 5

Lexical Frequency Profiles in Ukrainian EFL students' academic writing according to word frequency lists

	2 nd year of study				4 th year of study			
Student	K1	K2	AWL	OLW	K1	K2	AWL	OLW
Student 1	73.60 %	4.00%	11.60 %	10.80 %	68.40 %	4.00%	14.00%	13.60%
Student 2	76.00 %	3.20%	8.40%	12.40 %	66.53 %	3.98%	14.74%	14.74%
Student 3	69.20 %	4.00%	19.20 %	7.60%	62.40 %	10.00 %	11.60%	16.00%
Student 4	69.60 %	0.80%	11.60 %	18.00 %	66.40 %	3.60%	14.80%	15.20%
Student 5	70.80 %	8.00%	8.00%	13.20 %	73.60 %	7.20%	6.40%	12.80%
Student 6	72.40 %	6.40%	10.00 %	11.20 %	75.20 %	1.60%	11.20%	12.00%
Student 7	70.80 %	3.60%	14.40 %	11.20 %	69.20 %	6.00%	14.80%	10.00%
Student 8	70.00 %	3.60%	5.20%	14.00 %	74.40 %	4.00%	15.20%	6.40%
Student 9	68.00 %	6.40%	14.40 %	11.20 %	74.00 %	3.60%	14.00%	8.40%
Student 10	74.40 %	2.00%	8.80%	14.80 %	66.00 %	5.60%	9.20%	19.20%

Student 11	73.20 %	4.00%	14.40 %	8.40%	65.60 %	3.60%	11.20%	19.60%
Mean	71.64 %	4.18%	11.45 %	12.07 %	69.25 %	4.83%	12.47%	13.45%
Standard deviation	2.46	2.06	3.94	2.92	4.36	2.27	2.82	4.14

Appendix 6

Lexical diversity in Ukrainian EFL students' creative writing

Student	2nd year of study	4th year of study
Student 1	66%	65%
Student 2	66%	64%
Student 3	57%	61%
Student 4	63%	68%
Student 5	68%	70%
Student 6	68%	67%
Student 7	63%	61%
Student 8	62%	65%
Student 9	54%	59%
Student 10	59%	60%
Student 11	65%	64%
Mean	62.82%	64%
Standard deviation	4.53	3.49

Appendix 7

Lexical Frequency Profiles in Ukrainian EFL students' creative writing according to the number of tokens and types

	2 nd year of study			4 th year of study		
Student	Types	TTR	TPT	Types	TTR	TPT
Student 1	155	0.62	1.61	156	0.62	1.60
Student 2	162	0.65	1.54	160	0.64	1.56
Student 3	138	0.55	1.81	152	0.61	1.64
Student 4	151	0.60	1.66	165	0.66	1.52
Student 5	161	0.64	1.55	172	0.69	1.45
Student 6	162	0.65	1.54	162	0.64	1.56
Student 7	154	0.62	1.62	152	0.61	1.64
Student 8	145	0.58	1.72	155	0.62	1.61
Student 9	132	0.53	1.89	146	0.58	1.71
Student 10	145	0.58	1.72	145	0.58	1.72
Student 11	153	0.61	1.63	159	0.64	1.57
Mean	150.73	0.60	1.66	156.7	0.63	1.60
Standard deviation	9.86	0.04	0.11	8.04	0.03	0.08

Appendix 8

Lexical Frequency Profiles in Ukrainian EFL students' academic writing according to word frequency lists

	2 nd year of study				4 th year of study			
Student	K1	K2	AWL	OLW	K1	K2	AWL	OLW
Student 1	87.60 %	4.80%	2.80%	4.80%	86.80 %	6.00%	0.40%	6.80%
Student 2	80.00 %	4.80%	10.00 %	5.20%	82.40 %	5.20%	6.00%	6.40%
Student 3	74.40 %	14.00 %	1.60%	10.00 %	84.80 %	4.80%	4.80%	5.60%
Student 4	86.40 %	5.60%	4.80%	3.20%	80.80 %	6.80%	1.60%	10.80%
Student 5	84.00 %	3.60%	7.60%	4.80%	74.00 %	8.80%	2.80%	14.40%
Student 6	82.00 %	7.60%	2.00%	8.40%	84.19 %	8.70%	0.40%	6.72%
Student 7	84.00 %	3.20%	8.40%	4.40%	86.00 %	4.00%	0.80%	9.20%
Student 8	87.20 %	2.80%	5.20%	4.80%	86.00 %	4.00%	1.20%	8.80%
Student 9	92.00 %	2.40%	2.40%	3.20%	82.80 %	9.60%	1.20%	6.40%
Student 10	86.00 %	4.80%	5.60%	3.60%	86.00 %	4.40%	2.40%	7.20%

Student 11	86.80 %	2.40%	4.80%	6.00%	77.60 %	8.80%	2.00%	11.60%
Mean	84.58 %	5.09%	5.02%	5.31%	82.85 %	6.46%	2.15%	8.54%
Standard deviation	4.62	3.34	2.76	2.13	4.02	2.16	1.80	2.74

Appendix 9

Lexical density and lexical diversity of the second-year Ukrainian EFL students' writing

	Academic writing		Creative writing	
	Lexical Density	Lexical Diversity	Lexical Density	Lexical Diversity
Student 1	46%	56%	38%	66%
Student 2	50%	55%	42%	66%
Student 3	47%	56%	34%	57%
Student 4	51%	54%	33%	63%
Student 5	53%	52%	40%	68%
Student 6	48%	51%	40%	68%
Student 7	47%	57%	38%	63%
Student 8	49%	62%	37%	62%
Student 9	50%	50%	36%	54%
Student 10	48%	57%	42%	59%
Student 11	51%	53%	36%	65%
Mean	49%	54.82%	38%	62.82%

Appendix 10

Lexical density and lexical diversity of the fourth-year Ukrainian EFL students' writing

	Academic writing		Creative writing	
	Lexical Density	Lexical Diversity	Lexical Density	Lexical Diversity
Student 1	45%	62%	44%	65%
Student 2	51%	69%	38%	64%
Student 3	51%	61%	34%	61%
Student 4	49%	59%	40%	68%
Student 5	49%	51%	40%	70%
Student 6	49%	51%	41%	67%
Student 7	48%	62%	36%	61%
Student 8	50%	61%	39%	65%
Student 9	53%	51%	37%	59%
Student 10	46%	55%	45%	60%
Student 11	49%	49%	40%	64%
Mean	49%	57.36%	39.45%	64%

