



TECHNICAL UNIVERSITY OF COTOPAXI

GRADUATE SCHOOL

MASTER'S DEGREE IN APPLIED LINGUISTICS TO

TEACHING ENGLISH AS A FOREIGN LANGUAGE

MODALITY: RESEARCH REPORT

Theme:

LEXICAL RICHNESS AND READABILITY OF ENGLISH DEGREE STUDENT'S WRITING

Research dissertation before obtaining the master's degree in Applied Linguistics
to Teaching English as a Foreign Language.

Author:

Romero Mayorga Nataly Monserrat Ing.

Tutor:

Tovar Viera Rodrigo Vicente Mg.

LATACUNGA –ECUADOR

2021

TUTOR'S ENDORSEMENT

In my capacity as a Supervisor of the Research dissertation titled LEXICAL RICHNESS AND READABILITY OF ENGLISH DEGREE STUDENTS' WRITING investigated by Ing. Romero Mayorga Nataly Monserrat, for obtaining the master's degree in Applied Linguistics to Teaching English as a Foreign Language.

I CERTIFY THAT:

This research dissertation has been fully revised and has the requirements and merits to be submitted for evaluation by the assigned Revision Committee and its presentation and public defense.

Latacunga, June 1st, 2021.



.....
Tovar Viera Rodrigo Vicente Mg.
ID: 0502414089

COMMITTEE APPROVAL

This research dissertation: Lexical richness and readability of English degree students' writing, has been revised, approved and authorized for printing and binding, before obtaining a Master's degree in Applied Linguistics to Teaching English as a Foreign Language; This meets the substantive and formal requirements to hand in for the presentation and defense.

Latacunga, June 24th, 2021



.....
Fabiola Soledad Cando Guanoluisa Mg.

ID: 0502884604

Committee president



.....
José Ignacio Andrade Moran Mg.

ID: 0503101040

Committee Member 1



.....
Olga Lorena Gonzalez Ortiz Ph.D

ID: 1002377271

Committee Member 2

DEDICATION

I dedicate this research work to my parents and my family since they have always been the linchpin of my life. Your support and love are the most valuable treasure. Thanks for giving me your faith.

I also dedicate this thesis to Néstor who has been there to encourage me in this journey. Thanks for your love and help, this is the beginning to start a life together.

Nataly

ACKNOWLEDGEMENT

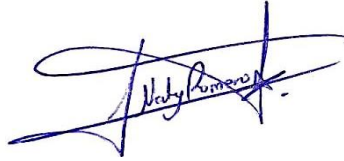
Thanks to God for giving me the strengths to end up this process and also aiding me to overcome problems. Special thanks to Mg. Rodrigo Tovar whose comments and advice were the motivation to continue researching this interesting topic. I truly appreciate you and the time you spent helping me on many occasions. I would like to recognize to Technical University of Cotopaxi for allowing the consecution of this study.

Nataly Romero M.

AUTHORSHIP

I, Romero Mayorga Nataly Monserrat declare to be the author of the concepts, procedures, and findings in this research dissertation.

Latacunga, July 5th, 2021

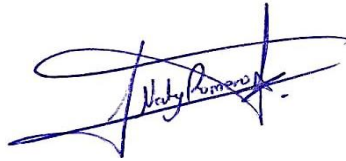


.....
Nataly Monserrat Romero Mayorga
ID: 0503615726

COPYRIGHT REFUSE

I, Romero Mayorga Nataly Monserrat confer the rights of this graduate dissertation and authorize its total reproduction or part of it, as long as it is under the regulations of the Technical University of Cotopaxi.

Latacunga, July 5th, 2021.

A handwritten signature in blue ink, appearing to read 'Nataly Romero', with a large, stylized flourish above it.

.....
Romero Mayorga Nataly Monserrat Ing.

ID: 0503615726

COMMITTEE PRESIDENT'S APPROVAL

I, Fabiola Soledad Cando Guanoluisa Mg., declare that this research dissertation: Lexical richness and readability of English degree students' writing, has the corrections and comments suggested by the members of the committee in the scientific session.

Latacunga, July 5th, 2021.



.....
Fabiola Soledad Cando Guanoluisa Mg.

ID: 0502884604

Committee President

**TECHNICAL UNIVERSITY OF COTOPAXI
GRADUATE DEPARTMENT**

**MASTER'S DEGREE IN APPLIED LINGUISTICS TO
TEACHING ENGLISH AS A FOREIGN LANGUAGE**

**THEME: LEXICAL RICHNESS AND READABILITY OF ENGLISH
DEGREE STUDENTS' WRITING**

Author: Romero Mayorga Nataly Monserrat Ing.

Tutor: Tovar Viera Rodrigo Vicente Mg.

ABSTRACT

This study aims to measure lexical richness and readability in the written production of thirty-nine pre-service teachers at Technical University of Cotopaxi in the November 2020 - March 2021 Academic Cycle. The research is descriptive with a mixed design of qualitative and quantitative approaches (QUAN/qual). The analysis follows a corpus linguistics methodology to clarify the relationship between the main variables. Lex-tutor Vocab profile and Coh-Metrix were employed to gauge lexical richness and readability. Lexical richness examination demonstrated that 83.59% of the corpus is composed of the 2000 most frequent words in English. In terms of readability, the corpus is very difficult to read. The correlational analysis revealed that lexical density, lexical sophistication, and lexical diversity explain in different proportions the variation in the readability of English degree students' writing. After analyzing the linguistic features in the corpus, it was possible to identify lexical weaknesses to create a pedagogical booklet. This material suggests activities and learning strategies for enhancing the use of vocabulary in academic tasks. The proposal was validated by experts and users to ensure its feasibility for application. To conclude, pre-service English teachers require vocabulary instruction about less frequent and academic words to increment their written products' quality. Lexical richness is only one element of the text that produces variations in readability. For this reason, the present study should be replicated in the future to analyze which other components might affect reading ease.

KEYWORDS: Lexical Richness; readability; academic writing; vocabulary learning strategies, corpus analysis.

**UNIVERSIDAD TÉCNICA DE COTOPAXI
DIRECCIÓN DE POSGRADO**

**MAESTRÍA EN LINGÜÍSTICA APLICADA A LA ENSEÑANZA DEL
IDIOMA INGLÉS COMO LENGUA EXTRANJERA**

**Título: LA RIQUEZA LÉXICA Y LEGIBILIDAD DE LA ESCRITURA DE
LOS ESTUDIANTES DE PREGRADO DE PEDAGOGÍA DEL IDIOMA
INGLÉS.**

Autor: Romero Mayorga Nataly Monserrat

Tutor: Tovar Viera Rodrigo Vicente Mg.

RESUMEN

El presente estudio tiene por objeto medir la riqueza léxica y la legibilidad de la escritura de treinta y nueve docentes en formación de la Universidad Técnica de Cotopaxi durante el ciclo académico Noviembre 2020 – Marzo 2021. La investigación es descriptiva con un diseño mixto de enfoque cuantitativo y cualitativo (QUAN/qual). El análisis sigue una metodología de lingüística de corpus para aclarar la relación entre las variables principales. Se emplearon los programas Lex-tutor Vocab profile y Coh-Metrix para medir la riqueza léxica y la legibilidad correspondientemente. La medición de la riqueza léxica demostró que el 83,59% del corpus está compuesto por las 2000 palabras más frecuentes en inglés. En términos de legibilidad, el corpus es muy difícil de leer. El análisis correlacional reveló que la densidad, sofisticación y diversidad léxica explican en diferentes proporciones la variación en la legibilidad de la escritura de los estudiantes de pregrado. Después de analizar las características lingüísticas del corpus, fue posible identificar las debilidades léxicas existentes, mismas que permitieron la creación de un folleto pedagógico con actividades y estrategias de aprendizaje que mejoren el uso de vocabulario durante la realización de tareas académicas. La propuesta fue validada por expertos y usuarios asegurando su viabilidad para posterior aplicación. Para concluir, los docentes en formación requieren aprender léxico menos frecuente y académico para incrementar la calidad de sus productos escritos. La riqueza léxica es uno de los elementos del texto que produce variaciones en la legibilidad. Por esta razón, este estudio debería replicarse para analizar qué otros componentes podrían afectar la facilidad de lectura.

PALABRAS CLAVE: Riqueza léxica; legibilidad; escritura académica; estrategias de aprendizaje de vocabulario, análisis de corpus.

INDEX

COVER	
TUTOR'S ENDORSEMENT	ii
COMMITTEE APPROVAL.....	iii
DEDICATION	iv
ACKNOWLEDGEMENT	v
AUTHORSHIP	vi
COPYRIGHT REFUSE.....	vii
COMMITTEE PRESIDENT'S APPROVAL.....	viii
ABSTRACT.....	ix
RESUMEN.....	x
INDEX	xi
INTRODUCTION	1
CHAPTER I THEORETICAL FRAMEWORK	11
1.1 Research Background.....	11
1.2 Epistemological foundation	13
1.2.1 Lexical Richness	13
1.2.2 Measuring lexical richness.....	15
1.2.3 Vocabulary size.....	17
1.2.4 Readability	18
1.2.5 Measures of readability	20
1.2.6 Relationship between lexical richness and readability.....	22
1.2.7 Explicit and incidental vocabulary learning.....	22
1.2.8 How to bring vocabulary into productive use	23
1.2.9 Vocabulary learning strategies and activities.....	24
1.2.10 Learn vocabulary from reading	26
1.2.11 Task-based learning for writing	26
1.3 Foundation of the state of the art.....	27
1.4 Chapter I Conclusions	30

CHAPTER II PROPOSAL	31
2.1 Proposal topic.....	31
2.2 Objectives.....	31
2.3 Justification	31
2.4 Proposal development	35
2.4.1 Proposal components	36
2.4.2 Proposal explanation	36
2.4.3 Premises for its implementation.....	38
2.5 Chapter II Conclusions.....	74
CHAPTER III. APPLICATION AND/OR VALIDATION OF THE PROPOSAL	75
3.1 Expert’s evaluation.....	75
3.2 User’s evaluation.....	76
3.3 Impact or results evaluation	77
3.4 Proposal Results	78
3.5 Chapter III Conclusions	79
General Conclusions	80
Recommendations	80
References	81
Appendices	90
Appendix 1	90
Appendix 2	98
Appendix 3	99
Appendix 4	101
Appendix 5	102
Appendix 6	104
Appendix 7	106
Appendix 8	111
Appendix 9	113
Appendix 10.....	121

Appendix 11	123
Appendix 12	124
Appendix 13	135
Appendix 14	139
Appendix 15	140
Appendix 16	142

INTRODUCTION

Education is the basis of an organized and sustainable society. Throughout history, people assimilated knowledge and transformed it to solve their common problems. This information has passed from generation to generation employing oral communication to preserve traditions, values, and the understanding of the world from each civilization's point of view. However, the exchange of ideas and technology between cultures transformed the reality over the years, developing a worldwide community in which English is the lingua franca. For this reason, in Ecuador, the use and learning of this foreign language constitute an opportunity for human interaction that breaks social progress barriers at many levels.

The article 350 of the Constitution of the Republic of Ecuador, mentions that the Ecuadorian Higher Education System has as purposes to form academically and professionally with a humanistic and scientific vision; scientific and technological research; innovation; promotion, development, and diffusion of ancestral knowledge; and building solutions for the country's problems related to the objectives of development of the regimen. These responsibilities are regulated by the Organic Law of Higher Education (henceforth OLHE) which controls universities, polytechnic schools, pedagogical, technical, technological institutes, and arts conservatories. This legislation was reformed in 2010 and has concordance with the information indicated in the Constitution. Article number 5, settles the student's rights, item b) remarks that students own the right to access to higher education of quality, which allows them to start an academic or professional career with equal opportunities.

The standards of higher education of Ecuador established in article number 84 of the Organic Law of Higher Education also mention that universities can include bylaw academic and disciplinary requisites for the approval of courses and third-level education degrees. Second language learning at high-level of education in Ecuador is mandatory for English degree students according to article 80 of the Regulation of Academic Regime (2020). The standards for in-service English teachers who work in public establishments are based on the document developed

by the Teachers of English to Speakers of Other Languages (TESOL) as the organization's K-12 ESL Teacher Standards (2009). In 2012, the government announced that it required English language teachers to reach B2 level according to the Common European Framework of Reference for Languages (CEFR). So, the outcome profile of pre-service English teachers established by the university is B2 level of English proficiency.

The reforms in OLHE also brought the enhancement of research output for students and teachers. There were assigned resources and facilities to support students' participation in research projects offering scholarships and leveling courses to be more competitive in this area. The article 8 of the OLHE in item f), describes that higher education will promote and execute investigation programs in the scientific, technological and pedagogical domains to assist the improvement and protection of the environment and national sustainable development. The statute of Technical University of Cotopaxi also recognizes in its article 5, item f) the relevance of research as a responsibility for the institution. For that reason, were approved Research Regulations in 2017 to reform a previous version according to the Plan for Strengthening and Academic Excellence.

The article 50 of the Regulation of Academic Regime specifies that Scientific dissemination consists of transmitting results, advances, ideas, hypotheses, theories, concepts, and in general any scientific activity or technological for the society; using the right channels, resources and languages that society can understand and assimilate. Consequently, "university students must be actively involved in research during their study, and be trained in the reading and analysis of research papers in English because it covers a broader view of the scientific developments around the globe" (Feyen, 2019, p. 11). English is considered the global academic language and pre-service teachers in Cotopaxi province at Technical University of Cotopaxi, require good writing skills and precise vocabulary in English to disseminate their ideas in the scientific community.

The problem statement is related to the learners, who entail proficiency not only in teaching skills but also in comprehending and using a foreign language. They usually deal with problems during writing related to grammar structures,

punctuation, coherence, and the usage of a lexical background resource from a different language that most of them have not experienced in context. Indeed, learning a foreign language usually begins with acquiring vocabulary through listening and reading to produce quality communication via speaking and writing. None skill could be developed without a vocabulary base. Instructors aim that a learner use words to reflect the meaning and intention that a native speaker will give in a natural conversation or written expression. Nation mentioned that “Vocabulary choice is a strong indicator of whether the writer has adopted the conventions of the relevant discourse community” (2001, p. 276). Pre-service English teachers frequently use complex terminology in their works, producing lexically dense texts with subordinate clauses and passive expressions. Excluding the grammatical aspects, other elements could distort the comprehension of the written productions (Sağlamel Kayaoğlu, 2015). For example, misuse of terminology because of lack of experience as the writers are not native speakers or misunderstanding the actual signified of the words.

The problem is formulated as a question: Is lexical richness related to the readability of English degree students’ writing? The following sub-questions will support this central research question; Are lexical density, lexical diversity, and lexical sophistication related to the ease of reading scores? Does the number of Academic Word List (henceforth AWL) families in the text influence readability? Which are the independent variables related to lexical richness that produce variability in the readability of the text? These interrogations require to be answered to understand their interaction and relevance in academic written products. If writers do not employ vocabulary according to an academic setting or use a limited lexical resource, that might interfere with the academic text’s purpose and the idea they want to transmit. This observation allows assuming that “clearly vocabulary plays a significant role in the assessment of the quality of written work” (Nation, 2001, p. 277).

After defining the problem to be solved, **the research’s main objective** will be to measure lexical richness and readability in English degree students’ writing for its assessment and improvement. This investigation of variables using technological tools can offer a broader view of the vocabulary employed and strategies to enhance

these features in pre-service English teachers' writing. The accomplishment of this goal requires **specific objectives**, which are the following:

- To set the theoretical basis and previous works in the field related to the problem under study.
- To gauge lexical richness and readability in an academic corpus through a technological lens.
- To validate a proposal design that aims to enhance the academic writing of pre-service English teachers.

There is a system of tasks concerning the specific objectives described in the following table:

Table 1. System of tasks concerning the specific objectives

Specific objectives	System of tasks
<p>Objective 1: To set the theoretical basis and previous work in the analysis of lexical richness and readability through a corpus linguistics methodology.</p>	<ul style="list-style-type: none"> ● Investigate previous work related to the epistemic foundation. ● Systematize the variables that explain the problem under study to structure a theoretical framework ● Explore the impact of contemporary research to exhibit the probable impact of the investigation.
<p>Objective 2: To gauge lexical richness and readability in an academic corpus through technological tool lens.</p>	<ul style="list-style-type: none"> ● Diagnose written receptive vocabulary knowledge of pre-service English teachers applying a multiple-choice test to determine the participants' level of vocabulary knowledge.

- Prepare the corpus by asking English degree students to produce an academic text with finite length.
- Establish parameters for measuring lexical richness and readability in a cleaned corpus.
- Measure lexical richness and readability of texts with technological tools (Lex-tutor Vocab profile and Coh-Metrix).
- Obtain values to understand the influence of lexical richness and readability in the quality of writing.
- Analyze data statistically to determine the relationship between variables.
- Interpret results to suggest a solution proposal.
- Propose a booklet of strategies to improve the written production of English degree students.
- Evaluate the probable results of the booklet with a valuation from users' and experts' opinions.
- Elaborate general conclusions of the research

Objective 3: To validate a proposal design that aims to enhance pre-service English teachers' academic writing.

Author: Nataly Romero

The problem passed through some stages over the years:

Table 2. Stages of the problem under study

Stage	Description
Primitive Corpus linguistics	<p>“The studies of child language in the diary studies period of language acquisition research (roughly 1876 - 1926) were based on carefully composed parental diaries recording the child’s locutions” (McEnery et al., 2001, p. 3). These studies were the beginning of employing corpus analysis to enhance the learning of a language. Of course, there were many studies during the period of Comparative Linguistics. “Eaton’s study in 1940, comparing the frequency of word meanings in Dutch, French, German, and Italian. The work is very sophisticated” (McEnery et al., 2001, p. 4).</p>
Early Corpus linguistics	<p>Fries and Traver (1940) and Bongers (1947) are examples of linguists who used the corpus in foreign language pedagogy research. These studies were limited because of the time and effort that the analysis required.</p>
Rationalism influence	<p>“Noam Chomsky changed the object of linguistic inquiry from an abstract description of language to theories which reflected a psychological reality, cognitively plausible models of language. In doing so, he apparently invalidates the corpus as a source of evidence in linguistic inquiry” (McEnery et al., 200, p.6). In this stage, just a few researchers use corpora. They employed it in phonetic studies because of the underestimation of performance at that time.</p>
Modern corpus - new parameters	<p>“Pioneers worked on with corpus data throughout the 50s, 60s, and 70s, and it is largely through their efforts that</p>

and computers
program design

corpus work revived” (McEnery et al., 2001, p.20). The type-token ratios (henceforth TTR) were the first ratios used in the attempt to gauge lexical diversity. The critical knot emerged in the analysis of long length text samples where the TTR values decreased if the number of tokens was large since “the number of word tokens can increase infinitely, and although the same is true for word types, it is often necessary for the writer or speaker to re-use several function words in order to produce one new (lexical) word” (Johanson, 2008, p. 63). These quantitative difficulties have been partially solved using computers, statistical methods, and qualitative analysis. “In the field of applied linguistics, more and more researchers and practitioners treasure what corpus linguistics has to offer to language pedagogy” (Romer, 2011, p. 205).

Corpus
Linguistics

“It is an area of computational linguistics that provides large quantities of empirical language databases accumulated systematically from various fields of actual language use following some statistical methods and techniques of data sampling. It also provides sophisticated devices to analyze these corpora to extract linguistic data, examples, and information necessary in applied linguistics, computational linguistics, and artificial intelligence for understanding human language in a better way as well as for applying this data and information in various fields of human knowledge” (Dash, 2008, p .3)

Author: Nataly Romero

After the description of the problem's background, it is necessary to give a **justification** for the development of the present research. Pre-service English teachers in the third level of education need competence using academic language to write their thesis work that is a requirement to achieve their degree and because it is a valuable tool for performing international tests like TOEFL or FCE. These evaluations are commonly demanded by institutions that request staff or a pre-requirement to get scholarships. Hence, this analysis is expected to establish students' vocabulary knowledge and to assess their written skills to strengthen them.

“Lexical richness in writing will be useful for determining the factors that affect judgments of quality in writing” (Laufer and Nation, 1995, p.1). For this reason, the present study encourages, on the one hand, the examination of the lexical profile employed by the students in their written works, which could stand as an indicator of the readability features of their creations and improvement. On the other hand, the investigation proposal seeks to analyze the relationship between lexical richness and readability in academic texts. Thus a “lexically diverse text is usually regarded as being more competent and more persuasive in its effect than an equivalent low-diversity reproduction of the same text” (Kakkonen, 2012. p.1). Another study showed that texts with high lexical density, not always obtain higher values of readability (To et al., 2013, p. 69). Consequently, the correlation between these variables might get a definition after this investigation's consecution.

Little research that uses samples of text written by English Degree students has been carried out before at Technical University of Cotopaxi. It is an innovative way of incorporating applied linguistics throughout the investigation process to enrich students' training who are coursing Academic Writing. It is feasible because the study uses written pieces produced by students. Therefore, to apply this methodology, it is not necessary to have direct contact with the participants preventing any contagious related to the Covid-19 virus, which is a pandemic at the moment. The study does not require extra expenses to be ongoing; it will directly benefit linguistics in Cotopaxi and the pre-service English teachers. Thus, it is possible to receive the university's support to collect the data and apply the proposal in the future.

The research is **descriptive**, with a mixed design of quantitative and qualitative approaches (QUAN/qual). It applies specific techniques to collect, process, and analyze data with specialized software. In this case, estimating the relation between the variables will be correlated for later generate a qualitative interpretation through a logical analysis of the numbers as a product of a conversion mixed design. The study describes both variables and their components. On one side, lexical richness being the first variable that functions as an umbrella term for lexical diversity, lexical sophistication, and lexical density. On the other side, the readability of academic texts produced by pre-service English teachers.

The population selected for the research is constituted by the students of Technical University of Cotopaxi. The sample is delimited to thirty-nine English degree students coursing the seventh level of the career and studying Academic Writing in the November 2020- March 2021 Academic Cycle. The scheme of the instruments applied for the diagnostic phase was designed and sent with the application Google Forms online survey to determine the writers' vocabulary level size. The data source was 39 authentic academic texts written by pre-service English teachers converted into a computer-readable TXT file. These students are Ecuadorian Spanish speakers learning English as a foreign language to get their professionalization. Each text has 350 to 400 words to keep the corpus balance and describes a thesis topic related to language teaching. This sample was selected because it reflects the writers' productive language knowledge regarding lexical richness and readability.

The collection of the data was completed through e-mail, where the participants sent their works and considered a full range of variability of the population. Each text was carefully revised, and the researcher determined that they have different lengths. To solve this inconvenience, they were reduced by employing the online software Text Compactor. This tool simplifies the word count, preserving the main idea of the written product. After this process, the corpus collected 14.937 words and overpassed a pre-process in which references, citations, repetitive table headings were deleted to prepare the text for the analysis. It was demarcated only external criteria to guarantee representativeness and balance in the corpus of study. These criteria included the primary channel, the creators who are pre-service

English teachers, the purpose, and the genre. Antfile Converter was used to transform the main file into TXT format to make it readable in the previously selected software systems. The last step was cleaning the file to check spelling mistakes or writing errors that might occur to obtain data that represents purely English language form.

The methods of data analysis included using proven and validated computational text analysis websites such as Lextutor Vocab profile and Cohmetrix which have been employed in many types of language analysis research and offered a quantitative interpretation of the variables after examined each production and the total corpus. In the first stage of analysis, texts were individually introduced into the software to distinguish their lexical density. After this, there was a reduction of individual text length, and the individual analysis of texts was done again. **(See appendix 4)** After the corpus composition, the researcher analyzed lexical richness, lexical density, and lexical sophistication. There was also an analysis in Cohmetrix Software using the same corpus that offers readability indexes and easability indicators.

Then, it was a statistical analysis of data to determine the correlation between the variables. There was a conversion of the quantitative data into a qualitative narrative for the research's inferential stage to offer extra information about the multi-level text analysis. Finally, the researcher will suggest improvement strategies for the products through the development of a pedagogical proposal.

CHAPTER I THEORETICAL FRAMEWORK

1.1 Research Background

After the revision of similar research projects in different repositories from Europe, Asia, the United States, and Ecuadorian universities, it is possible to state there is a research gap to be filled with the present investigation on what refers to the study of lexical richness and readability of written corpus. Because this analysis has not yet been clarified, there are mentioned similar related works that analyze the main variables by employing the same methodology but in a separate form.

In 2013, a study named Lexical Density and Readability on non-English majored freshmen's writing in Vietnamese Context from the researchers, To & Thi from KOTO Foundation and Hanoi National University of Education, analyzed twenty-six written products. They demonstrated that the students had an average level of lexical density and readability. In this research, they applied a statistical design to calculate lexical density and readability using the software Lexical Complexity Analyzer and the integrated tool of Microsoft Office that employs Flesch Formulas of readability.

The research conclusions suggested the increment of training to enhance and improve vocabulary use through writing models and activities about grammatical structures. To improve writing skills, this study suggests arranging groups of study to exchange writings for feedback on grammar and vocabulary. It also recommends to teachers develop awareness about the importance of grammatical structures and vocabulary in written compositions through the organization of groups to provide the learners with printed and online resources for self-study. These materials should contain good samples of written language that can be adapted.

Sari (2019) studied the lexical richness and syntactic complexity of the English exam papers of Joint Entrance Selection of State Universities, Seleksi Bersama

Masuk Perguruan Tinggi Negeri (henceforth SBMPTN) using a corpus from 2009 to 2018. This research from Airlangga University had a quantitative approach. He analyzed lexical density, lexical variation, and complexity using tools like Vocabprofile, AntWord, and other computerized programs. He found that lexical density had the most outstanding values in 2010 and that the sophistication of the writing was fitted to college-level English writing and undergraduate research articles. There is not a pedagogical implication as it is a linguistics study that is not focused on teaching or in developing a proposal. However, the author suggests improving sophistication, use academic words, and avoid repetition to enhance writing skills before taking this kind of test.

Many studies associate lexical richness in texts with academic success and quality in writing. Melanie Gonzalez (2013) Studied the intricate relationship between measures of vocabulary size and lexical diversity, as evidenced in non-native and native speaker academic compositions. She used the Measure of Textual Lexical Diversity, Voc-D, and CELEX from the computational program Coh- Metrix to gauge lexical diversity and vocabulary size. At the same time, the scores of writings came from a rubric.

The corpus consisted of 104 essays from advanced-level non-native writers and 68 essays from native speakers in the first year of composition. She found that vocabulary size increases the writing scores at the lower proficiency levels, but in an essay is lexical diversity which promotes higher scores. The pedagogical implication is that teachers should instruct medium frequency words and similar synonyms in a similar frequency to enhance writing proficiency, avoiding less-frequent vocabulary to prevent confusion.

Readability in connection with lexical richness has also been an object of research. Malverdi & Heidari (2018) published an article named; EFL Textbook Evaluation: An Analysis of Readability and Vocabulary Profiler of Four Corners Book Series. They investigated if there is a significant relationship between readability and vocabulary profile of the Four Corners Book Series. In this analysis, there was no significant correlation relationship between readability and most frequent words. Then, if the text is more difficult, the number of k1 words does not change, the

relationship between academic words and readability was also observed. The conclusion shows that in this study there was no significant correlation between them. So, according to the results, some other factors such as sentence length, syntactic complexity, and learner's background knowledge might contribute to the difficulty of the texts. The pedagogical implication summarizes that teachers and learners should take into account sentence length and syntactic complexity in the execution of productive activities to increase readability and during the selection of texts to offer appropriate resources for students, this study does not include a proposal.

Nouri & Badia (2018) studied the lexical frequency effect on reading comprehension and recall. This study evaluated the vocabulary size and the basic reading comprehension of 80 EFL students. The results demonstrated that text coverage is related to comprehension. So, the use of familiar or more frequent words in a text, increases the reading ease, suggesting the crucial role of vocabulary, especially semantic frequency in reading comprehension and recall. EFL teachers should systematically teach and test vocabulary, especially words in association since they are stored in the mental lexicon as a network. The pedagogical implications included exposing students to both, high and low synonyms to increase their vocabulary size. The training in reading and vocabulary learning strategies to help them with tools that expose them to a wide range of lexis.

1.2 Epistemological foundation

This section explains the dependent and independent variables under research and other concepts that help to develop the methodology of analysis. It also contains the most relevant theories on which the development of the proposal will be based.

1.2.1 Lexical Richness

“Lexical richness is defined as the superordinate term for the effective use of vocabulary in good writing” (Read, 2000, p. 192). It combines a wide range of characteristics present in a text, such as lexical sophistication (LS), lexical variation (LV), lexical density (LD), a low number of errors, and even lexical individuality. (Laufer & Nation, 1995, p. 209). This feature can be valuable in academic writing as it creates more attractive texts preventing tediousness. One of the reasons is

mentioned by Laufer & Nation (1995), “a well-used rich vocabulary is likely to have a positive effect on the reader” (p. 307). Another reason is that “knowledge of diverse words enables learners to avoid repetition of words by synonyms, superordinates, and other related terminology” (Read, 2000, p. 16-37). Other authors, like Arnaud in 1984, considered lexical richness as synonymous with lexical diversity.

Laufer & Nation (1995) describe lexical sophistication as the number of advanced words present in the text, taking into account what the investigator considers advanced. LS calculation is achieved by the division between the number of lexical tokens, also called content words. It depends on the investigation’s purpose or the level of education and specialization of the participants. For this reason, it is limited if the analysis includes a comparison of the writing of students with different systems of education. (p. 309-310). For example, a specialized person may achieve higher values of the ratio because of their lack of experience in specific topics or common jargon because of lower-frequency vocabulary. In English, “it is often associated with technical terminology, greater precision, abstraction, semantic, and morphological complexity” (Richards et al., 2008, p. 5-6).

According to Laufer & Nation (1995), lexical diversity is the type-token ratio (TTR), which is calculated by dividing the total number of words (types) divided by the total number of words (tokens). (p. 310). The interpretation of the value will consider that “The more types there are in comparison to the number of tokens, then the more varied is the vocabulary” (Graham Wilson, 2014). This ratio generates problems in the contrast of long length text samples where the TTR values decrease if the number of tokens is large since “the number of word tokens can increase infinitely, and although the same is true for word types, it is often necessary” (Johansson, 2008, p. 63). Another aspect taken into account during measuring lexical richness is lexical density. This index reflects how many different lexical tokens exist in the text concerning the total number of tokens. Lexical words are also called content words and include nouns, verbs, adjectives, adverbs, and it is a common matter in lexical analysis.

1.2.2 Measuring lexical richness

Measuring lexical richness in the text might be a complex activity since particular elements are not always considered on the calculus of quantitative indices. Breeze mentioned in her study, “There are many factors besides vocabulary size that could affect lexical richness in writing. These could include familiarity with the topic, skill in writing, and communicative purpose” (2008, p. 308). Considering these aspects, the study participants should have the same level of instruction and the same purpose. It is also necessary to diagnose the initial vocabulary level to begin the measuring because it allows determining the validity of the research indicators. As Aviad-Levitzky & Laufer mentioned, “effective vocabulary use in writing has been found to have a positive influence on measures of the quality of writing and one’s general language level” (2013, p. 128).

Indeed, vocabulary is a significant element of writing in a second language, even more, if we refer to academic writing as the terminology and style are different. In their study, Tidball & Treffers-Daller established that “many measures of lexical richness are based on the assumption that the key factor behind the difficulty of a lexical item is its frequency” (2008, p. 299). The theory of word-frequency distribution began with the studies of Jean Baptiste Estoup in 1912. He ordered the frequency of words found in a long article, beginning with those which appeared the most in the text to the word that appeared the least number of times. Estoup assigned a serial number for each absolute frequency and realized that “the absolute frequency product of a word and its corresponding serial number r is generally stable under constant K , represented as $n_r \times r = K$ ” (Qiu et al., 2017, p.124). However, George Zipf, in 1948, developed the mechanism of Zipf’s Law based on the principle of the least effort and revolutionized research in more than one field.

The less effort principle in linguistics is connected to the simplification in the use of language. During oral or written expression, the sender and the receptor want to exchange meaning using a small vocabulary of common words to understand the message easily. Thus, certain words are more employed than others in a language. Zipf’s law establishes a regularity in the structure of word frequencies in a corpus that generates a constant value if we multiply the rank by the word frequency. Then,

the word frequency is inversely proportional to its frequency rank, and in a graphical representation, the vocabulary presents a picture of the hyperbola (1949, p. 25-40). After this law, many researchers performed studies to apply Zipf's formula to improve and describe frequency distribution of other languages.

Zipf's law inspired investigators like Charles Muller (1964) to develop lexical statistics for solving problems associated with the use of vocabulary and its structural analysis. There was a proposed formula that produced; as a result of enough values to develop a theoretical curve of vocabulary growth. This demonstrates that specialized vocabularies grow with the corpus and suggested the relationship between text length and vocabulary size (Muller, 1977, p. 33-34). These text difficulty elements were also studied to obtain a measure of the lexical richness of a text called Lexical Frequency Profile, which gauges productive vocabulary knowledge. This quantitative diagnosis reflects the writer's vocabulary size represented by the proportion of words a person uses to form a list of words classified according to the frequency of use in English. "The LFP shows the percentage of words a learner uses at different vocabulary frequency levels" (Laufer & Nation, 1995, p. 311). The vocabulary employed in the written corpus is divided into four groups, called layers, the first 1.000 more frequent words in English, the second 1.000 more frequent words, words from the academic word list, and less-frequent words which do not appear in the three layers mentioned before.

Coxhead (2000) demonstrated that academic texts characterize having a high percentage of words from the Academic Word List compared to other genres. He found that the proportions of the coverage by different kinds of vocabulary in the academic corpus were structured as follows; the first and second 1000 most frequent words constituted 76%, AWL covered 10%, and the less frequent words the rest of the corpus (p. 222 - 224). The lexical profile reveals the percentage of words in the text concerning each group mentioned before and offers a guide to instruct the students in vocabulary items that might enhance their productive skills with academic purposes. Laufer & Nation (1995) concluded that "the LFP is a reliable measure of the quality of lexis in writing that correlates well with other lexical measures, discriminates between learners of different proficiency levels, and is relatively stable across two pieces of writing by the same learner" (p. 319).

1.2.3 Vocabulary size

The estimation of vocabulary size can determine the vocabulary employed by pre-service English teachers on their written products and establish the relationship between receptive and productive vocabulary. Nation mentioned that “receptive vocabulary use involves perceiving the form of a word while listening or reading and retrieving its meaning. Productive vocabulary use involves wanting to express meaning through speaking or writing and retrieving and producing the appropriate word form” (2001, p. 38).

These terms are often treated as equals with passive and active vocabulary; the distinction between them sits on the association of knowledge.

Active vocabulary is easily accessed from anywhere in the vocabulary network, and in its turn allows easy access to other parts of the system too. Passive vocabulary, on the other hand, comprises vocabulary items that are part of the overall system but which cannot be reached from other parts of the network. (Meara, 1990, p. 3)

In general, passive vocabulary includes all the words that people comprehend and recognize; nevertheless, they do not commonly handle during speaking or writing and require external stimuli to be brought up. Active vocabulary is composed of words that connect with other words and are usually activated because of these associations existing in the speaker’s brain. Vocabulary size and knowledge seem to impact on writing and reading performances of EFL students. In 2017, Karakoç & Köse demonstrated that receptive vocabulary knowledge (passive) was more extensive than productive vocabulary knowledge (active). The lexical level significantly correlated with the participants’ productive vocabulary knowledge. The research showed a positive correlation between 2000 more frequent word family levels of productive vocabulary knowledge and reading comprehension, and the relationship between 2000 more frequent word family levels of productive vocabulary and writing scores were moderated.

The vocabulary size test was designed to assess passive vocabulary knowledge concerning written word form, form-meaning connection, and a minor degree of concept knowledge. VST also measures the necessary knowledge for reading as the

better the words are known, the easier the reading will be. It tests up to the 14th 1000-word family level due to multiple-choice questions without context; each item represents a language's frequency level. The VST requires choosing the most accurate option to define a word from four options, and the result estimates the test-taker total vocabulary size by multiplying the number of correct answers by 100. "The vocabulary size of undergraduate non-native speakers of non-European background successfully coping with study at an English-speaking university achieves 5000 to 6000-word families vocabulary size" (Nation & Beglar, 2007, p. 9-12).

Vocabulary Level Test (VLT) diagnoses written receptive vocabulary knowledge according to frequency levels. VLT has been modified since Paul Nation created it in 1983. The newest version is the New Vocabulary Level Test (NVLT), developed by Stuart Mclean and Brandon Kramer in 2015. This test reflects learners' knowledge at six frequency levels (1000, 2000, 3000, 4000, 5000, and AWL). It contains 120 multiple-choice items that measure the 5000 most frequent families and 30 items to measure Academic Word List knowledge. The interpretation of the test considers that "for mastering a single 1000-word level should be at least 23 out of 24 correct items and mastering the most frequent academic vocabulary should be defined as correctly answering 29 or more of the 30 AWL items" (McLean & Kramer, 2015, p.6)

1.2.4 Readability

Another aspect that students should take into account during writing is the readability of academic texts. "Readability is what makes some texts easier to read than others. It is often confused with legibility, which concerns typeface and layout" (Dubay, 2004, p. 3). So, it has a relation with the comprehension of what is written. Writing aims to allow communication and convey meaning through text, but it loses its value if the written product is not clear. Then, all the time expended in writing and editing will be lost. For that reason, many experts have developed ways to measure this critical feature in writing for enhancement.

Many factors can influence readability in texts. Gray & Leary (1935), mentioned four major categories that collect the features that make a book readable according

to experts in the matter. Their study revealed that vocabulary and sentences constitute essential elements of expression. If these are readable, they might ensure readable paragraphs and chapters in the text. (p. 10). Then, vocabulary constitutes the basis of language, and the kind of words covering a text also affects reading comprehension. It means that written productions with a high number of unknown words might be more difficult to read. However, it is not mandatory to know each lexical item to understand a whole text. It is also possible to infer the meaning from the context. Nation (2006), Van Zeeland & Schmitt (2012), and Schmitt et al. (2015) studied the proportion of known words that a reader requires to understand written language. Their results showed that 95% to 98% of coverage (known words) could be enough for reading comprehension.

The length of sentences and diversity in texts is also examined to settle readability. In 1880, Lucius Sherman applied a statistical approach to literature studies. He analyzed sentence length, concluding that “there is such a thing as unconscious sentence-rhythm, is it constant in different works of the same author, especially when written in different styles and at widely distant intervals” (Sherman, 1893, p. 260). This consistency in the length of sentences in written productions from the same writer in different periods validated the use of text samples to predict readability. Authors with more years of experience produced more simple texts with shorter sentences and some spoken language characteristics. It also encouraged Sherman to propose that shorter sentences and specific vocabulary increase the ease of reading.

Chall (1981) stated that written products are easier to read if the writer uses a low number of different words, short known words, and short, simple sentences. (p. 2) Many formulas use these elements to set the ease of reading. Lexical diversity is an indicator of cohesion; when more words are used in a text, it implies a higher cohesion level that influences the reader’s understanding. The American Press Institute study developed between 2006 – 2009 showed that the average number of words in a sentence correlates with readability. Sentences with eight or fewer words were 100% understood by readers, while the sentences with fourteen words were understood in 90%. However, in long sentences with 43 words or longer, the understanding of information decreased to 10%. For this reason, Cutts (2013)

wrote, “Better to aim for an average of 15-20 words” (p. 2). It means that it is not a specific length for insurance reading comprehension of a sentence; however, there is a gap that writers should contemplate.

1.2.5 Measures of readability

The measure of readability in the 1920s was related to improving texts for foreign students and finding solutions for educational problems. Edward Thorndike organized the most frequent 10,000 words used in English, usually found in general literature samples, and compiled them in a book called “The teacher’s word book”. “Lexical units were alphabetically classified according to their difficulty to help teachers estimate the commonness and importance of them according to their students’ reality” (Thorndike, 1921, p.4). Vocabulary is a central element of language, and it is necessary to teach what is going to be useful for students and select material that augments their lexical repertoire. In this study, Thorndike proposed a mathematical formula to determine the difficulty of words and created this guide to instruct vocabulary, which set him as a pioneer in readability studies.

Readability formulas have been used in many kinds of research and studies with some limitations. (Dubay, 2004, p.2). These formulas are defined as “an equation that gives an estimate of the readability of a text. The estimate is generally in terms of the number of years of education one needs to have to comprehend that text” (Kondru, 2006, p. 3). The Flesch Reading Ease formula is one of the most employed for readability; Rudolf Flesch developed it in 1948. This equation uses calculus elements such as the average sentence length (ASL) and the average number of syllables per word (ASW) to offer a readability score from 0 to 100, where a higher value indicates an easy-to-read text. The formula is provided below:

$$\text{Flesch Reading Ease} = 206.835 - (1.015 * \text{ASL}) - (84.6 * \text{ASW})$$

However, formulas sometimes fail to analyze characteristics beyond text length or word frequency and influence comprehension. The use of computers has facilitated linguistic analysis, and one of the more asserted software to accomplish this function is Coh-Metrix T.E.R.A, (McNamara et al., 2010). This automated text analysis system studies multiple levels of cohesion and text difficulty that well employed might facilitate comprehension during reading. Halliday and Hasan

defined cohesion from a semantic view as “the relation of meaning that exists within the text, and that defines it is a text” (Halliday & Hassan, 1976, p.5). There are two types, the first one is lexical cohesion that is translated as adequate use of vocabulary, and the second one is grammatical cohesion. Both are necessary to convey meaning in a written piece. Then, cohesion is “a set of semantic resources for linking a sentence with what has gone before” (Halliday and Hasan, 1976, p.10).

Academic writing requires cohesion to be understood by the reader. Cohesion is expressed partly through the grammar and partly through the vocabulary. (Halliday and Hasan, 1976, p.5). High cohesion texts are beneficial for low-knowledge readers because they show a clear connection between ideas creating a solid setting easier to follow for the reader who can omit re-reading the same passage for comprehension (McNamara, 2001). Thus, coherence is also an important property of quality in writing that can be measured with computational aid, saving time, and offering high reliability. “Coh-Metrix T.E.R.A. offers information about five major factors that systematically vary as a function of types of texts and grade level: narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion” (Graesser et al., 2014, p. 210)

These five dimensions show the easability and readability of the written product. Narrativity is intuitive and connected with conversation. It includes characters, events, places, and familiar things for the reader. Syntactic simplicity refers to simple sentence structures, fewer words per sentence, and repetitions of structures that create a text easier to understand. Word concreteness is the employment of words that evoke mental images in the reader from real experiences, while abstract words are less meaningful and may distort the text’s comprehension. Referential cohesion is the overlap across the sentences and ideas in the entire text that creates a connection between the writing. Deep cohesion measures connectiveness between the text. It helps the reader to create a deep understanding thanks to the ties within the writing. (Graesser et al., 2014, p. 215)

1.2.6 Relationship between lexical richness and readability

There is a limited scope of previous research about the relationship between these variables. However, vocabulary takes part in many readability measures and language itself is a selection of lexical items to develop communicative acts with the agreement of the speaking community. So, vocabulary is important to convey meaning in a written or spoken way. According to Halliday and Hassan, some elements in a text depend on others to be decoded or interpreted. “Meaning is put into wording, and wording into sound or writing” (1976, p. 5). The dependence that exists between some elements in a text to achieve meaning is called cohesion. Texts require this kind of “glue” to be differentiated from a group of sentences without any connection or purpose.

Halliday and Matthiessen identified three ways by which cohesion is created in English; conjunctions, reference, and lexical cohesion. “Lexical cohesion in a text comes about through the selection of lexical items that are related in some way to those that have gone before” (2013, p. 642). Grammatical and lexical cohesion produce effects on readability and also in written performance. Rezaee & Norouzi (2011) found in their research that “There is a significant correlation between the readability of passages and the learners’ performance. The learners’ performance correlated significantly with grammatical cohesion markers at intermediate level and with lexical cohesion markers at advanced level” (p. 1005). Then, word choice is relevant for writing to make products more understandable for a target audience.

1.2.7 Explicit and incidental vocabulary learning

Vocabulary learning has traditionally been considered as a part of other skills development such as writing, speaking, reading, or listening. Vocabulary is a relevant element of language learning, Nation (2001) explains how lexicon can be instructed in a foreign language. “The core idea is that through careful analysis of both the target language and the needs of particular groups of learners, instructable portions of a second lexicon can be identified and the effects of knowing them predicted” (Cobb, 2001, p.477).

Learning vocabulary is a cumulative process that requires both explicit and incidental learning. On one hand, explicit learning is a conscious process that

intends to push a word from short-term memory to long-term memory through the application of a variety of techniques, independent learning strategies, etc. On the other hand, “incidental learning is the process of learning something without the intention of doing so. It is also learning one thing while intending to learn another” (Richards & Schmidt, 2002, p. 276). The words are not presented in isolation but in context without any request for memorization.

There are many types of vocabulary to learn. Teachers should consider students’ necessities to define the priority in which words should be taught. High-frequency words are the focus of learning to achieve basic proficiency. However, pre-service English teachers also require to understand specialized and sophisticated words as these terms cover a good proportion of texts, thesis, and investigative reports. Consequently, “words in the academic vocabulary are useful for learning humanities, science or commerce” (Nation, 2001, p. 24). Low-frequency words are usually left apart during direct instruction and are mostly learned in context due to specialization or experience. Then, teachers should develop some strategies to learn this dense vocabulary in the classroom and generate practices to enhance remembering and use them during language production.

1.2.8 How to bring vocabulary into productive use

According to Nation (2001), “Written vocabulary can be increased by a general focus on vocabulary size and by focus on particular words for particular activities” (p. 281). It is necessary to dedicate time to pre-teaching the vocabulary we intend to transform into an active one. Another important element to succeed in vocabulary activation is motivation, the creation of opportunities to use a word. Paul Nation & James Coady (1988), analyzed some studies related to the effect of pre-teaching vocabulary on reading comprehension, highlighting some essential factors to achieve success in learning the meaning of an unknown word and using it in a sentence context.

If vocabulary instruction is to influence comprehension it must develop fluency of access to word meaning and must integrate the learned words into existing semantic networks, such instruction takes considerable time.

Word repetition helps some learners in increasing comprehension. Pre-teaching of vocabulary has added an effect of increasing the saliency of a word when it is met during reading. This improves recall of propositions in the text containing the instructed words. (p.100)

Then, the pre-taught words should be connected with previous vocabulary already controlled by the learner which could be the instruction of more frequent words' synonyms and academic words. Duin and Graves (1987) examined the effects on vocabulary knowledge, use of vocabulary in writing, and writing performance of pre-teaching 13 words over a six-day period. "The vocabulary and writing group outperformed the traditional vocabulary teaching group" (p. 311). Then, a combination of vocabulary pre-teaching and writing activities could be a booster for learning.

1.2.9 Vocabulary learning strategies and activities

Nation (2001) defined a taxonomy of vocabulary learning strategies dividing them into planning, source, and processes categories. **Planning strategies** include the selection of words to be learned, the aspects of the word (form, meaning, pronunciation), choosing the strategies to learn, and planning repetition. **Source strategies** are those which facilitate getting information about the word, analyzing its parts, using it in context, consulting a referent in a dictionary, and using parallels in native and non-native language. **Processes strategies** include creating an understanding of the word through noticing, retrieving, and generating strategies. This classification divides strategies from easier to more complex in terms of cognition.

Table 3. Vocabulary Learning Strategies

General class of strategies	Types of strategies	Strategies selected by the researcher
Planning: Choosing what to focus on and when to focus on it.	<ul style="list-style-type: none"> ● Choosing words ● Choosing the aspects of word knowledge ● Choosing strategies ● Planning repetition 	
Sources: Finding information about words	<ul style="list-style-type: none"> ● Analyzing the word ● Using context ● Consulting a reference source in L1 or L2 ● Using parallels in L1 and L2 	<ul style="list-style-type: none"> ● Dictionary use ● Semantic mapping ● Guess meaning from context
Processes: Establishing Knowledge	<ul style="list-style-type: none"> ● Noticing: Give attention to an item. ● Retrieving: Recall previously met items ● Generating: Establish vocabulary knowledge about the item 	<p>Retrieving strategy</p> <ul style="list-style-type: none"> ● Reading and sentence completion ● Paraphrase <p>Generating strategy:</p> <ul style="list-style-type: none"> ● Reading like a writer ● Using written input to affect vocabulary use in writing

Note. Adapted from (2001). Learning vocabulary in another language (p. 353-354.) by I.S. Nation, 2001, Copyright 2000 by Cambridge University Press.

1.2.10 Learn vocabulary from reading

According to Hu and Nation (2000), the relationship between vocabulary knowledge and reading comprehension is complex and dynamic. “When learners begin to read, vocabulary supports their reading comprehension but when they have gained control of many of the skills of reading, it becomes a means of vocabulary growth” (p. 403). Students read texts to complete academic activities such as writing, summarizing, studying, researching, etc. These reading activities can be an opportunity for new vocabulary learning. For non-native EFL students reading can be employed to increase language proficiency in three ways;

Intensive reading where learners work through texts containing unfamiliar language features, but where learners are assisted by teacher guidance and discussion or by the use of dictionaries, glossaries, or elaboration in the text.

Extensive reading for language growth where learners read texts containing some unfamiliar language features but cope with these with only minor interruptions to understanding the message of the text. **Extensive reading for developing fluency in reading** skills where learners read texts containing no unknown language features and where reading should not be interrupted by the need to deal with unfamiliar language features. (Hu & Nation, 2000, p. 423)

1.2.11 Task-based learning for writing

Task-based Learning (henceforth TBL) is considered an approach for foreign language teaching that generates an opportunity to produce the language in a natural context environment. TBL methodology reinforces the completion of tasks to learn the language. Prabhu (1987) defined a task as “an activity which required learners to arrive at an outcome from given information through some process of thought and which allowed teachers to control and regulate that process” (p. 24).

Academic writing can adopt this approach to assemble the multiple skills required in the style into a task that includes brainstorming, planning, discussing, and

organizing ideas before the writing cycle development. This view of writing as a process “reflects academic practice more authentically in the sense that writing takes time to produce because it requires thinking and drafting before it can be presented as a final, public product.” (Esfandiari et al., 2012, p.15)

The TBL framework proposed by Willis (1996) offers flexibility in application. The lesson can be divided into three phases: **Pre-task**, task cycle, and Language focus. During the pre-task, the teacher introduces the topic and explains the task. Students can solve an exercise that might prepare them for the task or experience a model for the task elaboration. The instructor monitors this stage and offers help. The **task cycle** begins with the task in which students can communicate. So, it is recommended to develop it in pairs to increase motivation. Planning is also part of the task cycle, students start drafting, planning, outlining, writing, and revising their reports. The teacher can provide advice during this phase about specific language items and encourage clarity, organization, and accuracy. Peer-edition and use of dictionaries are also allowed during writing tasks. Finally, **the report** closes the cycle with a presentation of the task, it could be an oral reading or a written composition. It is not necessary that all the students present, however, the teacher might offer suggestions, make comparisons or ask questions in general terms. **Language Focus** stage encourages the analysis of forms, grammatical structures, or clarification of meaning to explain concepts. There could be also a practice, sentence completion, games, or other reinforcement.

1.3 Foundation of the state of the art

The analysis of language through computational linguistic research dates back to the fifties. Since then, many authors have used it to explore various theories and corpus techniques to infer regularities in language, later described as new theories or norms. In the field, some studies are increasing the attention of researchers for the improvement of writing assessment. It is possible to mention Scott Crossley’s work, who researched second language writing utilizing computational tools and machine learning techniques. In this study, he employed natural language processing tools to analyze longitudinal writing development, to predict essay quality, differentiate between L1 and L2 writers, to provide a deeper understanding

of how second language writing is developed and the elements that affect production and the quality of writing with a functional view. For the pedagogical area, the author suggests employing technological tools for intelligent writing feedback as they offer information for textual and corpus discourse about the linguistic properties that affect text comprehension and production processes. (Crossley, 2013, p. 163-171)

According to Gregori-Signes and Clavel-Arroitia, written products are a central part of the assessment of linguistic competence. They studied lexical density and diversity in texts produced by Spanish native speakers with a C2 level of English employing technological tools such as Textalyser and RANGE. They suggest that quantitative and qualitative measures of lexical richness might provide an accurate picture of a students' lexical progress and their assessment. Lexical Frequency Profile in their study correlated well with other independent measures of vocabulary size and helped teachers diagnose the use of low and high-frequency vocabulary in texts from pupils and use these results to improve pedagogic instruments according to learners' needs. Teachers should employ this methodology to develop awareness about the material design, learner assessment and to detect progress in vocabulary learning. (Gregori-Signes & Clavel-Arroitia, 2015, p. 555)

Dakhi and Hutabarat in 2018 researched about language effectiveness and factors influencing scientific writing of Indonesian undergraduate thesis. This study reported lexical density and grammatical complexity of forty-two academic writing samples. The lexical density and grammatical complexity were different within each text, and the products were lexically acceptable but required to enhance grammatical complexity. The conclusions explained that some variables are affecting academic writing. Not only linguistic factors are crucial but also psychological, experiential, and sociocultural. (Dakhi & Hutabarat, 2018, p. 61) Therefore, a computational analysis should be combined with holistic features from a text's human semantic analysis to achieve a complete vision. For pedagogical improvement, it is recommended to create a product-based approach to writing which should be well-defined, planned, and conducted to develop thesis products with a scientific level.

Vuković Stamatović, Bratić, and Lakić analyzed the lexical level, lexical variation, and academic vocabulary of graduation theses from University of Montenegro students and compared the same characteristics from theses written by native speakers of American English. The abstract of their research summarizes that:

Montenegrin theses are readable at 4,000 words; it means that B2 learners can read them reasonably, while native speakers' theses can be read at 7,000 words only by high-level readers (C level). Students from the University of Montenegro displayed a sufficient vocabulary size and underuse of academic vocabulary. (Vuković Stamatović et al., 2020, p.1)

This result suggested that academic word use is significant for the quality of writing in research products and that students can produce readable pieces. The pedagogical implications of this study are that Montenegrin students should be motivated to master and use more academic vocabulary and to increase the exposure to any type of vocabulary such as undergraduate seminars and graduation papers to become more used to academic vocabulary and good models of academic writing. It is also suggested to employ the academic vocabulary list to use the time efficiently concerning the large coverage that these words provide in academic texts.

William Gyasi applied a corpus linguistic methodology to analyze the readability of different handbooks of three Ghanaian universities by employing classic readability indexes and automated tools. The study reflected that these three handbooks were very difficult to read in some cases, they were written for being understood by university graduates and above. There is variation in the measuring of readability from index to index requiring extra research to understand better the origin of these changes. The pedagogical suggestion is to employ new handbooks taking into account student's reading level, considering that readability is only one element in text selection and aiming to provide useful and functional information to the students.

Readability is the high importance for academic writing as the intended is to create products with an impact on the reader and the researcher's community. Ilina

Doykova developed in 2016 a study in a medical corpus of research articles published in international journals by non-native English users. It was focused on analyzing sentence structure, the accuracy of the message, and the economy of words in the published papers. The findings include complex sentences, use of modifiers, excessive use of general vocabulary, inappropriate word choice, and verbal excess. The pedagogical suggestions for the authors are to improve writing skills, academic language revision, through increasing training in textual collocations to reduce the complexity of the sentence and avoiding direct translation from the native language.

1.4 Chapter I Conclusions

- This chapter synthesizes fundamental concepts about lexical richness and readability, parameters for methodology development including the description of the diagnostic instrument, and tools.
- Background research exhibited the development of readability and lexical richness research in computational linguistics around the world revealing that the present research fulfills a research gap in the national context. These works showed interesting results that inspired procedures and possible results to be expected in this investigation.
- There were explained the leading theories on which this investigation is based for the proposal's construction offering a clear idea about how students learn vocabulary and the variety of strategies to be selected to provide instruction.

CHAPTER II PROPOSAL

2.1 Proposal topic

Booklet for academic writing improvement based on lexical and readability analysis of written corpus.

2.2 Objectives

General Objective:

To design a booklet for pre-service English teachers to improve academic writing in terms of vocabulary use and ease of reading through tasks development.

Specific Objectives:

- To identify which lexical and grammatical items require attention to increase the readability of pre-service teachers' writing.
- To provide a reference guide for academic style improvement in terms of vocabulary.
- To train students in vocabulary learning strategies to increase lexical richness in written products.
- To offer writing tasks to bring vocabulary knowledge into productive use.

2.3 Justification

The diagnosis was executed through a Vocabulary Level Test (Mc Lean & Kramer, 2015) applied to the 39 students studying the module of Academic Writing at Technical University of Cotopaxi, using the application Google forms to estimate the total number of words that pre-service English teachers know. This information reveals the students' level of proficiency and the level of mastering the most frequent lexical items in English that will be compared with their lexical repertoire. Students had 60 minutes to complete the multiple-choice test that contains 150

items with one correct answer and three distractors. The app instantaneously closes the form when the time ends. Two students did not complete the test during the established period, and the application erased the answers. For this reason, 37 answers were shown in **Appendix 2**. According to the Common European Framework of Reference, the Technical University of Cotopaxi curriculum sets that pre-service English teachers in the seventh semester of their career should master a B1.2 level of English. As this diagnosis is focused on vocabulary size, there would be an analysis and comparison about the number of correct answers obtained in the test and their English level based on previous research related to vocabulary coverage in academic texts.

To analyze the results from the NVLT, the writer considered Coxhead's study (2000), which described the coverage of words in an academic corpus written in English. Pre-service teachers will require enough receptive vocabulary to read these kinds of books. For that reason, their language proficiency is linked to their vocabulary knowledge in certain bands of frequency. In academic texts, the coverage of the first 1,000 more frequent words is 71.4%, the second 1,000 more frequent words cover an extra 4.7%, the AWL cover 10% of the text. The rest of the vocabulary will cover the last 13.9%. Therefore, according to the results achieved on each frequency band based on the previous percentages, there is the sum of total vocabulary for each pre-service teacher detailed in *Appendix 2*.

“Usually, the 2,000-word level has been set as the most suitable limit for high-frequency words that cover a little more than the 80% of the running word in academic texts” (Nation, 2001, p .22). Then, non-native learners with less than 1000-word families vocabulary plus the number of right answers in the AWL items (less than 71.40% in the NVLT) might belong to the A1 level. These learners would not be able to understand an academic text completely since their vocabulary covers less than 80%, and that is not enough to reach comprehension, according to Hu and Nation (2000). Learners who achieved less than 2000-word families plus the academic word list items (less than 86.10 but more than 71.39 in the NVLT) fit the A2 level because they would have a better understanding of academic texts as AWL increases 570-word families to their vocabulary and 10% of coverage in this type of corpus. “Nation & Wang presented evidence that graded readers schemes need

to go up to the 5,000-word level to make easier the transition from graded readers to unsimplified text easier that provide coverage of the Academic word list” (1999, p. 355).

Consequently, a student that knows less than 5,000-word families and more than 3,000 plus the AWL (86.10% to 97.99% in the NVLT) may be situated in a B1 level as they will develop meaning in reading through focused input and achieve a 95% of coverage of academic text achieving acceptable comprehension. It is also possible to infer that a non-native student who knows more than 5,000-word families, including the AWL (98% to 100% in the NVLT), might be located in a B2 level of English as they will know the required high-frequency and academic vocabulary that covers the at least 98% of academic texts and allows adequate comprehension of them. Nation mentioned that “highly educated non-native speakers of English who are studying advanced degrees through the medium of English indicate a receptive vocabulary size of around 8,000 to 9,000 word-families” (2006, p .60). So, this number of families might be considered the number of word-families that a C2 and C1 level student should know in the CEFR. Though the output profile for pre-service English teachers specified in the curriculum is B2, the table does not specify ranges for these levels.

Table 4. Pre-service teachers CEFR English Level

CEFR English Level and vocabulary size		NVLT Result	N° Pre-service teachers	Percentage
A1	1000	0% - 71,40%	2	5,41%
A2	2000+AWL	71,41% - 86,09%	7	18,92%
B1	>2000+AWL	86,10% - 97,99%	24	64,86%
B2	5000+AWL	98% - 100%	4	10,81%
Total			37	100%

Elaborated by: Nataly Romero

Source: NVLT results

The results of the vocabulary level test exhibited that 24 students have a B1 level. This number of students constitutes 64.86% of the total, indicating that these pre-service English teachers are expected according to the university's curriculum planning. They might understand the meaning of most of the words that commonly cover academic texts. The 10.81% of the students are in a higher B2 level, which is the pre-service English teachers' outcome profile. They might understand almost all the words that commonly cover academic texts and guess unknown words due to inference. These students could achieve even higher levels in the following years of training, improving the outcome profile for the career and their language skills. None of the students obtained the highest score in the Vocabulary Level Test, meaning that they will require instruction to learn academic and specialized lexicon.

18.92% of the students belong to A2 level of vocabulary about the established parameters, and 5.41% of pre-service English teachers. Together, these percentages constitute 24.33% of the student sample which reflects basic vocabulary knowledge that might affect their written production. Although the first 1000 and 2000 more frequent word families contain the basis of language, teachers require specialized and less frequent kinds of vocabulary to write, read and perform successfully in an academic context. Morris & Cobb studied the vocabulary profile of TESL trainees. They found that "trainees with the best command of AWL words and the richest expressive vocabularies perform better in the representation of grammatical knowledge and oral presentation to a target audience" (2004, p.87).

Therefore, it is compulsory to find strategies to increase the collection of the most frequent lexical items and academic vocabulary of pre-service teachers to improve their writing and reading comprehension skills. These strategies should focus on students who demonstrated lower knowledge in the VLT, so they would understand words with a higher proportion of coverage in texts and reach the lexical proficiency required to become English teachers. Many academic word list items are related to Spanish cognates; however, none of the students achieved the maximum grade in this part of the vocabulary level test. Consequently, students might have difficulties with reading comprehension and might require extra training in using the academic word list (**See appendix 2**).

The analysis of lexical richness considers lexical density, lexical diversity, and lexical sophistication in the corpus (**See appendix 5**). Authors like Ure (1971) and Halliday (1989) have found that lexical density is higher in writing than in speaking. However, the lexical density of the language sample represented in the corpus was 0.52. Demonstrating that there is a low number of content words in the corpus to provide meaning and information. The analyzed texts present a high quantity of function words that are not as descriptive and meaningful as academic writing requires. According to the classifications proposed by Syarif & Putri, the corpus is less dense. (2008, p .20) This result suggests that pre-service teachers might have vocabulary limitations that should be reinforced with lexical teaching. Comparing the results in the VLT with the vocabulary profile, it is possible to infer that in the written expression, they remain as intermediate producers of the language.

The readability analysis of the corpus developed with COH-METRIX shows that it has a high level of difficulty. “Written is usually more complex than spoken language” (Halliday, 1989, p.29-45). Intricacy in academic texts regularly derives from employing less-frequent words, dense vocabulary, and lexical diversity. However, the source of difficulty of the studied corpus written by pre-service English teachers originates in excessive usage of abstract words, complex sentence structures, and lack of explicit links that offer connectivity to reflect contrast and adversativity. These inferences match with the LFP investigation as it demonstrated a high quantity of function words and low sophistication (**See appendix 4**). In this way, vocabulary accuracy also requires enhancement in selecting appropriated lexical items according to the style to increase lexical density and diversity in the corpus, improve the use of academic words, develop awareness about sentence length, enhance the use of additive, adversative, and comparative connectors.

2.4 Proposal development

It is said that many teachers employ corpus analysis to diagnose patterns and linguistic features from their student’s products. This tool offers answers for multiple language teaching questions and provides evidence to implement strategies to solve their problems and validate or propose learning theories. After analyzing the written corpus, it was identified that pre-service English teachers require to

increase lexical richness in the text by using synonyms of words, employ different content words (AWL) that provide meaning and clarification, and more sophisticated words to evoke the characteristics of academic writing. Their texts also require to increase their readability through the counting of word length, employing grammatical and lexical items that help increasing reading ease such as conjunctions, linking, and comparison terms. These terms are not in their productive lexical repertoire, for this reason, it is also necessary to instruct them in how to acquire this less-frequent, academic and grammatical vocabulary that increases the quality of texts. The instruction will require vocabulary learning strategies to fulfill this requirement of vocabulary growth.

2.4.1 Proposal components

The proposal is constituted by theoretical and practical components. The theoretical component includes an explanation of the approach to use this pedagogical resource, the writing process, strategies, and planned activities. The practical component of the booklet collects eight main topics related to the enhancement of lexical richness and readability in academic writing. Each topic collects a pre-task and a task to develop TBL learning model. The selection of the thematic for each unit was based on the results of the previous corpus analysis. Each unit specifies the learning outcome expected to obtain with the completion of the tasks. The booklet includes an explanation for each topic and content related to the linguistics field to develop a controlled practice. In this way, pre-service English teachers will strengthen useful concepts for their professional development. The material presents tasks to bring vocabulary to production. It also suggests vocabulary learning strategies that students can apply to learn new academic words and promote the acquisition of new, less-frequent vocabulary. The final evaluation will be developed with a productive activity that the teacher must assess to offer feedback.

2.4.2 Proposal explanation

The present proposal aims to support academic writing development by enhancing vocabulary to increase the readability of the text. Deep analysis of corpus allows a better understanding of students' necessities beyond a vocabulary level test. The

vocabulary-level test measures which is the vocabulary size of the students allowing establishing a parameter for comparison. A corpus linguistics analysis increases comprehension about how they use language and their writing's readability. This booklet might solve writing weaknesses because it offers strategies and examples of activities to overcome those lexical and grammatical limitations present in the corpus. It extends advice that can be applied to complete different kinds of academic writing tasks and examples of practices to reinforce the knowledge acquired in academic writing lessons. The Academic Style unit seeks to highlight the most common errors committed during the writing and which actions might solve these inconveniences that reduce the quality of the text.

There are some units focused on lexical richness improvement, for instance, the unit about preventing redundancy and repetition was designed to increase lexical diversity in the text. It was found that students repeat the same terminology instead of employing less frequent terms that might help to improve sophistication and diversity. The Academic Vocabulary unit proposes lexical diversification and also provides content words that are connected to the academic setting that contribute to acquiring academic style referring to the lexicon. How to increase less frequent and academic vocabulary unit explains three different vocabulary learning strategies that can be applied by students with teacher support aiming to transform receptive vocabulary into active through writing activities.

Other units are meant to increase readability in the text, connective words unit was developed to help students remembering conjunctions use. Conjunctions are a cohesive device that allows the reader to establish meaningful relationships between the ideas in the text. Therefore, these connectors help to make the text more readable. The proposed activities encourage identifying and using these words to connect sentences and improve the information in the text. The comparison unit helps to recall words to establish contrast relationships in the text. There was a low level of this terminology which increased the difficulty of the text as they help to clarify opposition meaning. Syntactic simplicity is related to sentence length. This unit intends to aware students of paragraph composition to produce different length sentences increasing the readability of the text and creating a good effect on the reader. The nominalization unit offers a warning about how to use concrete and

abstract words. Concrete words increase readability while abstract words reduce it. However, nominalization is a common practice in academic texts that requires to be exercised by young writers to clarify their written ideas.

Vocabulary learning strategies are an extended process that should be shared between teachers and learners. For this reason, this proposal explains how students can intentionally continue practicing and learning vocabulary outside the classroom. Active strategy users achieve larger vocabulary size and proficiency than those who do not use them or only try memorizing words. So, strategies foster meaningful knowledge in students despite their age or the language they are learning (Nation, 2001). The booklet's activities solution encourages higher-order thinking skills like synthesizing, analyzing, reasoning, applying, and evaluating. Therefore, students will use critical thinking to simultaneously learn new words, academic style features, and thinking in a second language. "Vygotsky believed that, if situations were designed to allow students to use their critical thinking skills, new knowledge would be acquired" (Swaran & Marappan, 2020, p. 741). This kind of training can achieve more effective long-term results in learning unfamiliar terms than direct study.

The proposal includes extracts of linguistics texts, different academic words such as nouns, verbs, and adjectives. They might be incidentally learned while reading or in the activities' solving process this learning is encouraged by a guess from the context strategy that will be applied as a constant during the proposal's implementation. "Incidental vocabulary is only one of the various kinds of learning that can occur when learners read. Not only can they begin to learn new words and enrich known ones, but they can also improve grammatical knowledge, become more familiar with text structure, improve reading skills, learn new information, and learn that reading can be an enjoyable activity" (Nation, 2001, p. 374).

2.4.3 Premises for its implementation

Something to consider when talking about improving writing is that it requires time and practice. "From a vocabulary perspective, second language learners might require about to two years to gain control of the two thousand high-frequency general service words, and three to five years more to gain control of the academic

vocabulary and other relevant low frequency and technical words” (Nation, 2001, p. 190). It is hopeless that the proposal’s implementation accomplishes excellent results quickly or after just a few lessons. For this reason, the author suggests the socialization of the booklet to the users, which was developed through an online workshop. Subsequently, the proposal might be applied in future research development on this investigative line to verify its accuracy on vocabulary growth by employing the same methodology.

The proposal’s implementation was designed for an eight-week program in which the teacher should cover each week the content explanation of a topic from the booklet and the student at the same time should apply a vocabulary learning strategy. The first part of the proposal includes the application of TBL approach employing the booklet. The pre-task includes the introduction to the topic and examples or models of vocabulary use that might be analyzed in pairs during the class. Students as active learners should complete the writing task to create the product for a presentation, this activity might be developed in pairs or individually according to the learner’s will to encourage communication during the execution of the task. The instructor also should assign time for a post-task stage to provide feedback and peer-review opportunities.

Vocabulary learning strategies should be applied to learn the words from the lists detailed in **appendix 11** each week. These lists include academic words that the students did not use in the corpus which are considered less frequent. The lists were built with Lex-tutor Vocab profile program. Students must exercise the indicated strategy using the time to learn 2 words per day with the semantic mapping strategy (8 words per week), and 5 words per day with the use of the dictionary strategy (20 words per week). The evaluation of this component will be developed once a week after finishing the TBL stages. It will consist of writing a sentence employing in context a word from the lists of study. There would be a random selection of terms to be evaluated each week.

Guessing from context strategy should be instructed by the teacher at the beginning of the implementation and applied during the eight weeks. As incidental vocabulary is difficult to measure this process will be self-assessed by the student with the

instrument provided in the booklet. The guessing of words will be complemented with the use of the dictionary to amplify the possibilities of recalling the terms. Students should write each time they guess the meaning of a word while reading and confirm if the guess was right at the end of the week by revising the definition in a dictionary.

The teacher will spend on week 0, two extra hours to introduce the booklet use, explain some features of academic writing, and instruct about the use of vocabulary learning strategies. After this, they should invest two hours in class per week, an hour for instruction, and another for the post-task stage and vocabulary evaluation. This might complete a total of 16 hours. Students must dedicate four hours for vocabulary learning strategies application and one hour to complete the writing task presented in the booklet for a total of 40 hours. The final evaluation should take two weeks to be completed to develop the writing process and offer support in the classroom. The following chart reflects the planning of the proposal.

Table 5. Timetable for applying the proposal

Period	Booklet Topic	Students' activity			Teacher Control	
		Strategy	Words to be learned	Task to be developed	Vocabulary Evaluation	Feedback to the task and practice
Week 0	Academic Style	Strategies explanation - Semantic Mapping	List 1	Task 1		
Week 1	Connective Words	Semantic mapping	List 2	Task 2	Sentences using words in context List 1	Academic Style task

Week 2	Comparison Words	Semantic mapping	List 3	Task 3	Sentences using words in context List 2	Connective Words task
Week 3	Preventing redundancy	Using dictionaries	List 4	Task 4	Sentences using words in context List 3	Comparison Words task
Week 4	Sentence length	Using dictionaries	List 5	Task 5	Sentences using words in context List 4	Preventing redundancy task
Week 5	Academic Vocabulary	Using dictionaries	List 6	Task 6	Sentences using words in context List 5	Sentence length task
Week 6	Nominalization				Sentences using words in context List 6	Academic Vocabulary task
Weeks 7-8	Evaluation: Research Proposal					Nominalization task Peer-review Research Proposal Draft
Weeks 1-8	Guessing from context will be applied during academic activities. Self-assessed by students employing the chart in the booklet. (Identify guessed word, at the end of the week verify correct guesses, report number of right guesses to the instructor)					

Elaborated by: Nataly Romero

The process for vocabulary learning strategies implementation is also described in Nation (2001):

The teacher models the strategy for the learners in the classroom. The steps in the strategy are practiced separately. Learners report back on the application of the steps in the strategy. Learners report on their difficulties and successes in using the strategy when they use it outside class time. Teachers systematically test learners on strategy use and give them feedback. Learners consult the teacher on their use of the strategy, seeking advice where necessary. (p. 359)

Teacher's Role. - The critical activity of the teacher is to raise awareness about the importance and value of exercising vocabulary learning strategies outside the classroom. Teachers take precious time from the primary educative process to instruct the topics in the booklet by applying a TBL approach and to clarify the process for applying the use of the dictionary and semantic mapping strategies. Task-based learning requires extra effort from the teacher to be updated in the target vocabulary to be taught, to manage the time and inconveniences that might occur in any of the stages of the training. The trainer also requires to generate activities that activate the production of academic texts and provide periodic feedback. It is also important to maintain control about how many students employed the learned terms according to the context of the sentences produced in the weekly evaluation.

Learners' Role. - The student is an active element in the learning process. In this proposal, the increasing vocabulary and quality of writing will depend on the student's compromise and motivation to continue learning. "For each of the strategies like guessing from context, using words parts, dictionary use or direct learning, learners need to spend a total of at least four or five hours per strategy over several weeks" (Nation, 2001, p. 358). All this time can be completed as a student often spends many hours reading, writing, and investigating their assignments each week.

Vocabulary in Academic Writing

Introduction

Academic writing requires extra effort from the writer to read, summarize, analyze and connect the information in a formal product. The writer's thoughts and knowledge follow certain features given by the genre, but these alignments do not apply to all research disciplines. A fundamental element in all kinds of writing is effective vocabulary use. Tovar Viera R. states that "Vocabulary knowledge of foreign language is necessary; it provides learners a broader ability to produce well-structured written texts and contributes to the comprehension of utterances as well" (2017, p.89). Quality writing entails proper vocabulary but also an adequate structure use to facilitate reading. Although students receive training in academic writing when they start, the implementation of rules and conveying ideas at the same time complicate reaching an effective style. For this reason, this booklet presents some tips and practices that pre-service teachers can apply after class aiming to develop good habits during composition tasks.

General Objective: To improve pre-service English teachers' writing through the application of academic style features and vocabulary learning strategies.

Specific Objectives:

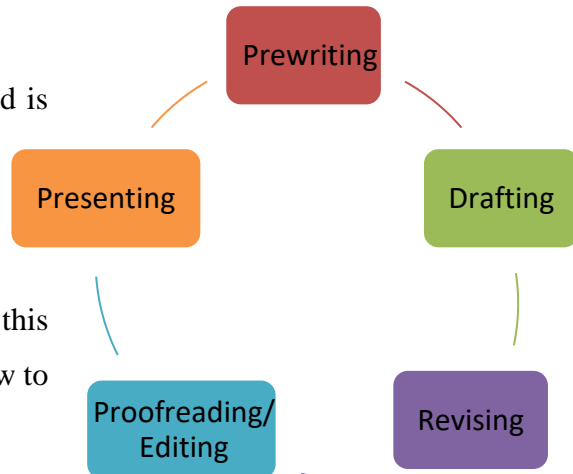
- To identify cores aspects of academic style
- To overview strategies that enhance academic vocabulary learning.
- To provide writing activities encouraging lexical richness and readability development in texts.

Approach

This pedagogical aid follows a Task-Based Learning approach (henceforth TBL) since it is learner-centered. Students are responsible for their own learning which is promoted as a result of tasks completion. Learners require knowledge about writing academically because as teachers they will be immersed in research, academic and educative fields that require experience in completing writing tasks. “TBL is then suggested as a suitable approach to the teaching of academic writing because of the humanistic educational principles that underlie its pedagogy” (Esfandiari et al., 2012, p.2). TBL also demonstrated being effective for teaching specialized vocabulary. Sarani & Farzaneh Sahebi (2012) studied the impact of TBL on vocabulary learning in English for specific purposes. “In pre-task phase, the researcher tried to activate the ESL learner’s schemata related to the text with new technical vocabularies to motivate them to read. In the during task phase, the students were engaged in completing different kinds of tasks, and in the post-task phase, they gave a report, repeated the tasks, and even dealt with language focus tasks” (p. 121). They found that students who have been taught vocabulary through this approach outperformed those who followed the traditional one.

The writing process

The objective of this pedagogical aid is to increase students' vocabulary to produce a positive effect on writing quality. The booklet contains writing tasks for each unit, for this reason, it is imperative to explain how to develop the writing process.



- 1. Pre-writing** consists of planning what is going to be written. Collect information about the topic, brainstorm ideas, participate in a small discussion with a classmate about the topic and employ graphic organizers to organize your notes.
- 2. Drafting** is the beginning of writing, use the structure and information planned to develop ideas. It does not matter if the sentence is not perfect, do not lose the main idea. This step might take more time than expected, this is the creative process.
- 3. Revising** can be a shared step. It includes thinking about the person you wrote for and if your text is going to be understandable. This booklet offers good advice about using some words to make your texts more readable. It is possible to ask a friend to revise your writing, remember to keep it clear and nice for the reader. If it is a formal composition be very careful.
- 4. Editing** requires being focused and remember grammatical rules, spelling, punctuation, sentence structure, and the most important part; word choice. Try to change all the elements which are not standard English.
- 5. Presenting or Publishing** is the objective of any writer. If you write you should publish to share all your effort and work. In this case, present the written task to your teacher. An expert might offer you feedback to improve your product. The best part is that process starts again and you have learned many new tips for continuing writing.

Timetable

The following chart shows which vocabulary learning strategies you should apply each week and the number of words for daily learning. The time required to apply the strategy will be one hour per day for 4 days. Evaluation of this learning will be developed once a week. The strategies should be registered in sheets of paper or it is possible to reproduce the tables proposed in unit 8 to be printed. The writing tasks should be developed in pairs out of the classroom but presented each week to receive feedback from the instructor or to organize peer-reviewing.

Students' activity				
	Strategy	Words to be learned	Time	Task to be developed
Week 0	Strategies explanation - Semantic Mapping	List 1 (8 words)	Apply the strategy for an hour. Four times a week. (2 words each day)	Task 1
Week 1	Semantic mapping	List 2 (8 words)	Apply the strategy for an hour. Four times a week. (2 words each day)	Task 2
Week 2	Semantic mapping	List 3 (8 words)	Apply the strategy for an hour. Four times a week. (2 words each day)	Task 3
Week 3	Using dictionaries	List 4 (20 words)	Apply the strategy for an hour. Four times a week. (5 words each day)	Task 4
Week 4	Using dictionaries	List 5 (20 words)	Apply the strategy for an hour. Four times a week. (5 words each day)	Task 5
Week 5	Using dictionaries	List 6 (20 words)	Apply the strategy for an hour. Four times a week. (5 words each day)	Task 6
Week 6				Evaluation: Task 7
Weeks 7- 8				
Weeks 1-8	Guessing from context			

Strategies and activities developed with the use of the booklet

Vocabulary is vital for language learning. This booklet proposes the application of different strategies in different levels to learn academic and less frequent vocabulary. Source strategies will require time from the student to be developed each day while generating and retrieving strategies will be exercised during the completion of pre-tasks planned in the booklet. Here are explained the different strategies that students will exercise:

Source Strategies

These strategies are focused on learning the form of a word by analyzing sources that provide information about the term, for example, how it is written or pronounced. Some examples are:

- Dictionary use

Learners need to be trained in dictionary use so that they can readily find words that they need in their writing. The length of an entry was seen as the major challenge in finding needed information about a word.

- Semantic mapping

The learners work to develop a semantic map around a term. The teacher deliberately introduces several target vocabulary items, learners research the form and meaning of the word. The learners then use the semantic map to do a piece of writing. (Nation, 2001, p. 284-285)

- Guessing meaning from context strategy

Learners guess the meaning of an unknown word by clues provided from the context such as parts of speech, relationship with known words, the position of it in the sentence, etc.

Retrieving Strategy

These strategies benefit the learner to recall previously learned words which prevent forgetting them. The recalling should be oral or written but without checking the original source.

- Reading and sentence completion

There are several varieties of completion activities that can follow a reading text and use words that occur in the text. The completions can range from copying from the text to having to use the words with a different inflection or derivational affix or to express an idea, not in the text.

- Paraphrase

The learners read sentences that they then have to re-express using the target word which is provided for them. The teacher will need to model the use of the word first or provide some example sentences. (Nation,2001, p. 283-284)

Generating Strategies

These strategies entail making vocabulary knowledge be remembered by the student. The process helps the brain to transform new words into productive vocabulary. For example:

- Reading like a writer

Learners work through a reading text noting features of the text that typify that style of writing. From a vocabulary perspective, these features can include the degree of formality of the vocabulary, the use of lexical chains, lexical cohesion through the use of related words, and signals of changes in

the stages of the text. The learners are encouraged to use some of the features in a writing task.

- Using written input to affect vocabulary use in writing

This activity includes providing plenty of written input to the task, designing the task to make use of the written input, and using recall, and adaptation of the input to encourage generative use”. (Nation,2001, p. 283-284)

1. Academic Style

Learning outcome: Create sentences with appropriate style and recognize those that require improvement.

Academic writing requires a lot of practice, constancy, and discipline to be perfected. It presents some regularities that allow achieving the expected style and form easily. Check these features the next time you have to write papers, essays, or research reports. Nevertheless, remember that these rules can be broken on some occasions depending on the text’s purpose.

Do

- ✓ Be objective and descriptive.
- ✓ Employ formal register.
- ✓ Write clear ideas.
- ✓ Employ academic vocabulary.
- ✓ Be consistent.
- ✓ Combine sentence structures.
- ✓ Cite sources.

Don't

- ✗ Be emotive or personal.
- ✗ Use colloquialism or offensive language.
- ✗ Wordiness.
- ✗ Use shortened words.
- ✗ Use different fonts or numeration
- ✗ Write very long sentences without connectors.
- ✗ Write other authors’ ideas without properly citing and reference.

Pre-task 1

Strategy: Retrieving-Paraphrase

Identify which of the following sentences have poor academic style and rewrite them appropriately.

- a. When we think about sociolinguists, we think they study the relationship between language and society.
- b. It's ok to think that social context influences people's talk.
- c. Sociolinguistics concerns the relationship between language and context.
- d. Lots of people use different styles in different social contexts.
- e. In the 1960s, linguists started to research language use.
- f. Surprisingly, language provides information about the social relationship in a community.

Task 1

Strategy: Reading like a writer

Time: 1 hour

You are interested in assisting to a one-month training course about foreign language teaching. Discuss with your partner why this course would be useful for you since you finished University. Using these ideas write a short essay to obtain a scholarship to go. In your essay offer details of your academic background, say which is your motivation to assist, and explain how you will help society after returning from the training course. Write at least 150 words, remember academic writing features.

2. Connective words

Learning outcome: Apply connective words according to their function.

Conjunctions link clauses, paragraphs, or other words together. These words allow the flow of writing, the text becomes more informative and stimulating for the reader by preventing chopped sentences that interrupt the reading task. There are three forms of conjunctions; coordinating, correlative and subordinating.

Coordinating conjunctions: They join two words, phrases, or independent clauses (sentences that express a complete thought). We can employ a fun acronym to remember them:



Examples:

Linking words:

Speech sounds are classified into vowels **and** consonants.

Linking Phrases:

The research was irrefutably fascinating **yet** highly questionable.

Linking clauses:

All languages share features, **but** languages have only a limited range of sounds.

Writing Tip If you join two independent clauses together, use a comma before the conjunction. It is not necessary to use a comma if you link only **two** words or phrases.

- ✓ Learning a second language might be difficult, **but** it is not impossible to learn English.
- ✗ I have learned Russian, **and** French.

Correlative Conjunctions: They are sets of words or phrases that relate one element of a sentence with another. These elements have the same relevance or grammatical form.



Examples:

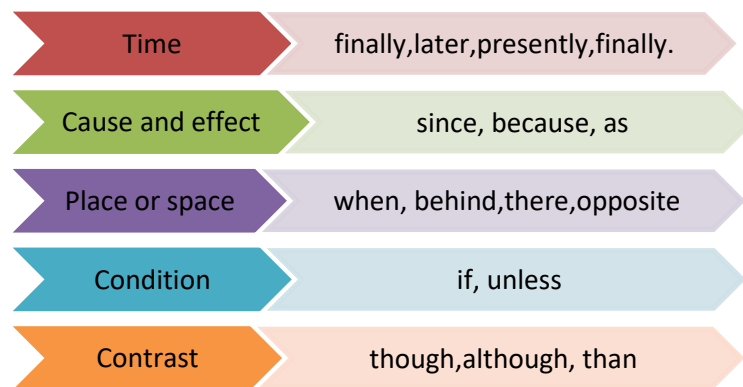
Vowels occur either alone or combined.

It is important not only where people are talking but also how they are feeling.

Writing
Tip

If you are using correlative conjunctions, you do not need commas.

Subordinating conjunctions: They are words to link clauses with a dependence relation. When a dependent clause comes first, it is necessary to add a comma before writing the independent clause. If the independent sentence comes first, the comma addition before writing the conditional sentence is unnecessary.



Examples:

Using information about the domain in a community is helpful because it summarizes the norms of language use.

People may select a particular code if they are discussing a particular topic.

If they are discussing a particular topic, people may select a particular code.

Writing
Tip

Avoid using conjunctions to start a sentence in academic writing.

❌ But people's speech provides clues about their social experiences.

Pre-task 2

Strategy: Retrieving- Reading and sentence completion

Read the following passage and underline the conjunctions. Then, classify them according to their type or function.

“A brief look at the dictionaries available now and in the past shows that lexicography is full of interdisciplinarity. Lexicography is not a sub-discipline of linguistics or lexicology, but rather a discipline in its own right with its own research object, namely the dictionary. Lexicography concerns the development of theoretical and practical principles as well as the production of lexicographical tools, so several disciplines are involved in any dictionary project. Almost all dictionaries contain words or terms from more than one discipline and subject field. Therefore, knowledge about and cooperation with specialists of, for instance, translation, copyediting, knowledge management, operative skills, and text production are necessary” (Fuertes-Olivera,2018, p. 102)

Conjunction	Type/ Function

Task 2

Strategy: Generating - Using written input to affect vocabulary use in writing

Time: 1 hour

Discuss with a friend possible causes and solutions about school dropout during the pandemic. Develop a brainstorm with your ideas and write an essay about this topic. Include the teacher’s role in this problem. Write at least 200 words, remember employing connective words to link ideas.

3. Comparison

Learning outcome: Produce comparative sentences and paragraphs to interpret charts.

Comparison is usually required in academic writing. It helps to understand, summarize and analyze; two or more topics, ideas, individuals, or characteristics. The organization and structures employed in the comparison help to clarify the written outcome. The use of comparative structures is not difficult to achieve if you follow these pieces of advice:

- ✓ Employ transition words and expressions.
- ✓ Structure the presentations of your ideas. You can mention all the similarities first. After that, you can write the differences or compare each aspect at a time.
- ✓ Use comparative and superlative grammatical structures.

Transition words to express difference

whereas, while, by contrast, on the other hand, conversely

Transition words to express similarity

Similarly, likewise, In the same way, both

Examples:

The purpose of an approach is to realize the goals of education, **whereas** the purpose of a method is to make effective presentations of subjects and content in the classroom.

A strategy differs from a method. The first term is new and belongs to educational technology, **while** a method is an old word related to pedagogy.

A strategy **compared with** a method is more flexible during application.

Freewriting is a strategy that encourages exploring ideas. By contrast, the debate is a strategy that develops oral presentation. **Both of them** require creativity and effort from students.

Pre-task 3

Strategy: Retrieving- Reading and sentence completion

Complete the following sentences by selecting one of the transition words to establish the proposed relation.

Sentence	Relation
1. Controlled practice is used to describe exercises that require a particular answer _____ free practice is used to allow the students to practice the language point in an unrestricted manner.	Difference
2. _____ skimming and scanning refer to visually reading a text.	Similarity
3. Skimming _____ scanning because it tries to get a feel for what the piece of text is about. _____ scanning is looking for specific information within the text.	Difference
4. Context of culture is related to genre _____ context of situation is related to register, and co-text to the discourse itself.	Difference
5. A morpheme is the smallest meaningful unit of a language. A phoneme, _____, is the smallest unit of speech.	Difference
6. The form _____ the meaning of a word are significant for vocabulary learning.	Similarity

Task 3

Strategy: Generating- Using written input to affect vocabulary use in writing

Time: 1 hour

Table 1. shows some characteristics from different languages. Analyze which language might be considered the most difficult to learn. Summarize the information by selecting and reporting the main features, and make comparisons where relevant. Remember employing transition words.

Table 1. Comparative chart of language characteristics

	English	Danish	Hindi	Arabic
No. of consonants	21	20	36	25
No. of Vowels	5	9	14	3
Numerals	10	10	10	10
Cursive style	No	No	Yes	Yes
Ligatures	No	No	Yes, a lot 504 variations of conjuncts	No/few
Diacritic marks	No	Not compulsory	Yes, and necessary	Yes, but few
Hyphens and other special characters	Yes	Yes, but seldom	No	Yes

Note. Retrieved from Analyzing cultural usability of mobile keypad and displays for textual communication in internationalization and localization perspectives by Orngreen et al., *Human Work Interaction Design: Usability in Social, Cultural and Organizational Contexts*, p.120. CC Copyright by Orngreen et al.

4. Preventing redundancy and repetition

Learning outcome: Identify redundancy and repetition to improve academic writing.

Redundancy is the reiteration of ideas using different sentences or words that do not offer extra meaning to the text. On the other hand, repetition implies using a word multiple times within the same sentence or paragraph. These practices are appreciated in poetry, but academic writing is better to use a diverse vocabulary generating more readable and informative products. Proofread your writing or ask

a friend to do it, so you can identify if there are recurrent terms or thoughts. Using synonyms is also a suitable option. Try to choose them from the academic word list and be careful that these words transmit what you mean. Be careful about employing two words that have the same meaning to prevent redundancy.

Examples:

- ✘ A phoneme is the smallest fragment of speech with meaning. The phoneme consists of some similar sounds. These similar sounds are also called allophones.
- ✔ A phoneme is the smallest fragment of speech with meaning. It consists of some similar sounds, also called allophones.
- ✘ In the year 1945, American varieties of English became more accepted or preferred over received pronunciation.
- ✔ In 1945, American varieties of English became more accepted or even preferred over received pronunciation.

Pre-task 4

Strategy: Retrieving-Paraphrase

Underline the redundant or unnecessary word in the sentence:

1. The 2000 more frequent words in English are sufficient enough for basic communication.
2. The conjunction is a word that joins parts of a sentence together.
3. Students dropped out of school because of the fact that it is necessary for them to support his family.
4. Homophones are words that sound exactly almost the same but are different in meaning or spelling.
5. Using Mobile Assisted Language Learning is considered as an advance forward from traditional approaches.

Task 4.1

Strategy: Generating-Reading like a writer

Time: 20 minutes

Read with a partner the following paragraph and discuss which words might be considered redundant or repetitive. Finally, rewrite together the paragraph and discuss the improvements.

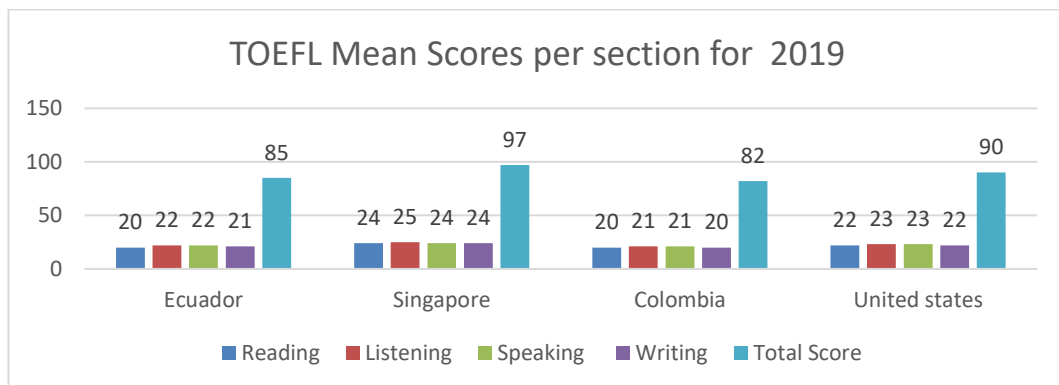
Misconceptions about language are very widespread, very often they derive from writing the language. For instance, some people think that English is more difficult to learn than other languages and they suggest developing a new innovative international alphabet to reduce confusion. However, writing is only a symbolization of language, then spelling systems can evolve over time.

Task 4.2

Strategy: Generating - Reading like a writer

Time: 40 minutes

Ask your classmates if they have taken or heard about international language tests. What did they hear about? The charts below show mean scores obtained per section in the Test of English as a Foreign language (TOEFL) in 2019 and the mean total score from some countries. Write a short report for a university lecturer describing the information shown below. Avoid redundancy or repetition. You should write at least 150 words.



Note. Retrieved from TOEFL iBT Test and Score Data Summary 2019

5. Sentence length

Learning outcome: Organize short and long sentences to convey relevant information in a paragraph.

Long sentences may be complicated to understand, while the recurrent use of short sentences creates chunks that are not pleasant for the reader either. The best idea is to combine them along the writing employing short sentences at the beginning of the paragraph to emphasize and state the main idea. Longer sentences might be utilized to give explanations or make comparisons.

Examples:

Short sentence (<15 words)

Long Sentence (<25 words)

"The nature of foreign accent is determined by a learner's native language. Thus, speakers of English are able to recognize Spanish accents, Russian accents, or Chinese accents. This ability indicates that the sound patterns of the native language are being transferred into the second language. In other words, there is often no one-to-one correspondence between the sounds people hear and the letters on a page."

(Avery & Ehrlich, 1992, p.1)

Writing
Tip

Sentences longer than 30 words are very difficult to read. Try to split them to increase your text's readability.

Pre-task 5

Strategy: Retrieving-Paraphrase

Rewrite the following paragraph to combine short and long sentences. You can use connectors to reorganize the ideas.

The pronunciation of grammatical endings differs depending on the sound that precedes them. One example of this is the past tense or plurals. Pronunciation of the grammatical ending is entirely predictable. Some rules describe how to pronounce suffixes. The pronunciation of sounds will vary depending on the

phonetic context in which they occur. These variations can be stated by a rule. Native speakers are not always aware of rules.

Task 5

Strategy: Generating - Reading like a writer

Time: 1 hour

You are completing an English proficiency test. The writing part asks to write a paragraph based to answer a question. Employ a combination of short and long sentences to offer an emphasis in your answer. Before writing, discuss with a classmate: Do you agree or disagree with the following statement? Use specific reasons and examples to support your answer, you should write at least 150 words.

It is more important that teachers dominate knowledge of the subject being taught over their ability to relate well with students.

6. Academic Vocabulary

Learning outcome: Substitute high-frequency words for academic and specialized vocabulary in writing activities.

Academic words are essential for academic writing. Not only for essays, research reports, or investigations but also for understanding while reading magazines and books and even watching programs on television. Therefore, learning, employing, and remembering these words can help you to improve your proficiency. The Academic Word List includes 570-word families, which are common in academic texts. The following exercises gave you an idea about their use in academic writing.

6.1 Academic Nouns

Nouns classify into concrete and abstract. Concrete nouns describe persons, animals, objects, places that are easier to represent in our minds because there is a physical reference, such as the words; woman, dog, book, London. Abstract nouns are concepts or ideas that require extra effort to be represented in our minds, for instance, the words creativity, strongness, opportunity.

Pre-task 6.1

Strategy: Retrieving- Reading and sentence completion

Underline the best noun to complete each sentence.

- a. Constructivism and Behaviorism are **models/theories/issues** of language learning.
- b. Evaluation is a broad **topic/scope/number** involved in the curriculum design process.
- c. Filling a timetable while listening to a conversation is a good **opportunity/model/aspect** to learn vocabulary and exercise listening skills.
- d. Sociolinguists mentioned that it is important to use authentic visuals to distinguish vocabulary from native and target **theme/culture/issue**.
- e. A common **nature/feature/theme** of research is the interaction between language and society.

6.2 Academic verbs

Academic writing requires formal verbs, not those that we regularly use during conversation. For this reason, you should elude utilizing two-word verbs; they are typically related to orality. If you intend to write an understandable and easy-to-follow investigation report, avoid phrasal verbs. They are difficult to understand for second language learners because they use suffixes which meaning is not always understood literally.

Students **pick up** new terms faster if they are young.

✓ Students learn new terms faster if they are

Writing
Tip

Avoid using phrasal verbs in academic texts.

Pick up as a phrasal verb means “to learn”, but it also means “to collect” or “lift”.

young.

Include verbs from Table 3. to enrich your writing. These words are not synonyms; their meanings are different. A helpful activity to understand their meaning is through reading academic texts and identify the possible contexts in which you can use them.

Table 2. Common Verbs in Academic Writing

Showing Change or difference	Increase: broaden, enlarge, exceed, expand, generate, improve, maximize, optimize
	Decrease: decline, deteriorate, erode, minimize, narrow, reduce, worsen
	Difference or varying: alter, contrast, convert, deviate, differ, differentiate, distinguish, diverge, evolve, modify, revise, transform
Shows stability	maintain sustain
Shows keeping within a certain range/keeping under a certain level	confine, inhibit, prohibit, restrict
Shows in-depth study	analyze, examine, investigate, observe, survey
Stating, Restating, or emphasizing ideas/ concepts	Stating: acknowledge, argue, attribute, comment, propose, establish, identify, mention, note, observe, state
	Restating: elaborate, expand
	Emphasizing: emphasize, stress
Describes a phenomenon or data	Describes phenomena: acquire, define, impact, signify, symbolize
	Describes data: approximate, demonstrate, indicate, levels off, reflect
Stating Position	Positive advocate, hold the view that, hypothesize, propose
	Negative/ Contradict: deny, dispute, negate, reject
Showing uncertainty or an extrapolation of information	Uncertainty: Predict, speculate
	Extrapolation of information: deduce, imply, infer, project
Shows components	comprise, consist, constitute, incorporate

Note: Retrieved from Verbs in Academic Writing, by Khoo, E., 2005, *Academic Vocabulary Series*, p. 1 – 2. Copyright 2005 by the writing center, University of Toronto at Scarborough.

Pre-task 6.2

Strategy: Source

Match the academic synonyms in the following list.

Nouns		Verbs	
area	advantage	accelerate	change
authority	part	achieve	help
behavior	argument	alter	question
beliefs	disadvantage	analyze	suggestion
benefit	tendency	assist	explain
category	field	attach	evolve
component	source	challenge	examine
concept	emotion	claim	establish
controversy	target	clarify	insist
drawback	explanation	concentrate on	speed up
expansion	conduct	confine	take apart
feeling	topic	develop	join
framework	possibility	eliminate	reach
goal	ethics	evaluate	decrease
hypothesis	production	found	demonstrate
interpretation	research	maintain	increase
issue	theory	predict	cite
method	increase	prohibit	reinforce
option	idea	quote	remove
quotation	citation	raise	focus on
results	figures	reduce	forecast
statistics	type	respond	ban
study	structure	retain	limit
trend	system	show	keep
output	findings	strengthen	reply

Note: Retrieved from Academic Writing A handbook for International Students by Bailey S., 2011, *Routledge, Third Ed., p.223*. Copyright 2011 by Stephen Bailey.

6.3 Academic adjectives

Adjectives increase the understanding of nouns. They give a broader description that benefits the reader to visualize them for comprehension. Some adjectives are mostly presented in writing with specific nouns. Table 2 includes a summary of common adjectives that you can use.

Table 3. Common adjectives in academic writing

Relating to:	Common Adjectives	Examples of use in Academic Writing
Importance	important/salient/ relevant necessary significant	significant relation necessary skill
Size/ amount / intensity / frequency	high/ increasing low/ declining adequate/ sufficient prime/ main/ primary / major/ only/ sole annual/ hourly / indiscriminate	increasing complexity declining interest primary consideration only restriction indiscriminate depletion annual evaluation
Quality	new/ innovative consistent sustainable abstract hierarchical	innovative strategy consistent representation sustainable solution abstract concept hierarchical organization
Variation	different/ alternative variable	alternative meanings variable context
Probability	likely/ possible sure/definite/ inevitable impossible	likely paradigm possible scenario inevitable outcome

Note: Adapted from Adjectives and adverbs in Academic Writing by Khoo, E.,2005, *Academic Vocabulary Series, p. 1*. Copyright 2005 by the writing center, University of Toronto at Scarborough.

Pre-task 6.3

Strategy: Source

- Graphic context for a word in a story is a **new/significant/high** factor for vocabulary learning.
- The Vocabulary Level Test does not include **alternative/likely/increasing** questions.
- The finding is not **consistent/major/inevitable** as some studies suggest that long-term learning is not associated with keyword technique.

- d. The research presents two **sustainable/sole/possible** scenarios, a positive and a negative.

Task 6

Strategy: Generating - Using written input to affect vocabulary use in writing

Time: 1 hour

Your teacher asked you to revise some paragraphs you wrote last week. The instructor suggests it should include more academic vocabulary and increase its formality. Re-write the paragraph to improve it. Finally, present your topic to the class orally and discuss the relevance of the research in pairs.

Many EFL students find it difficult to speak aloud in class or when the teacher asks them for their opinion. After a survey, many of them said that their problem is lack of confidence and lack of vocabulary. In both cases, novels, are great solutions for gaining confidence, but not alone. Some students said that their problem is that they cannot practice with confidence in class or with the teacher, although they cannot practice alone because they do not find it useful.

The proposal of the investigation is that if students practice reading with a friend of confidence who masters the language. During these sessions, both can read a book in which they are interested and while reading they can practice oral expression freely and at the same time. They can acquire vocabulary while reading. This can be an interesting activity for both, but especially for the student who enjoys the practice and speak aloud.

Summing up, oral expression can be intimidating for students because they do not perform enough practice on their own. The possible solution is that students work in pairs in short reading sessions, where one of them performs better the language and both can enjoy reading, acquiring new vocabulary about their own interests, and practicing oral expression.

7. Nominalization

Learning outcome: Employ nominalization to improve writing quality.

Nominalization is the process in which a verb converts into an abstract noun. This transformation sometimes decreases the quality of the text. So, be aware when using this practice in your academic papers.

✘ The researcher developed an investigation about the main factors that dismiss learning a second language in children.

The verb "developed" reduces the vigor of the verb "investigate," unnecessarily blurring the clarity of the sentence. Eliminate nominalization and simplify the idea.

✔ The researcher investigated the main factors that dismiss learning a second language in children.

If nominalization is correctly applied, it may increase formality, as in the following example:

✘ I analyzed the corpus, and it revealed that syntactic simplicity influenced the readability of the text.

✔ The corpus analysis revealed the influence of syntactic simplicity over texts' readability.

Pre-task 7

Strategy: Retrieving-Paraphrase

Rewrite the following sentences to nominalize the highlighted word.

- a. The number of people who learn English as a second language increased significantly in the 15th century.
- b. The teaching strategies applied revealed that the oral production had enhanced.
- c. Lexical sophistication is covering the texts irregularly.
- d. Lexical density in the text was increased by 5%. This caused a diminution of readability.

Task 7

Strategy: Generating - Using written input to affect vocabulary use in writing

Time: 1 hour

Use nouns to explain a teaching approach you are interested in. You can develop your own definitions, explain the framework or possibilities for its application with students. You can employ all the nouns listed below or you can choose only some of them to help you clarify your writing.

1. Identification, 2. Interpretation, 3. Specification, 4. Context, 5. Conception, 6. Significance, 7. Assumption, 8. Procedure, 9. Assessment, 10. Approach

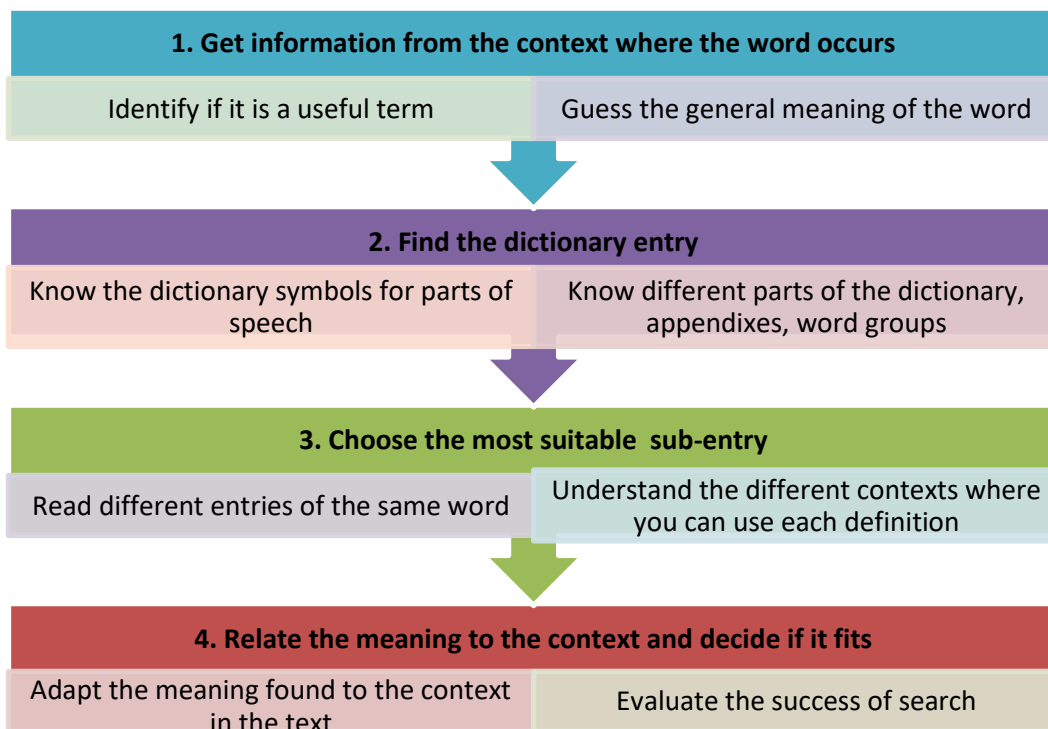
8. How to increase academic and less frequent vocabulary

Learning outcome: Apply strategies to learn academic and low-frequency vocabulary autonomously.

Some words are less frequent to find in standard texts or daily conversations. These words are considered more problematic to learn because we do not often hear them or read them. However, it is possible to apply some strategies to learn these terms and increase the quality of the language employed in our academic writing tasks. Nation mentions that “The teacher aims to train learners in using strategies to deal with less frequent or academic vocabulary. These strategies include guessing using context clues, using word parts to help remember words, using vocabulary cards, and using dictionaries” (2001, p. 30)

➤ **Using dictionaries**

An online dictionary might be our best friend during writing tasks. We use them to check word spelling, to avoid repetition, or maybe to find definitions. Dictionaries can also be devoted to vocabulary learning if we study new, unfamiliar items or search for the different meanings that a term can have. According to I.S.P. Nation, the following graphic shows the steps and skills required to exercise the strategy. (2001, p .454)



A fun activity you can do is search in the dictionary for a word from the academic word list every day and try to produce a sentence or a short story with it. There are also many free apps to download on your cellphone and provide good academic vocabulary, like English vocabulary builder-Test prep which collects words from international proficiency tests.

Pre-task 8.1

Strategy: Source -Use of dictionaries

Look up the highlighted words in the dictionary. Check how many sub-entries has each word and write them in the table. Finally, underline the meaning that fits better with the sentence context.

- a. The evidence produced an **argument** between those who followed Chomsky's ideas and those who supported Skinner's believes.
- b. The Ecuadorian ministry of education **establishes** academic standards through the national curriculum guidelines.
- c. Academic papers pass through a strict **review** process before being published in a journal.

- d. The budget reduction for scholarships in Ecuador stopped many teachers from achieving a master's **degree**.

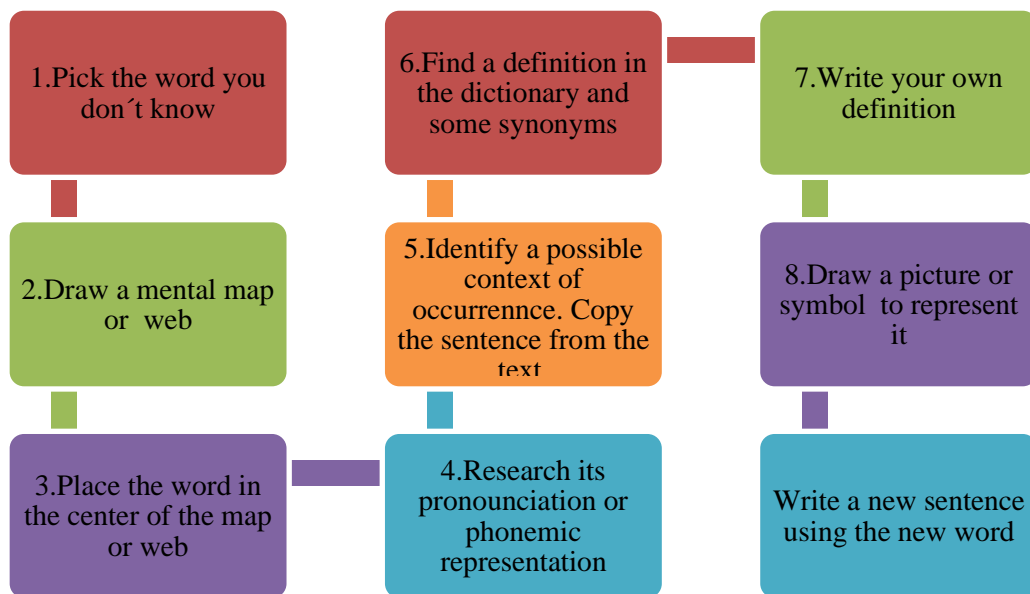
Word	Meanings

Note: Use this chart to exercise at home.

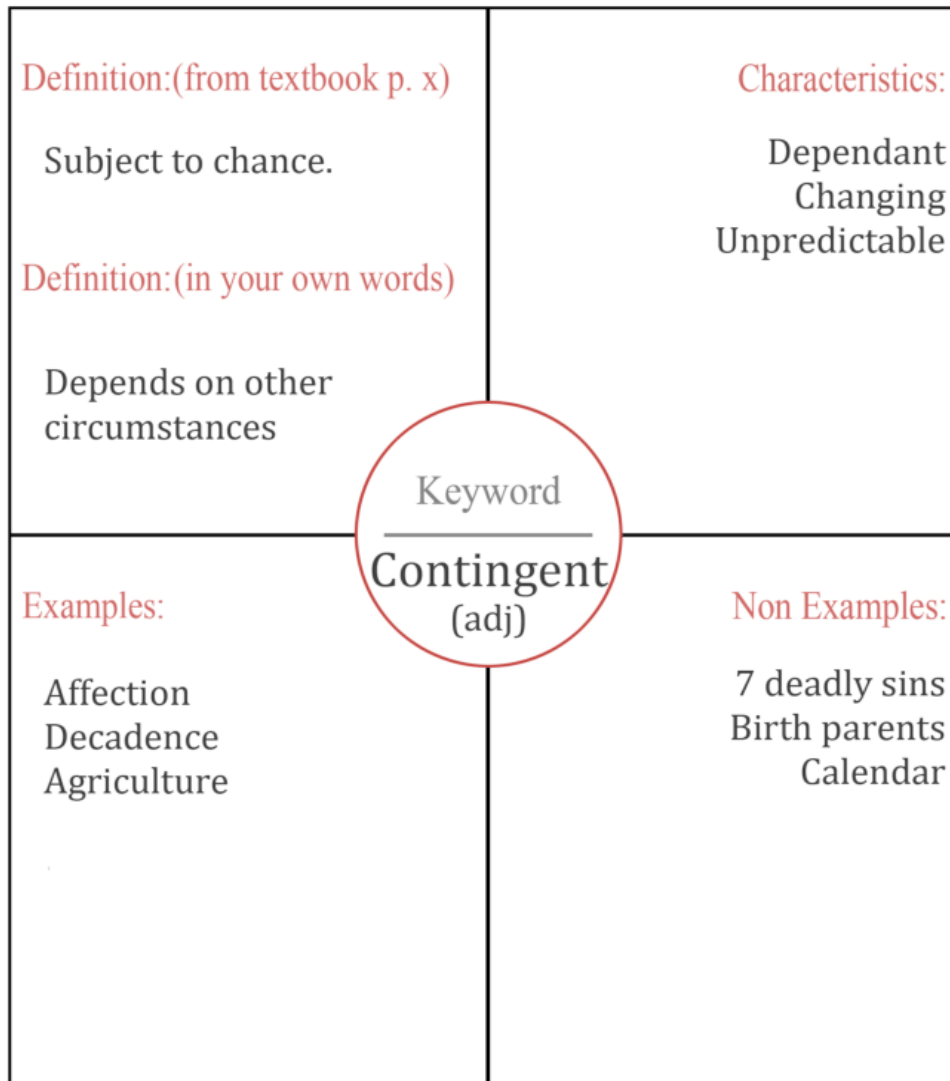
➤ **Semantic Mapping**

This strategy might allow you to associate new words with previous vocabulary knowledge. It is easier to understand a word if you related it with a familiar term, image, or thought. For this reason, when you listen to or read an unknown word, pay attention to it and write it down. Later, you can create a semantic map to help you remember. Semantic maps are similar to graphic organizers and they can be adapted according to your needs. The following picture shows the process to apply the strategy:

Steps



Example:



Note: Teaching and Learning Academic Vocabulary - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-Frayer-Model-for-the-word-Contingent_fig2_305072985 [accessed 13 May 2021]

Pre-task 8.2

Strategy: Source – Semantic Mapping

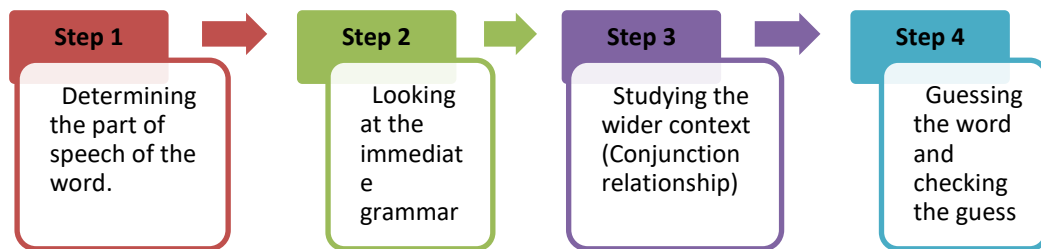
Identify five unknown words that you have read. Then, develop a semantic map for each one using the model presented below:

Definition from the dictionary	Sentence in which the word was used in the text:
Word	
Synonyms	Own sentences
Antonyms	
Definition (in own words)	Picture / Symbol

Note: Use this model to exercise at home.

➤ Guessing from the context

Academic life demands reading a great number of texts. This task is the perfect opportunity to learn new aspects of academic words using context to understand their meaning. You will try to use the words and structures already known to deduce a guess. The following graphic describes the steps to apply the strategy:



Note: Adapted from Clark & Nation, 1980, p. 1.

Techniques:

- ✓ Identify if the unknown word is a noun, verb, adjective. (step 1)
- ✓ Recognize if the unknown word is the synonym of another known word in the same sentence.
- ✓ Find if the unknown word is an antonym of another word in the sentence.
- ✓ Analyze if there is a cause-effect relationship between the unknown word and a known word.
- ✓ Separate the unknown term to analyze if there are affixes that modify it.

You also can carry out this process with a partner as a game. Once each person offers a possible meaning, you can check the dictionary and find out who approached the most to the actual definition. Practice makes perfect!

Pre-task 8.3

Strategy: Source – Guess meaning from context

Try to guess the meaning of the highlighted words. Please write what you think they mean. After that, check the dictionary definition to verify your guess.

- a. Linguists already have plenty to do to understand some language theories. They are surrounded by a wealth of **baffling** data that requires to be explained more easily.
- b. The research follows a set of twelve steps. In itself, it is not a **futile** procedure but a different procedure that the researcher aimed.
- c. Once people **attune** themselves to new concepts, they realize changes are not as difficult as they think.
- d. Many authors consider that vocabulary is the **linchpin** of second language acquisition.

Word	Guessing	Meaning from dictionary

. **Note: Use this chart to exercise at home.**

Evaluation:

Evaluation activity

The best way to evaluate how many lexical items you have learned and assess academic writing improvement is by producing a text. Remember to include academic features and formal vocabulary. Write an academic essay from 350 to 400 words describing a thesis topic related to English Language teaching. The essay's purpose is to communicate to the reader an overview of a research proposal's main points. Include title, introduction, body, conclusions, and implications.

2.5 Chapter II Conclusions

- The proposal results from a well-applied methodology of corpus analysis that provided precise information about pre-service teachers' writing difficulties such as low readability and low lexical richness.
- The booklet is designed to improve academic writing quality by enhancing the use of lexical and grammatical items in the text to make it more readable.
- Vocabulary learning strategies can be adapted in a Task-Based Learning approach to increase students' lexicon through the development of writing.

CHAPTER III. APPLICATION AND/OR VALIDATION OF THE PROPOSAL

3.1 Expert's evaluation

The proposal explained and developed in chapter two passed through a validation process. Three experts in the matter executed it; they employed their proficiency and knowledge to evaluate the booklet. They can identify if the activities, learning outcomes, and content in the proposal are adequate and pedagogical thanks to their vast experience teaching English to young adults. The assessment follows a rating scale with three parameters; excellent, good, and terrible, which estimates twelve predetermined points. It also includes four open questions related to criteria such as temporality, content, selectivity, impact and a space for general comments and recommendations. **(See appendix 12)** The following paragraphs describe the profiles and estimation of each professional.

The first expert is Mg. Patricia Marcela Chacón Porras, an English teacher in the Language Center at Technical University of Cotopaxi, I.D. number 0502211196. She achieved a Master's degree in applied linguistics in Bilingual Teaching (Spanish-English). She evaluated the proposal as excellent in 12 out of 12 parameters, considering it easy to use with good methodological structure and practical activities to develop appropriate academic writing. The expert evaluated the booklet content and rated it as understandable and relevant with accurate terminology. She recognized that the proposal contributes to vocabulary acquisition through learning strategies for academic writing improvement. She suggests that vocabulary is essential for academic writing, especially during selecting descriptive words to help readers envision what researchers are describing. The recommendation was to implement and disseminate the proposal in English

language learning programs promoting the acquisition of new and less frequent vocabulary while performing academic writing activities.

The second professional is Mg. Miryan Consuelo Salazar Tobar, an English teacher at Technical University of Ambato, I.D. number 1802840833. Her professional background includes a Master's degree in teaching English as a foreign language and a Master's degree in professional pedagogy and educational management. After evaluating the booklet, the expert ranks as excellent 11 out of 12 parameters, validating the proposal for its implementation by virtue of its relevant contribution to the field, proper structure and research process, well-achieved argumentation, and format. She considers that it can be applied as a guide to improve pre-service teachers' practice and develop their academic skills. She suggests training students in writing with different strategies based on lexical and readability analysis of written corpus. The specialist highlights the proposal's feasibility as a tool to implement a strategic approach to improve academic writing performance and benefit educators.

The third expert is Mg. Nelly Patricia Galora Moya, an English teacher at Technical University of Ambato, I.D. number 1803104601. Her academic achievements include a Master's degree in teaching English as a foreign language and a Master's degree in professional pedagogy and educational management. She rated the booklet as excellent in 12 out of 12 parameters considering it a valid and feasible proposal. The professional emphasizes a well-developed methodology as a guarantee for its application. Some characteristics in the booklet, such as illustrations, author reflections, clear content, and harmony between the objectives and goals, gave her reasons to support its significance. She suggests employing the proposal as a contribution to English language learning and teaching.

3.2 User's evaluation

The booklet focuses on pre-service English teachers as users. For this reason, the evaluators are twenty-nine students from the group who wrote the texts that built the analyzed corpus. They are studying the seventh semester in the major of English Language at Technical University of Cotopaxi. The evaluation includes seven

points valued under a rating scale with three parameters; excellent, acceptable, or deficient, and three open questions to understand their opinions and suggestions about the research product. **(See appendix 12)** The users consider that the booklet is a valid proposal that is current and offers a relevant contribution. The first point, which qualifies these characteristics in the proposal, was graded as excellent by twenty-six students, and four students graded it as acceptable. Twenty-seven pre-service teachers evaluated the proposal as excellent, whereas three think it is acceptable in the second point. They consider that the booklet results from a long research process that shows a complete concept and is critically contrasted with similar investigations.

In the analysis of features that give value to the proposal, such as originality of product and reflections of the author, twenty-two users rated these characteristics as excellent while eight considered them acceptable. The fourth point evaluates if the proposal has vocabulary, spelling, and language appropriate for the level of training. Twenty-nine students chose the option of excellent, while only one user chose the option acceptable. The graphics, relevance, clarity, and significance were evaluated in the fifth point. Twenty-seven students value these elements with excellent, whereas three evaluate these features with a rate of acceptable. All the users think there is harmony between the proposal objectives and the results and qualify this point with excellent. Finally, in the last point, users evaluated the structure of reflections and ideas in the booklet, twenty-seven of the evaluators graded these aspects as excellent, and three chose the option acceptable. This information provides a broader view of the user's opinions and their acceptance of the proposal. **(See appendix 13)**

3.3 Impact or results evaluation

There was a theoretical-practical workshop to present the booklet to the users and socialize its objectives, theory background, strategies, and activities. Thirty-five pre-service teachers formed the audience for the presentation. The aim was to provide some practical and helpful guidance in using the product and observe the effect of the strategies in vocabulary learning through developing exercises. The workshop's theoretical component included lexical richness, lexical density, lexical

diversity, lexical sophistication, academic writing features, and the explanation of vocabulary learning strategies such as guessing from context and using dictionaries. The practice was focused on applying the strategies mentioned before. Students exercised their abilities to infer vocabulary meaning employing the techniques explained before. After this virtual interaction space, there was a moment to revise the booklet and exchange opinions about the material. Finally, the evaluation form was shared to offer feedback to the author after deeply reading the content, exercises, and format.

The impact of the proposal was measured using the evaluation forms from users and experts. **(See appendixes 12 and 13)** The experts anticipated that the booklet might have a local impact due to pre-service English teachers were the primary beneficiaries of the proposal since they were the generators of the analyzed corpus and the workshop recipients. There was a direct contribution to enhancing their most recurrent weaknesses in academic writing and improving their lexical repertoire. They were able to recognize how essential is the vocabulary for the quality of academic writing. Most of the users agreed with the experts and concluded that the impact might be local because they think the proposal is helping them to surf writing difficulties during their research project execution and understanding the language. Other users considered a more extensive impact, such as regional, national or international, owing to the fact that the booklet is easy to share and understand. They consider that it allows producing new ideas about corpus analysis for educative purposes and teacher's training in Ecuador and Latin America. Academic writing is an indispensable subject for researchers, educators, students, and every person who intends to get an academic degree in English or international education. For this reason, the proposal and methodology can be applied in multiple contexts.

3.4 Proposal Results

During the presentation of the proposal, there were three sets of activities to practice the strategies for vocabulary learning and appreciate their feasibility. Guessing vocabulary from context strategy was applied in two sets of activities. The first set included techniques to guess meaning to learn new words, like identifying

synonyms, general knowledge, and parts of speech. Students achieved 70%, 71%, and 77% of success using these cues on each attempt. The second set applied strategies such as; identifying a word's definition, identifying antonyms and contrast, and identifying parts of speech. The achievement of word guessing applying these techniques increased to 83%, 100%, and 82% on each activity. (**See appendix 14**) Another strategy that was employed to learn vocabulary was the use of the dictionary. For this strategy, it is relevant to recognize which entry from the dictionary associates better with the actual use of the word in context. In this set of activities, the accuracy in choosing the best definition for an unknown word was improved thanks to continuous practice. The accuracy in associating meaning and context of use with an unknown academic word was 78%, 69%, 83%, and 100%. So, students applied the strategy correctly to learn new vocabulary and selected the appropriate context of use, which is a good indicator of the success of a strategy that increases quality in written texts and readability. The students understood how to apply the strategies to get the meaning and the steps they might follow to consolidate that knowledge through production. They showed interest in applying strategies and completing activities in the booklet because of its usefulness and clarity. These preliminary results are a sample of how effective the proposal could be for lexical training and academic writing enhancement if applied for extended periods.

3.5 Chapter III Conclusions

- The proposal is valid and feasible for application according to the experts in the field and users.
- Pre-service teachers at Technical University of Cotopaxi can use the booklet strategies and activities to enhance their academic writing skills.
- Students can employ strategies autonomously for learning vocabulary, they achieved high percentages of accuracy in understanding new words in different contexts, and the accuracy of their answers increases along with the experience.

General Conclusions

- Academic writing can be improved by employing an adequate combination of lexical and grammatical items.
- Word-frequency theory and corpus linguistics methodology are valuable for understanding students' needs in terms of vocabulary and assessing their written performance in a foreign language.
- The measurement of lexical richness revealed low lexical density, lexical sophistication, and lexical diversity in the corpus. Readability measurement exposed a difficult to read corpus because of low syntactic simplicity, low connectives, low verb cohesion, and low word concreteness, these elements increase the difficulty
- The booklet is a proposal that encourages the mutual beneficial relationship between writing tasks and vocabulary size growth.

Recommendations

- Applying the proposal is recommended to confirm the expert's validation and corroborate the expected outcomes. It is suggested to apply the same methodology for comparison.
- Teachers can adapt the proposal for different purposes, such as teaching high-frequency words or specialized vocabulary by adapting the examples in the booklet.
- It is suggested to research how effective could be applying Task-Based Learning Approach to teach academic writing in the higher education context.
- Further study should consider which elements besides lexical richness influence text's readability employing the same methodology.
- Corpus linguistics methodology should be employed by English teachers to identify student's needs. Many research studies can be derived from the present research concerning vocabulary size.

References

- Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation* (1st ed.) Oxford University Press. Papers from the International Symposium on Language Testing
- Arnaud, P. J. (1984). *The lexical richness of L2 written productions and the validity of vocabulary tests* (143). University Lyon. <https://files.eric.ed.gov/fulltext/ED275164.pdf>
- Bailey, S. (2011). *Academic Writing A Handbook for International Students* (3rd ed.). Routledge. ISBN13: 978-0-203-83165-6
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511621024>
- Björk, L., Knight, M., & Wikborg, E. (1992). *The writing process: Composition writing for University students*.
- Breeze, R., R. (n.d.). Researching simplicity and sophistication in student writing. *Revistas Científicas de la Universidad de Murcia International Journal of English Studies*, 8(1), 51-66. <https://revistas.um.es/ijes/article/view/49091>
- Cobb, T. (2002). Review of Paul Nation, *Learning vocabulary in another language*. Cambridge: Cambridge University Press. 2001. 477 + xiv pp. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 46(3-4), 242-245. <https://doi.org/10.1017/s0008413100018260>
- Cobb, T. *Web Vocabprofile* [accessed 15 March 2021 from <http://www.lexutor.ca/vp/>], an adaptation of Heatley, Nation & Coxhead's (2002) *Range*.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213. <https://doi.org/10.2307/3587951>
- Clarke, D., & Nation, I. (1980). Guessing the meanings of words from context: Strategy and techniques. *System*, 8(3), 211-220. [https://doi.org/10.1016/0346-251x\(80\)90003-2](https://doi.org/10.1016/0346-251x(80)90003-2)

- Chall, J. (1981). Readability: Conceptions and misconceptions. *Support for the learning and teaching of English - National Council of Teachers of English*, 1-4. <https://files.eric.ed.gov/fulltext/ED263620.pdf>
- Crossley, S. A. (2013). Advancing research in second language writing through computational tools and machine learning techniques: A research agenda. *Language Teaching*, 46(2), 256-271. <https://doi.org/10.1017/s0261444812000547>
- Cutts, M. (2013). *Oxford guide to plain English* (4th ed.). OUP Oxford. ISBN 9780199669172
- Dakhi, S., & Hutabarat, H. (2018). Language effectiveness and factors influencing scientific writing of Indonesian undergraduate thesis. *English Review: Journal of English Education*, 7(1), 61-74. <https://doi.org/10.25134/erjee.v7i1.1496>
- Dash, N. S. (2008). Introduction. In *Corpus linguistics: An introduction* (pp. 1-22). Pearson Education India. ISBN: 81-317-1603-1
- Doykova, I. (2016). The strive for readability in academic papers (EMP). *Knowledge International Journal*, 15(3), 1487-1490. <https://ikm.mk/ojs/index.php/KIJ/issue/view/74>
- DuBay, W. H. (2004). *The principles of readability*. Impact Information. <http://www.impact-information.com/impactinfo/readability02.pdf>
- Duin, A. H., & Graves, M. F. (1987). Intensive vocabulary instruction as a prewriting technique. *Reading Research Quarterly*, 22(3), 311. <https://doi.org/10.2307/747971>
- Esfandiari, M., Knight, P., Molinari, J., & Zacharias, S. (2012). *Task-Based Learning Applied* (1st ed.). The Nottingham Jubilee Press. ISBN 9781906235178
- Estoup, J. B. (1912). *Gammes st'enographiques : m'ethode et exercices pour l'acquisition de la vitesse* (3rd ed.). <https://gallica.bnf.fr/ark:/12148/bpt6k9440031>
- Feyen, J. (2019). Una LLAMADA DE ATENCIÓN Para Las Universidades Ecuatorianas. *MSKN 2019*, 10, 5-14. <https://doi.org/10.18537/mskn.10.02.01>

- Flesch, R. 1948. "A new readability yardstick." *Journal of Applied Psychology* 32:221-233.
- Fuertes-Olivera, P. A. (2018). *The Routledge handbook of lexicography* (1st ed.). Routledge. <https://doi.org/10.4324/9781315104942>
- Gill, A. K., & ., K. (2017). Teaching approaches, methods and strategy. *Scholarly Research Journal for Interdisciplinary Studies*, 4(36).
<https://doi.org/10.21922/srjis.v4i36.10014>
- Gonzalez, M. (2013). *The Intricate Relationship Between Measures Of Vocabulary Size And Lexical Diversity As Evidenced In Non-native And Native Speaker Academic Compositions* [Doctoral dissertation].
<https://stars.library.ucf.edu/etd/2633>
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202. <https://doi.org/10.3758/bf03195564>
- Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal*, 115(2), 210-229. <https://doi.org/10.1086/678293>
- Gray, W. S., & Leary, B. E. (1935). What makes a book readable?. *The English Journal*, 25(9), 358. <https://archive.org/details/whatmakesabookre028092mbp>
- Gregori-Signes, C., & Clavel-Arroitia, B. (2015). Analyzing lexical density and lexical diversity in University students' written discourse. *Procedia - Social and Behavioral Sciences*, 198, 546-556.
<https://doi.org/10.1016/j.sbspro.2015.07.477>
- Gyasi, W. K. (2013). Readability and academic communication: A comparative study of undergraduate students' and handbook of three Ghanaian universities. *IOSR Journal of Computer Engineering*, 13(6), 41-50.
<https://doi.org/10.9790/0661-1364150>
- Halliday, M., & Hasan, R. (1976). *Cohesion in English*. Longman.

- Halliday, M. A. (1989). *Spoken and written language* (2nd ed.). Oxford University Press. <https://doi.org/0194371530>
- Halliday, M., & Matthiessen, C. M. (2013). *Halliday's introduction to functional grammar* (4th ed.). Routledge. <https://doi.org/10.4324/9780203431269>
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a foreign language*, 13(1), 403-430. <https://openaccess.wgtn.ac.nz/ndownloader/files/23411972>
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working Papers*, (53), 61-79. <https://journals.lub.lu.se/LWPL/article/download/2273/1848/>
- Kakkonen, T. (2009, September). *TexComp - A text complexity analyzer for student texts* [Conference session]. Conference ICL2009 , Villach, Austria. <https://arxiv.org/abs/1206.6612v1>
- Karakoç, D., & Köse, G. (2017). The impact of vocabulary knowledge on reading, writing and proficiency scores of EFL learners. *Journal of Language and Linguistic Studies*, 13(1), 352-378. <https://files.eric.ed.gov/fulltext/EJ1140609.pdf>
- Khoo, E. (2005). *Academic style / Writing support*. University of Toronto Scarborough. Retrieved April 11, 2021, from <https://www.utoronto.ca/twc/academic-style>
- Kondru, J. (2006). *Using part of speech structure of text in the prediction of its readability* [Master's thesis]. <http://hdl.handle.net/10106/178>
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322. <https://doi.org/10.1093/applin/16.3.307>
- Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: single words and collocations. In *L2 vocabulary acquisition, knowledge and use. New perspectives on assessment and corpus analysis*. Bardel, C., C, Lindqvist, & B. Laufer (Eds.), *EUROSLA Monographs 2*, 127-148. <http://www.eurosla.org/monographs/EM02/EM02home.php>

- Malverdi, M., & Heidari Darani, L. (2018). EFL Textbook Evaluation: An Analysis of Readability and Vocabulary Profiler of Four Corners Book Series. *International Journal of Foreign Language Teaching & Research*, 6(22). http://jfl.iaun.ac.ir/article_601180.html
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics an Introduction* (2nd ed.). Edinburgh University Press.
- McLean, S., & Kramer, B. (2015). The creation of a New Vocabulary Levels Test. *Shiken*, 19(2), 1-11.
<https://www.lexutor.ca/tests/levels/recognition/nvlt/paper.pdf>
- McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 55(1), 51-62. <https://doi.org/10.1037/h0087352>
- McNamara, D. S., Louwse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-matrix: Capturing linguistic features of cohesion. *Discourse Processes*, 47(4), 292-330. <https://doi.org/10.1080/01638530902959943>
- Meara, P. (1990). A note on passive vocabulary. *Interlanguage studies bulletin (Utrecht)*, 6(2), 150-154. <https://doi.org/10.1177/026765839000600204>
- Morris, L., & Cobb, T. (2004). Vocabulary profiles as predictors of the academic performance of teaching English as a second language trainees. *System*, 32(1), 75-87. <https://doi.org/10.1016/j.system.2003.05.001>
- Muller, C. (1964). *Essai de statique lexicale: L'illusion comique de P. Corneille*. Klincksieck.
- Muller, C. (1977). *Principes et Méthodes de Statistique Lexicale*. Paris:Hachette. <https://doi.org/10.3406/lfr.1969.5419>
- Nation, I., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12, 355-380. <https://cutt.ly/mbvdw4q>
- Nation, I. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59-82. <https://doi.org/10.3138/cmlr.63.1.59>

- Nation, I., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Ngan, N.N.T. & Thao, N T.T. (2016). "Lexical Density and Readability of Non-English Majored Freshmen's Writing in Vietnamese Context." 7th International Conference on TESOL: Innovations in English Language Teaching and Learning, 11-13 August, Qui Nhon City, Vietnam.
Retrieved from
http://www.vnseameo.org/TESOLConference2016/materials/06_1.pdf
- Nouri, N., & Zerhouni, B. (2018). Lexical frequency effect on reading comprehension and recall. *Arab World English Journal*, 9(2), 234-250. <https://doi.org/10.31235/osf.io/t5n8h>
- Nunan, D. (1990). Designing tasks for the communicative classroom. *The Modern Language Journal*, 74(1), 89. <https://doi.org/10.2307/327949>
- Orngreen, R., Katre, D., & Sandeep, M. (2010). Analyzing cultural usability of mobile keypad and displays for textual communication in internationalization and localization perspectives. *Human Work Interaction Design: Usability in Social, Cultural and Organizational Contexts*, 115-131. https://doi.org/10.1007/978-3-642-11762-6_10
- Panavelil, A. (2015). Teaching and Learning to Write: Using a Task-Based Approach in an EFL Class. In R. Al-Mahrooqi, V. Vijay Singh, & A. Adrian Roscoe (Eds.), *Methodologies for effective writing instruction in EFL and ESL classrooms* (pp. 115-129). IGI Global. <https://doi.org/10.4018/978-1-4666-6619-1.ch008>
- Prabhu, N. S. (1987). Second language pedagogy. *Oxford University Press*, 1-156.
- Qiu, J., Zhao, R., Yang, S., & Dong, K. (2017). *Informetrics: Theory, methods and applications*. Springer.
- Read, J. (2000). Assessing vocabulary. *Cambridge University Press*. <https://doi.org/10.1017/cbo9780511732942>
- Rezaee, A. A., & Norouzi, M. H. (2011). Readability formulas and cohesive markers in reading comprehension. *Theory and Practice in Language Studies*, 1(8). <https://doi.org/10.4304/tpls.1.8.1005-1010>

- Richards, B., Malvern, D., & Graham, S. (2008). Word frequency and trends in the development of French vocabulary in lower intermediate students during Year 12 in English schools. *Language Learning Journal*, 36, 199-213. <https://doi.org/10.1080/09571730802390098>
- Richards, J. C., & Schmidt, R. W. (2002). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Longman Publishing Group.
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205-225. <https://doi.org/10.1017/s0267190511000055>
- Sağlamel, H., & Kayaoğlu, M. N. (2015). English major students' perceptions of academic writing: A struggle between writing to learn and learning to write. *Journal of History Culture and Art Research*, 4(3), 37. <https://doi.org/10.7596/taksad.v4i3.477>
- Sarani, A., & Farzaneh, L. (2012). The impact of task-based approach on vocabulary learning in ESP courses. *English Language Teaching*, 5(10), 118-128. <https://doi.org/10.5539/elt.v5n10p118>
- Sari, W. (2019). *Lexical Richness and Syntactic Complexity of the English exam papers of SBMPTN* [Master's thesis]. <http://repository.unair.ac.id/id/eprint/91510>
- Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2015). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), nation (2006) and Cobb (2007). *Language Teaching*, 50(2), 212-226. <https://doi.org/10.1017/S0261444815000075>
- Sherman, L. A. (1893). *Analytics of literature: A manual for the objective study of English prose and poetry*. Boston, Ginn. <https://archive.org/details/analyticsofliter00sherooft>
- Sigray, P., Matheson, G. J., Bjorn, C., & Thompson, W. (2017). Decision letter: The readability of scientific texts is decreasing over time. *e-Life*, 6(27725), 1-14. <https://doi.org/10.7554/elife.27725.028>
- Swaran, C., & Marappan, P. (2020). A review of research on the importance of higher order thinking skills (Hots) in teaching English language. *Journal of critical reviews*, 7(08). <https://doi.org/10.31838/jcr.07.08.161>

- Syarif, H., & Putri, R. E. (2018). *Lingua Didaktika: Jurnal Bahasa dan Pembelajaran Bahasa*. 86-94, 12(2).
<https://doi.org/10.24036/ld.v12i1.10408>
- Thorndike, E. L. (1921). *The teacher's word book of 3000 words*. Teachers College, Columbia University. <https://cutt.ly/qbvtMn4>
- Tidball, F., & Treffers-Daller, J. (2008). Analysing lexical richness in French learner language: What frequency lists and teacher judgements can tell us about basic and advanced words. *Journal of French Language Studies*, 18(3), 299-313. <https://doi.org/10.1017/s0959269508003463>
- TOEFL iBT. (2019). *Test and Score Data Summary 2019*.
https://www.ets.org/s/toefl/pdf/94227_unlweb.pdf
- Tovar Viera, R. (2017, December 30). Vocabulary knowledge in the production of written texts: a case study on EFL language learners. *Revista Tecnológica - ESPOL*, 30(3).
<http://www.rte.espol.edu.ec/index.php/tecnologica/article/view/628/377>
- Ure, J. (1971) *Lexical density and register differentiation*. In G. Perren and J.L.M. Trim (eds), *Applications of Linguistics*, Cambridge University Press. 443-452.
- Vajjala, S., & Meurers, D. (2012, June). *On Improving the Accuracy of Readability Classification using Insights from Second Language Acquisition* [Paper presentation]. The 7th Workshop on the Innovative Use of NLP for Building Educational Applications, Montreal, Canada.
<https://www.aclweb.org/anthology/W12-2019.pdf>
- Van Zeeland, H., & Schmitt, N. (2012). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457-479.
<https://doi.org/10.1093/applin/ams074>
- Vuković Stamatović, M., Bratić, V., & Lakić, I. (2020). Vocabulary of L1 and L2 graduation theses written by English philology students: Academic writing of Montenegrin and US students compared. *ELOPE: English Language*

Overseas Perspectives and Enquiries, 17(2), 101-113.

<https://doi.org/10.4312/elope.17.2.101-113>

Williamson, G. (2018, May 13). *Type-token ratio*. SLT info.

<https://www.sltinfo.com/type-token-ratio/>

Willis, J. (1996). *A flexible framework for task-based learning* (2nd ed.). An overview of a task-based framework for language teaching.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley Press.

Appendices

Appendix 1

The new vocabulary level test

This is a vocabulary test.

Please select the option a, b, c, or d which has the closest meaning to the word in **bold**.

Example question

see: They **saw** it.

- a. cut
- b. waited for
- c. looked at
- d. started

The correct answer is **c**.

If you do not know the word at all, please do not answer the question and continue to the next question.

However, if you think that you may know the word, please try to answer.

Let's begin.

Bilingual versions of this test are also available in Japanese and Chinese. If you are interested in making a version in a different language please feel free to do so, and contact the authors if necessary or required.

Section 1

1. time: They have a lot of **time**.

- a. money
- b. food
- c. hours
- d. friends

2. stone: She sat on a **stone**.

- a. hard thing
- b. kind of chair
- c. soft thing of the floor
- d. part of a tree

3. poor: We are **poor**.

- a. have no money
- b. happy
- c. very interested
- d. tall

4. drive: She **drives** fast.

- a. swims
- b. learns
- c. throws balls

d. uses a car

5. jump: She tried to **jump**.

- a. lie on top of the water
- b. get up off the ground
- c. stop the car on the road
- d. move very fast

6. shoe: Where is your other **shoe**?

- a. the person who looks after you
- b. the thing you keep your money in
- c. the thing you use for writing
- d. the thing you wear on your foot

7. test: We have a **test** in the morning.

- a. meeting
- b. travelling somewhere
- c. a set of questions
- d. an idea to do something

8. nothing: He said **nothing** to me.

- a. very bad things

- b. zero
- c. very good things
- d. something

9. cross: Don't **cross**.

- a. go to the other side
- b. push something
- c. eat too fast
- d. wait for something

10. actual: The **actual** one is larger.

- a. real
- b. old
- c. round
- d. other

11. any: Does she have **any** friends?

- a. some
- b. no
- c. good
- d. old

12. far: You have walked **far**!

- a. for a long time
- b. very fast
- c. a long way
- d. to your house

13. game: I like this **game**.

- a. food
- b. story
- c. group of people
- d. way of playing

14. cause: He **caused** the problem.

- a. made
- b. fixed
- c. explained
- d. understood

15. many: I have **many**.

- a. none
- b. enough
- c. a few
- d. a lot

16. where: **Where** did you go?

- a. at what time
- b. for what reason
- c. to what place
- d. in what way

17. school: This is a big **school**.

- a. where money is kept
- b. sea animal
- c. place for learning
- d. where people live

18. grow: All the children **grew**.

- a. drew pictures
- b. spoke
- c. became bigger
- d. cried a lot

19. flower: He gave me a **flower**.

- a. night clothes
- b. small clock
- c. beautiful plant
- d. type of food

20. handle: I can't **handle** it.

- a. open
- b. remember
- c. deal with
- d. believe

21. camp: He is in the **camp**.

- a. sea
- b. place outside where people enjoy nature
- c. hospital
- d. building where people sleep

22. lake: People like the **lake**.

- a. area of water
- b. very young child
- c. leader
- d. quiet place

23. past: It happened in the **past**.

- a. before now
- b. big surprise
- c. night
- d. summer

24. round: It is **round**.

- a. friendly
- b. very big
- c. very quick
- d. with no corners

Section 2

1. maintain: Can they **maintain** it?

- a. keep it like it is
- b. make it larger
- c. get a better one than it
- d. get it

2. period: It was a difficult **period**.

- a. small set of questions
- b. time
- c. thing to do
- d. book

3. standard: Her **standards** are very high.

- a. the back under her shoes
- b. test scores
- c. cost of something
- d. level of how good she wants things to be

4. basis: This was used as the **basis**.

- a. answer
- b. resting place
- c. next step
- d. main part

5. upset: I am **upset**.

- a. strong
- b. famous
- c. rich
- d. angry

6. drawer: The **drawer** was empty.

- a. box that goes in and out for clothes
- b. place to keep cars
- c. place used to keep things cold
- d. animal house

7. pub: They went to the **pub**.

- a. place where people drink and talk
- b. place that keeps money
- c. large building with many shops
- d. building for swimming

8. circle: Make a **circle**.

- a. rough picture
- b. space with nothing in it
- c. round shape
- d. large hole

9. pro: He's a **pro**.

- a. person who has the job to find important secrets
- b. stupid person
- c. person who writes articles
- d. someone who is very good at doing something and is paid to do it.

10. soldier: He is a **soldier**.

- a. person who works in business
- b. person who studies at school
- c. person who works with wood
- d. person who fights in a war

11. result: They were waiting for the **results**.

- a. right time
- b. questions
- c. money
- d. effects of something

12. resist: They **resisted** it.

- a. made it work again
- b. looked at it twice
- c. thought hard about
- d. acted against

13. lend: She often **lends** her books.

- a. lets people use them
- b. draws inside them
- c. cleans them
- d. writes her name on them

14. refuse: She **refused**.

- a. went back
- b. thought about something
- c. said no
- d. stayed late

15. speech: I enjoyed the **speech**.

- a. type of presentation
- b. very fast run
- c. short piece of music
- d. type of hot food

16. pressure: They used too much **pressure**.

- a. money
- b. time
- c. hard pushing
- d. bad words

17. refer: She **referred** to him.

- a. supported him
- b. let him go first
- c. talked about him
- d. answered him

18. army: They saw the **army**.

- a. black and white animal
- b. place where books are kept
- c. person who lives nearby
- d. people who protect a country

19. knee: Take care of your **knee**.

- a. small child
- b. part of your leg

- c. plan for spending money
- d. something that is yours

20. rope: He found a **rope**.

- a. thick and strong string
- b. something used to make holes
- c. strong box for keeping money
- d. metal tool used to climb up high

21. brand: This is a good **brand**.

- a. dance party
- b. first try
- c. place to wait for others
- d. name of a company

22. seal: They **sealed** it.

- a. fixed it
- b. closed it tightly
- c. looked at it carefully
- d. opened it quickly

23. warn: They were **warned**.

- a. pushed away
- b. welcomed inside
- c. told about bad things
- d. led into war

24. reserve: They have large **reserves**.

- a. things kept to use later
- b. machine for making bread
- c. money from other people
- d. group that runs a company

Section 3

1. restore: It has been **restored**.

- a. said again
- b. given to a different person
- c. given a lower price
- d. made like new again

2. compound: They made a new **compound**.

- a. agreement between two people
- b. thing made of two or more parts
- c. group that works together
- d. guess based on past experience

3. latter: I agree with the **latter**.

- a. man from the church
- b. reason given before
- c. second one of two things
- d. answer to the spoken question

4. pave: It was **paved**.

- a. stopped quickly
- b. divided into many parts
- c. given gold edges
- d. covered with a hard surface

5. remedy: We found a good **remedy**.

- a. way to fix a problem
- b. place to eat in public
- c. way to prepare food
- d. rule about numbers

6. bacterium: They didn't find a single **bacterium**.

- a. small living thing causing sickness
- b. plant with red or orange flowers
- c. animal that carries water on its back
- d. thing that has been stolen and sold to a shop

7. behavior: Look at her **behavior!**

- a. people who have come to listen
- b. the way she acts
- c. large amount of money
- d. small land with water around it

8. fuel: Do you have any **fuel?**

- a. material used to make energy
- b. a drug that stops pain
- c. clothing used to keep you warm
- d. a material put in walls to keep heat inside

9. silk: It's made of **silk.**

- a. smooth and soft cloth
- b. hard black wood
- c. animal fur
- d. very light metal

10. conceive: Who **conceived** the idea?

- a. told it to others
- b. explained it
- c. thought of it first
- d. said it was bad

11. legend: It is now a **legend.**

- a. building for keeping old things
- b. thing that is always done
- c. story from the past
- d. event that happens regularly

12. impose: This was **imposed.**

- a. completely changed
- b. in the middle of other things
- c. made to look like something else
- d. forced to happen by someone in power

13. solution: There is no **solution.**

- a. time
- b. support
- c. problem
- d. answer

14. celebrate: We have **celebrated** a lot recently.

- a. found something for the first time
- b. seen many new places
- c. worked very hard
- d. had a lot of parties

15. independence: He has too much **independence.**

- a. freedom from outside control
- b. time by himself
- c. physical strength
- d. feeling of being better than others

16. tunnel: We need a **tunnel** here.

- a. way through or under something
- b. long piece of wood or metal to hold
- c. mark on paper to show a short space
- d. piece of material to cover a window

17. reward: He got a good **reward.**

- a. things said about him by others
- b. someone to help him in the house
- c. money or gift for the things he did
- d. large group of people to listen to him

18. review: The committee **reviewed** the plan.

- a. examined it carefully for a decision

- b. agreed to allow
- c. made more just like it
- d. threw it away

19. mode: The **mode** of production has changed.

- a. type
- b. speed
- c. attitude
- d. amount

20. personnel: I don't like the **personnel** there.

- a. type of chair that folds
- b. machine that controls the heat
- c. people who work there
- d. person who owns a company

21. competent: She was very **competent.**

- a. very fast
- b. made angry easily
- c. able to do things
- d. easily hurt

22. devastate: The city was **devastated.**

- a. made beautiful for a special occasion
- b. separated from the rest of the world
- c. suffered great damage
- d. made dirty by small animals

23. constituent: This is an important **constituent.**

- a. building
- b. agreement
- c. idea
- d. part

24. weave: She knows how to **weave.**

- a. make cloth
- b. join pieces of metal together
- c. make people think something
- d. trick people

Section 4

1. patience: He has a lot of **patience.**

- a. ability to wait
- b. free time
- c. faith in God
- d. knowledge

2. strap: She broke the **strap.**

- a. promise
- b. top
- c. plate
- d. belt

3. weep: He **wept.**

- a. finished school
- b. cried
- c. died quickly
- d. thought deeply

4. haunt: The house is **haunted.**

- a. full of decorations
- b. allowed to be used for money
- c. completely empty
- d. full of ghosts

5. cube: I need one more **cube.**

- a. pin
- b. box
- c. cup
- d. postcard

6. peel: Shall I **peel** it?

- a. let it sit in water for a long time

- b. take the skin off it
- c. make it white
- d. cut it into thin pieces

7. distress: He felt **distressed**.

- a. unwanted
- b. satisfied
- c. unhappy
- d. energetic

8. depart: She **departed** yesterday.

- a. went away
- b. said no
- c. went down a hill
- d. got worse

9. romance: They had a short **romance**.

- a. difference of opinion
- b. holiday away from home
- c. serious discussion
- d. love relationship

10. ambition: He has no **ambition**.

- a. strong desire to do well
- b. ability to understand people's feelings
- c. ability to make new things
- d. enjoyment of life

11. dash: They **dashed** over it.

- a. ran quickly
- b. walked slowly
- c. fought bravely
- d. looked quickly

12. drown: People have **drowned** here.

- a. eaten outside
- b. died in water
- c. dug a hole
- d. cut down trees

13. originate: It **originated** here.

- a. grew very well
- b. changed shape
- c. remained
- d. first started

14. leaf: He touched the **leaf**.

- a. part of a plant
- b. soft shoe
- c. top of a bottle
- d. glass window

15. amateur: She is an **amateur** player.

- a. someone who plays for fun, not money
- b. player who replaces other hurt players
- c. player representing her country
- d. ball-sports player

16. evacuate: They were **evacuated**.

- a. moved to another place for safety
- b. searched for guns or knives
- c. frightened suddenly
- d. made to look like criminals

17. exert: Don't **exert** yourself!

- a. praise too much
- b. hurt yourself
- c. work too hard
- d. give yourself everything you want

18. marble: It was made of **marble**.

- a. hard stone
- b. hard wood

- c. soft metal
- d. soft cloth

19. diminish: It has **diminished**.

- a. become dark
- b. become less in size
- c. become cloudy
- d. grown colder

20. sheriff: The **sheriff** was friendly.

- a. pilot
- b. housekeeper
- c. policeman
- d. teacher

21. monarch: They saw the **monarch**.

- a. army group
- b. gate
- c. king or queen
- d. criminal

22. plunge: It **plunged**.

- a. danced around
- b. was made quiet
- c. dropped suddenly
- d. stayed still

23. mourn: They **mourned** for several years.

- a. performed on the street
- b. felt very sad
- c. worked hard
- d. used their money carefully

24. fragile: These things are very **fragile**.

- a. special
- b. hard to find
- c. popular
- d. easily broken

Section 5

1. scrub: He is **scrubbing** it.

- a. cleaning
- b. repairing
- c. worrying about
- d. drawing pictures

2. dinosaur: The children were pretending to be **dinosaurs**.

- a. people who look for gold
- b. small people that fly
- c. animals that make fire
- d. animals that lived a long time ago

3. nun: We saw a **nun**.

- a. small worm
- b. big accident
- c. woman who serves her religion
- d. strange light in the sky

4. compost: We need some **compost**.

- a. strong support
- b. mental help
- c. strong material that is used for building
- d. soil used to help the garden

5. miniature: It is a **miniature**.

- a. small version of something
- b. brick house
- c. very small living creature

d. detailed plan for a building

6. crab: Do you like **crabs**?

- a. small sea animals
- b. hard thin salty bread
- c. original copy of a piece of music
- d. insect which sings and jumps

7. vocabulary: You will need more **vocabulary**.

- a. words
- b. skills
- c. money
- d. guns

8. corpse: The **corpse** was found in the park.

- a. large and deep cup
- b. mobile phone
- c. artist's hat
- d. dead body

9. rove: He is **roving**.

- a. getting drunk
- b. traveling around
- c. making a musical sound with his lips
- d. working hard using his body

10. divert: The rivers were **diverted**.

- a. made to move in a different way
- b. given bridges
- c. made very dirty
- d. made wider and deeper

11. trench: They looked at the **trench**.

- a. mountain
- b. long hole
- c. pile of trash
- d. beautiful sight

12. technician: She is a **technician**.

- a. man with magical abilities
- b. person who works with and fixes machines
- c. doctor who cares for young children
- d. person who is good at music

13. query: I have a **query**.

- a. headache
- b. large amount of money
- c. question
- d. good idea

14. mug: This **mug** needs a wash.

- a. big cup
- b. old car you like
- c. clothes worn under other clothes
- d. area in front of the door where rain and wind cannot reach

15. static: It's **static** at the moment.

- a. not popular
- b. demanded by law
- c. often said
- d. not moving or changing

16. slaughter: We read about the **slaughter** in the paper.

- a. problem
- b. scientific research
- c. killing
- d. sports event

17. spider: We caught the **spider**.

- a. disease that gives red spots
- b. small animal with eight legs
- c. small public bus
- d. oily fish

18. circus: We went to the **circus**.

- a. place for people who love God
- b. traveling company of entertainers
- c. place where people run races
- d. music group

19. sofa: He bought a **sofa**.

- a. soft seat for two or more people
- b. cutting machine
- c. long pipe for putting water on the garden
- d. a small car with four wheels that a baby can ride in while someone pushes it

20. logo: They have a pretty **logo**.

- a. tree with red fruit
- b. reception
- c. picture or word that represents a company
- d. a holiday home

21. commemorate: We must **commemorate** his actions.

- a. remember something or someone
- b. pretend to agree with something
- c. protest against something
- d. say good things about him

22. crook: They were **crooks**.

- a. people who are not honest
- b. people who work at hospitals
- c. people who cannot walk
- d. people who design buildings

23. volt: How many **volts** were used?

- a. large envelope for business letters
- b. something used to add flavor to food
- c. units measuring electrical power
- d. material that attracts other metals

24. warfare: Modern **warfare** is frightening.

- a. crime
- b. dancing
- c. fighting
- d. pollution

Section 6

1. concept: This is a difficult **concept**.

- a. legal agreement
- b. idea about what something is
- c. way of doing things
- d. a written explanation of a law

2. similar: These articles are **similar**.

- a. about a certain thing
- b. of great quality
- c. easy to understand
- d. close to the same

3. item: The next **item** is very important.

- a. thing on a list
- b. question sheet
- c. meeting of people
- d. way something looks

4. component: Each **component** is very important.

- a. set of ideas which support something
- b. flat part that sits on top of another
- c. small part of something bigger
- d. the person you work with

5. compensate: The government should **compensate** the farmers.
 a. give something good to balance something bad
 b. stop them from joining a group
 c. find where they are
 d. bring them together
6. professional: She wants to be a **professional** musician.
 a. someone who stays at home
 b. someone who gets paid to play
 c. someone on a list
 d. someone known by many people
7. external: They worried about the **external** damage.
 a. not known
 b. outside
 c. based on facts
 d. following
8. clause: Please fix that **clause**.
 a. part of a sentence
 b. something you are trying to do
 c. large picture
 d. small object
9. migrate: The animals began to **migrate**.
 a. work together
 b. move together to a different place
 c. come together as a group
 d. change together
10. priority: That is our **priority**.
 a. deal between two people
 b. most important thing
 c. something that has been printed
 d. person who comes next
11. reverse: Try it in **reverse**.
 a. the other direction
 b. the way things are arranged
 c. with the correct sound
 d. at the correct time
12. arbitrary: Her decision was **arbitrary**.
 a. not chosen for a reason
 b. necessary for success
 c. not able to be changed
 d. good enough for a purpose
13. mutual: The feeling was **mutual**.
 a. easy to understand
 b. fully developed
 c. the same between two people
 d. kept under control
14. alternative: Is there an **alternative**?
 a. another choice
 b. thing to do
 c. something to say
 d. activity with many people
15. colleague: That is my **colleague**.
 a. something that people talk about
 b. plan of things to do
 c. person you work with
 d. piece of writing
16. legal: Is this meeting place **legal**?
 a. based on the law
 b. free to be used
 c. easy to see
 d. important to someone
17. site: He looked for a better **site**.
 a. basic part of something
 b. opinion about the price
 c. place where something is
 d. something brought from another country
18. institute: We must **institute** new changes.
 a. get with effort
 b. control with laws
 c. begin or create
 d. search for
19. retain: How will the club **retain** its members?
 a. mix them together
 b. help them develop
 c. help them work together
 d. keep them
20. phase: This is one **phase** of the new system.
 a. list of things in a special order
 b. short part of a process
 c. range of levels
 d. rule that controls what something is
21. pursue: This year she will **pursue** the group's goals.
 a. try to get
 b. change
 c. check over time
 d. make easier
22. recover: The men **recovered** their strength.
 a. showed other people
 b. used for a reason
 c. said that they know
 d. got back
23. diverse: Having **diverse** information is important.
 a. with no mistakes
 b. very small amount
 c. able to be changed
 d. having different types
24. hierarchy: This **hierarchy** is very common.
 a. set of ideas a group has
 b. group with people at different levels
 c. dangerous material
 d. popular way of dressing
25. distort: The image is **distorted**.
 a. having more than one meaning
 b. exactly the same as something else
 c. has a badly changed shape
 d. from recent times
26. accumulate: He **accumulated** many friends.
 a. understood the value
 b. got more and more
 c. said good things about
 d. became the same as

27. abandon: He **abandoned** the project.

- a. used it for his own gain
- b. controlled in a clever way
- c. stopped working on it
- d. made it as small as possible

28. rigid: These rules are **rigid**.

- a. how good something is
- b. happening at the same time
- c. continuing for a limited time
- d. not able to be changed

29. notwithstanding : **Notwithstanding** John's feelings, Allison went to France.

- a. without knowing
- b. giving back in the same way
- c. because of
- d. not being stopped by

30. perspective: You have a good **perspective**.

- a. events that happen again and again
- b. way of seeing things
- c. group of people you know
- d. how other people see you

Source: McLean, S., & Kramer, B. (2015). The creation of a New Vocabulary Levels Test. *Shiken*, 19(2), 1-11.

Appendix 2

NVLT Vocabulary Level Test Results

<i>Students</i>	<i>Section 1 (1000)</i>	<i>Section 2 (2000)</i>	<i>Section 3 (3000)</i>	<i>Section 4 (4000)</i>	<i>Section 5 (5000)</i>	<i>Section 6 (AWL)</i>	<i>1000</i>	<i>2000</i>	<i>>3000</i>	<i>AWL</i>	<i>TOTAL</i>	<i>English level</i>
Student 1	16	22	18	18	6	9	47,60	4,31	8,11	3,00	63,02	A1
Student 2	16	13	14	15	14	16	47,60	2,55	8,30	5,33	63,78	A1
Student 3	18	23	23	21	23	27	53,55	4,50	12,93	9,00	79,99	A2
Student 4	19	23	22	21	20	23	56,53	4,50	12,16	7,67	80,86	A2
Student 5	19	23	22	21	21	23	56,53	4,50	12,36	7,67	81,05	A2
Student 6	21	15	16	13	15	24	62,48	2,94	8,49	8,00	81,91	A2
Student 7	23	20	15	14	12	11	68,43	3,92	7,92	3,67	83,92	A2
Student 8	21	21	23	17	18	22	62,48	4,11	11,20	7,33	85,12	A2
Student 9	23	20	14	7	12	20	68,43	3,92	6,37	6,67	85,38	A2
Student 10	21	22	22	22	21	23	62,48	4,31	12,55	7,67	87,00	B1
Student 11	20	24	23	24	23	29	59,50	4,70	13,51	9,67	87,38	B1
Student 12	23	19	18	16	21	15	68,43	3,72	10,62	5,00	87,76	B1
Student 13	21	23	20	22	23	26	62,48	4,50	12,55	8,67	88,19	B1
Student 14	22	21	18	21	20	22	65,45	4,11	11,39	7,33	88,29	B1
Student 15	21	24	23	22	22	26	62,48	4,70	12,93	8,67	88,78	B1
Student 16	22	21	21	22	19	24	65,45	4,11	11,97	8,00	89,53	B1
Student 17	22	22	18	19	21	26	65,45	4,31	11,20	8,67	89,62	B1
Student 18	22	22	18	19	21	26	65,45	4,31	11,20	8,67	89,62	B1
Student 19	22	22	18	19	21	26	65,45	4,31	11,20	8,67	89,62	B1
Student 20	22	22	18	19	21	26	65,45	4,31	11,20	8,67	89,62	B1
Student 21	22	21	22	20	21	25	65,45	4,11	12,16	8,33	90,06	B1
Student 22	22	23	23	21	22	24	65,45	4,50	12,74	8,00	90,70	B1
Student 23	23	22	19	21	20	22	68,43	4,31	11,58	7,33	91,65	B1
Student 24	22	24	22	22	21	27	65,45	4,70	12,55	9,00	91,70	B1
Student 25	22	23	22	23	23	26	65,45	4,50	13,13	8,67	91,75	B1
Student 26	22	21	23	24	23	28	65,45	4,11	13,51	9,33	92,41	B1
Student 27	23	22	22	24	23	25	68,43	4,31	13,32	8,33	94,39	B1
Student 28	24	19	21	21	20	22	71,40	3,72	11,97	7,33	94,42	B1
Student 29	23	23	23	23	23	25	68,43	4,50	13,32	8,33	94,58	B1
Student 30	24	19	18	21	19	25	71,40	3,72	11,20	8,33	94,65	B1
Student 31	23	23	24	22	23	27	68,43	4,50	13,32	9,00	95,25	B1
Student 32	24	23	22	23	21	23	71,40	4,50	12,74	7,67	96,31	B1
Student 33	24	23	20	22	23	26	71,40	4,50	12,55	8,67	97,12	B1
Student 34	24	23	24	23	23	27	71,40	4,50	13,51	9,00	98,42	B2
Student 35	24	24	23	24	22	27	71,40	4,70	13,32	9,00	98,42	B2
Student 36	24	24	24	24	23	28	71,40	4,70	13,71	9,33	99,14	B2
Student 37	24	24	24	24	23	28	71,40	4,70	13,71	9,33	99,14	B2

Elaborated by: Nataly Romero

Source: Vocabulary Level Test Answers by McLean, S., & Kramer, B., 2015.

Appendix 3

Individual lexical analysis

Students	K1	K2	AWL	off-list words	Lexical Diversity	Lexical Density	K1 Coverage %	K2 Coverage %	AWL Coverage %	off-list words Coverage %	Text size
Student 1	128	15	38	181	0,43	0,49	81,59	4,07	12,02	2,33	517
Student 2	108	11	34	153	0,36	0,47	80,19	5,05	8,97	5,79	535
Student 3	96	9	21	126	0,42	0,51	79,35	4,53	9,57	6,55	399
Student 4	141	18	65	224	0,43	0,55	71,47	5,51	17,51	5,51	703
Student 5	131	14	33	178	0,39	0,49	82,21	4,45	8,9	4,45	561
Student 6	101	12	25	138	0,42	0,52	78,61	4,23	9,95	7,21	403
Student 7	95	13	26	134	0,39	0,5	79,96	3,72	9,09	7,23	484
Student 8	132	21	46	199	0,4	0,54	78,76	6,53	11,08	3,64	660
Student 9	87	5	30	122	0,35	0,51	80,86	2,39	14,83	1,91	416
Student 10	100	15	28	143	0,42	0,56	77,1	4,31	13,83	4,76	441
Student 11	117	15	41	173	0,46	0,52	75,69	4,35	12,06	7,91	502
Student 12	100	8	13	121	0,31	0,51	80,92	6,11	6,87	6,11	523
Student 13	104	6	28	138	0,43	0,57	78,45	4,36	10,9	6,3	410
Student 14	91	11	29	131	0,42	0,52	79,9	4,07	12,98	3,05	393
Student 15	91	10	13	114	0,42	0,52	82,08	4,16	4,68	9,09	384
Student 16	138	14	46	198	0,43	0,49	79,33	4,42	13,43	2,83	562
Student 17	130	20	33	183	0,42	0,54	78,14	4,66	7,72	9,49	619
Student 18	95	10	30	135	0,43	0,57	77,18	3,16	10,19	9,47	401
Student 19	141	15	45	201	0,35	0,51	79,6	2,78	12,45	5,17	754
Student 20	93	13	36	142	0,49	0,53	77,66	3,72	12,23	6,38	376
Student 21	85	6	27	118	0,43	0,53	78,4	3,2	9,6	8,8	371
Student 22	86	14	31	131	0,42	0,5	74,51	5,39	12,75	7,35	408
Student 23	119	10	31	160	0,4	0,51	78,39	3,54	14,73	3,34	510
Student 24	96	12	27	135	0,38	0,55	78,51	4,62	10,44	6,43	501
Student 25	61	7	17	85	0,52	0,52	79	3,5	13	4,5	199
Student 26	118	19	35	172	0,43	0,56	78,08	5,87	10,3	5,68	513
Student 27	98	15	21	134	0,45	0,53	77,14	8,29	8,79	5,78	397
Student 28	87	7	34	128	0,43	0,49	80,22	2,79	13,93	3,06	359
Student 29	95	13	26	134	0,39	0,5	79,96	3,72	9,09	7,23	617
Student 30	103	16	13	132	0,48	0,54	80,05	5,74	5,19	9,02	361
Student 31	103	12	39	154	0,36	0,53	77,61	5,14	14,68	2,57	544
Student 32	122	7	30	159	0,39	0,49	81,5	4,3	9,35	4,86	535

Student 33	131	12	11	154	0,45	0,5	88,69	3,33	4,88	3,1	451
Student 34	111	8	23	142	0,5	0,54	80,16	2,41	9,92	7,51	373
Student 35	100	12	22	134	0,44	0,52	79,95	3,47	11,14	5,45	404
Student 36	100	16	10	126	0,42	0,51	80,47	8,85	3,12	7,55	384
Student 37	82	14	28	124	0,44	0,58	75,68	4,92	13,39	6,01	364
Student 38	101	10	20	131	0,4	0,52	84,67	3,02	10,05	2,26	398
Student 39	79	8	26	113	0,47	0,52	75,63	3,8	13,92	6,65	314

Elaborated by: Nataly Romero

Source: LexTutor Vocab profiler first analysis

Appendix 4

Individual lexical analysis after Text Compactor application

Students	K1	K2	AWL	Off-list words	TTR	Lexical density
Student 1	107	8	28	143	0,44	0,49
Student 2	91	9	30	130	0,39	0,47
Student 3	96	9	22	127	0,41	0,51
Student 4	89	15	46	150	0,5	0,58
Student 5	104	9	22	135	0,41	0,47
Student 6	101	12	25	138	0,42	0,52
Student 7	166	17	50	233	0,39	0,51
Student 8	91	16	31	138	0,42	0,54
Student 9	84	5	29	118	0,36	0,52
Student 10	91	15	26	132	0,44	0,56
Student 11	85	11	35	131	0,51	0,52
Student 12	88	8	12	108	0,35	0,52
Student 13	104	6	28	138	0,43	0,57
Student 14	91	11	29	131	0,42	0,52
Student 15	91	10	13	114	0,41	0,52
Student 16	114	10	37	161	0,47	0,5
Student 17	91	16	23	130	0,45	0,51
Student 18	95	10	30	135	0,43	0,57
Student 19	93	5	27	125	0,42	0,51
Student 20	93	13	36	142	0,49	0,53
Student 21	84	6	27	117	0,42	0,52
Student 22	86	14	31	131	0,42	0,5
Student 23	108	7	25	140	0,45	0,51
Student 24	71	10	25	106	0,37	0,55
Student 25	61	7	17	85	0,52	0,52
Student 26	96	16	28	140	0,47	0,56
Student 27	97	15	21	133	0,45	0,53
Student 28	87	7	34	128	0,43	0,49
Student 29	122	9	30	161	0,48	0,51
Student 30	103	16	13	132	0,48	0,54
Student 31	86	10	32	128	0,41	0,54
Student 32	95	5	22	122	0,38	0,49
Student 33	121	11	11	143	0,47	0,5
Student 34	111	8	24	143	0,5	0,54
Student 35	100	12	22	134	0,44	0,52
Student 36	100	16	10	126	0,42	0,51
Student 37	82	14	28	124	0,44	0,58
Student 38	101	10	20	131	0,4	0,52
Student 39	79	8	26	113	0,47	0,52

Elaborated by: Nataly Romero Source: LexTutor Vocab profiler first analysis

Appendix 5

Vocabulary Profile of Pre-service English teachers' Corpus

FRAMEWORK VP-Classic

	<u>Families</u>	<u>Types</u>	<u>Tokens</u>	<u>Percent</u>
K1 Words (1-1000):	493	875	11848	79.07%
Function:	(7149)	(47.71%)
Content:	(4699)	(31.36%)
> Anglo-Sax	(1953)	(13.03%)
K2 Words (1001-2000):	131	194	678	4.52%
> Anglo-Sax	(221)	(1.47%)
1k+2k	(83.59%)
AWL Words:	248	415	1624	10.84%
> Anglo-Sax	(50)	(0.33%)
Off-List Words:	<u>?</u>	<u>330</u>	<u>834</u>	<u>5.57%</u>
	872+?	1814	14984	100%

Words in text (tokens):	14984
Different words (types):	1814
Type-token ratio:	0.12
Tokens per type:	8.26
Lex density (content words/total)	0.52

Pertaining to onlist only

Tokens:	14150
Types:	1484
Families:	872
Tokens per family:	16.23
Types per family:	1.70
Anglo-Sax Index:	
(A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%

Lexical diversity in the corpus is low, as shown in the TTR value, which is 0.12. This result reveals that there is a lot of word repetition because of two factors. The first one is text length, as there are many repeated function words, and the second reason might be unawareness during writing to include a higher variation of content words to diversify the texts with the use of synonyms and other academic words. Lack of opportunities to experience a natural English immersion environment and developing academic writing practice might be constraints for pre-service teachers who focus primarily on grammatical structures rather than varied vocabulary use.

Lexical sophistication commonly includes less frequent words, proper nouns, and specialized vocabulary related to the writers' field of knowledge. In this study, the thematic language teaching due to specialized terminology is related to education, teaching theories, tools, and others. However, less frequent vocabulary in the corpus covers only 5.57%, which discloses a middle grade of sophistication in the production. The off-list vocabulary usually covers higher portions in academic texts, as mentioned by Coxhead (2000, p. 22)

The corpus's lexical profile displays the coverage of different frequency word bands. K1 and K2 words cover 83.59% of the whole corpus, meaning that pre-service English teachers produced texts which do not reflect their receptive vocabulary knowledge. The coverage percentage of the academic word list in the corpus reports 10.84% that remains in the average quantity expected for academic writing. As the theory suggests, receptive vocabulary is higher than productive.

Appendix 6

Frequency word list arranged from the corpus (100 more frequent words)

- | | | |
|------------------|--------------------|--------------------------|
| 1. 1167_the | 35. 55_them | 69. 31_teacher |
| 2. 645_of | 36. 54_because | 70. 31_work |
| 3. 563_to | 37. 53_not | 71. 30_classroom |
| 4. 484_and | 38. 46_skills | 72. 30_information |
| 5. 478_in | 39. 46_through | 73. 30_tandem |
| 6. 313_is | 40. 45_have | 74. 29_development |
| 7. 306_a | 41. 45_since | 75. 29_people |
| 8. 266_that | 42. 44_education | 76. 28_also |
| 9. 241_learning | 43. 44_learn | 77. 28_importance |
| 10. 212_english | 44. 44_new | 78. 28_tools |
| 11. 207_this | 45. 42_by | 79. 27_all |
| 12. 206_language | 46. 42_these | 80. 27_been |
| 13. 168_students | 47. 41_educational | 81. 27_but |
| 14. 166_for | 48. 41_important | 82. 27_communicatio
n |
| 15. 163_it | 49. 41_reading | 83. 27_help |
| 16. 146_be | 50. 40_resources | 84. 27_therefore |
| 17. 120_teaching | 51. 39_i | 85. 27_time |
| 18. 118_will | 52. 39_improve | 86. 26_about |
| 19. 114_as | 53. 39_strategies | 87. 26_e |
| 20. 111_are | 54. 39_virtual | 88. 26_most |
| 21. 99_use | 55. 38_topic | 89. 26_motivation |
| 22. 92_research | 56. 37_other | 90. 26_must |
| 23. 88_their | 57. 36_knowledge | 91. 26_problem |
| 24. 84_on | 58. 35_know | 92. 25_finally |
| 25. 84_with | 59. 34_develop | 93. 25_platforms |
| 26. 81_can | 60. 34_different | 94. 25_there |
| 27. 78_they | 61. 34_so | 95. 25_used |
| 28. 74_teachers | 62. 33_more | 96. 25_very |
| 29. 73_which | 63. 33_student | 97. 24_both |
| 30. 72_or | 64. 33_we | 98. 24_how |
| 31. 70_an | 65. 32_one | 99. 24_order |
| 32. 66_process | 66. 32_when | 100.24_second |
| 33. 59_way | 67. 31_at | |
| 34. 55_has | 68. 31_such | |

Elaborated by: Nataly Romero

Source: The compleat Word lister Cobb, T. *Web compleat lister* [accessed 15 March 2021 from <http://www.lex Tutor.ca/freq/comp>]

The following table shows the most frequent content words in the corpus analyzed by employing frequency analysis from Lextutor.

Nouns	Freq.	Verbs	Freq.	Adjectives	Freq.	Adverbs	Freq.
Learning	241	use	99	English	212	so	34
Language	206	research	92	new	44	more	33
Students	168	process	66	educational	41	also	28
teaching	120	have	45	important	41	finally	25
teachers	74	learn	44	virtual	39	very	25

Table 6. Five most frequent content words from the corpus

Elaborated by: Nataly Romero

Source: Frequency Analysis

Table 6 illustrates how students employed basic vocabulary consisted mainly of high-frequency words, and only two of them fit into the AWL. These words can be replaced with more formal lexical units. For instance, instead of *use*, students might write *utilize*, *new* could be changed by *contemporary* to enrich the writing.

Appendix 7

Results of Corpus analysis with Cohmetrix

Number	Label	Label V2.x	Text	Full description
Descriptive				
1	DESPC	READNP	256	Paragraph count, number of paragraphs
2	DESSC	READNS	503	Sentence count, number of sentences
3	DESWC	READNW	14966	Word count, number of words
4	DESPL	READAPL	1.965	Paragraph length, number of sentences in a paragraph, mean
5	DESPLd	n/a	1.097	Paragraph length, number of sentences in a paragraph, standard deviation
6	DESSL	READASL	29.753	Sentence length, number of words, mean
7	DESSLd	n/a	18.989	Sentence length, number of words, standard deviation
8	DESWLsy	READASW	1.689	Word length, number of syllables, mean
9	DESWLsyd	n/a	1.007	Word length, number of syllables, standard deviation
10	DESWLlt	n/a	5.130	Word length, number of letters, mean
11	DESWLltd	n/a	2.950	Word length, number of letters, standard deviation
Text Easability Principle Component Scores				
12	PCNARz	n/a	-0.462	Text Easability PC Narrativity, z score
13	PCNARp	n/a	32.280	Text Easability PC Narrativity, percentile
14	PCSYNz	n/a	-0.756	Text Easability PC Syntactic simplicity, z score
15	PCSYNp	n/a	22.660	Text Easability PC Syntactic simplicity, percentile
16	PCCNCz	n/a	-0.922	Text Easability PC Word concreteness, z score
17	PCCNCp	n/a	17.880	Text Easability PC Word concreteness, percentile
18	PCREFz	n/a	0.885	Text Easability PC Referential cohesion, z score
19	PCREFp	n/a	81.060	Text Easability PC Referential cohesion, percentile
20	PCDCz	n/a	1.697	Text Easability PC Deep cohesion, z score
21	PCDCp	n/a	95.450	Text Easability PC Deep cohesion, percentile
22	PCVERBz	n/a	0.254	Text Easability PC Verb cohesion, z score
23	PCVERBp	n/a	59.870	Text Easability PC Verb cohesion, percentile
24	PCCONNz	n/a	-1.627	Text Easability PC Connectivity, z score
25	PCCONNp	n/a	5.260	Text Easability PC Connectivity, percentile
26	PCTEMPz	n/a	0.226	Text Easability PC Temporality, z score
27	PCTEMPp	n/a	58.710	Text Easability PC Temporality, percentile
Referential Cohesion				
28	CRFNO1	CRFBN1um	0.580	Noun overlap, adjacent sentences, binary, mean
29	CRFAO1	CRFBA1um	0.661	Argument overlap, adjacent sentences, binary, mean
30	CRFSO1	CRFBS1um	0.709	Stem overlap, adjacent sentences, binary, mean
31	CRFNOa	CRFBNaum	0.454	Noun overlap, all sentences, binary, mean
32	CRFAOa	CRFBAaum	0.548	Argument overlap, all sentences, binary, mean
33	CRFSOa	CRFBSaum	0.604	Stem overlap, all sentences, binary, mean
34	CRFCWO1	CRFPC1um	0.127	Content word overlap, adjacent sentences, proportional, mean
35	CRFCWO1d	n/a	0.114	Content word overlap, adjacent sentences, proportional, standard deviation

36	CRFCWOa	CRFPCaum	0.098	Content word overlap, all sentences, proportional, mean
37	CRFCWOad	n/a	0.108	Content word overlap, all sentences, proportional, standard deviation
LSA				
38	LSASS1	LSAassa	0.311	LSA overlap, adjacent sentences, mean
39	LSASS1d	LSAassd	0.177	LSA overlap, adjacent sentences, standard deviation
40	LSASSp	LSApssa	0.293	LSA overlap, all sentences in paragraph, mean
41	LSASSpd	LSApssd	0.168	LSA overlap, all sentences in paragraph, standard deviation
42	LSAPP1	LSAppa	0.425	LSA overlap, adjacent paragraphs, mean
43	LSAPP1d	LSAppd	0.186	LSA overlap, adjacent paragraphs, standard deviation
44	LSAGN	LSAGN	0.407	LSA given/new, sentences, mean
45	LSAGNd	n/a	0.095	LSA given/new, sentences, standard deviation
Lexical Diversity				
46	LDTTRc	TYPTOKc	0.236	Lexical diversity, type-token ratio, content word lemmas
47	LDTTRa	n/a	0.122	Lexical diversity, type-token ratio, all words
48	LDMTLD	LEXDIVTD	63.447	Lexical diversity, MTLT, all words
49	LDVOCD	LEXDIVVD	73.864	Lexical diversity, VOCD, all words
Connectives				
50	CNCAll	CONi	96.619	All connectives incidence
51	CNCCaus	CONCAUSi	35.614	Causal connectives incidence
52	CNCLogic	CONLOGi	45.704	Logical connectives incidence
53	CNCADC	CONADVCONi	10.958	Adversative and contrastive connectives incidence
54	CNCTemp	CONTEMPi	17.974	Temporal connectives incidence
55	CNCTempx	CONTEMPEXi	16.170	Expanded temporal connectives incidence
56	CNCAdd	CONADDi	50.715	Additive connectives incidence
57	CNCPos	n/a	0	Positive connectives incidence
58	CNCNeg	n/a	0	Negative connectives incidence
Situation Model				
59	SMCAUSv	CAUSV	11.560	Causal verb incidence
60	SMCAUSvp	CAUSVP	29.868	Causal verbs and causal particles incidence
61	SMINTEp	INTEi	6.281	Intentional verbs incidence
62	SMCAUSr	CAUSC	1.575	Ratio of casual particles to causal verbs
63	SMINTEr	INTEC	3.358	Ratio of intentional particles to intentional verbs
64	SMCAUSlsa	CAUSLSA	0.092	LSA verb overlap
65	SMCAUSwn	CAUSWN	0.465	WordNet verb overlap
66	SMTEMP	TEMPta	0.832	Temporal cohesion, tense and aspect repetition, mean
Syntactic Complexity				
67	SYNLE	SYNLE	5.022	Left embeddedness, words before main verb, mean
68	SYNNP	SYNNP	0.995	Number of modifiers per noun phrase, mean
69	SYNMEDpos	MEDwtm	0.610	Minimal Edit Distance, part of speech
70	SYNMEDwrd	MEDawm	0.851	Minimal Edit Distance, all words
71	SYNMEDlem	MEDalm	0.837	Minimal Edit Distance, lemmas
72	SYNSTRUTa	STRUTa	0.064	Sentence syntax similarity, adjacent sentences, mean
73	SYNSTRUTt	STRUTt	0.058	Sentence syntax similarity, all combinations, across paragraphs, mean
Syntactic Pattern Density				

74	DRNP	n/a	364.894	Noun phrase density, incidence
75	DRVP	n/a	214.018	Verb phrase density, incidence
76	DRAP	n/a	20.246	Adverbial phrase density, incidence
77	DRPP	n/a	125.017	Preposition phrase density, incidence
78	DRPVAL	AGLSPSVi	10.424	Agentless passive voice density, incidence
79	DRNEG	DENNEGi	4.343	Negation density, incidence
80	DRGERUND	GERUNDi	24.856	Gerund density, incidence
81	DRINF	INFi	28.865	Infinitive density, incidence
Word Information				
82	WRDNOUN	NOUNi	260.123	Noun incidence
83	WRDVERB	VERBi	126.153	Verb incidence
84	WRDADJ	ADJi	104.169	Adjective incidence
85	WRDADV	ADVi	33.943	Adverb incidence
86	WRDPRO	DENPRPi	37.285	Pronoun incidence
87	WRDPRP1s	n/a	3.742	First person singular pronoun incidence
88	WRDPRP1p	n/a	3.942	First person plural pronoun incidence
89	WRDPRP2	PRO2i	1.069	Second person pronoun incidence
90	WRDPRP3s	n/a	1.336	Third person singular pronoun incidence
91	WRDPRP3p	n/a	15.168	Third person plural pronoun incidence
92	WRDFRQc	FRCLacwm	2.169	CELEX word frequency for content words, mean
93	WRDFRQa	FRCLaewm	3.033	CELEX Log frequency for all words, mean
94	WRDFRQmc	FRCLmcsm	0.939	CELEX Log minimum frequency for content words, mean
95	WRDAOAc	WRDAacwm	393.292	Age of acquisition for content words, mean
96	WRDFAMc	WRDFacwm	571.042	Familiarity for content words, mean
97	WRDCNCc	WRDCacwm	351.917	Concreteness for content words, mean
98	WRDIMGc	WRDIacwm	382.930	Imagability for content words, mean
99	WRDMEAc	WRDMacwm	424.851	Meaningfulness, Colorado norms, content words, mean
100	WRDPOLc	POLm	3.599	Polysemy for content words, mean
101	WRDHYPn	HYNOUNaw	6.758	Hypernymy for nouns, mean
102	WRDHYPv	HYVERBaw	1.525	Hypernymy for verbs, mean
103	WRDHYPnv	HYPm	1.914	Hypernymy for nouns and verbs, mean
Readability				
104	RDFRE	READFRE	33.746	Flesch Reading Ease
105	RDFKGL	READFKGL	15.944	Flesch-Kincaid Grade level
106	RDL2	L2	13.683	Coh-Metrix L2 Readability

Elaborated by: Nataly Romero

Source: Coh-metrix output retrieved on March 22nd, 2021.

Cohmetrix analysis provided indexes that clarify the cohesion and difficulty of a text and complement the information obtained in readability scores. The corpus was studied in eight components and they are summarized in the following figure:

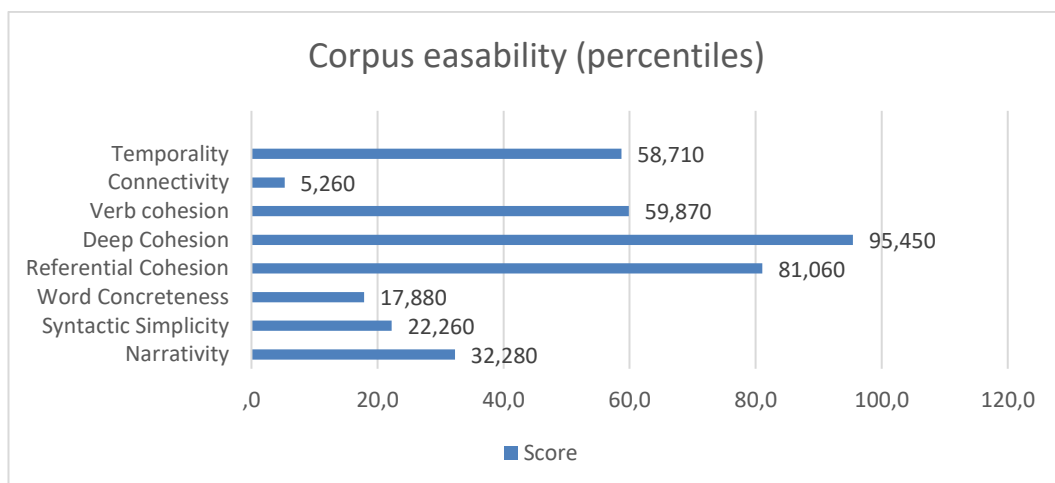


Figure 2. Easability components score summary

Elaborated by: Nataly Romero

Source: Results of Cohmetrix analysis

The temporality is associated with the tense and aspect of the text. The corpus is average in temporality. The connectivity is very low, there are not many adversative, additive, and comparative connectives to express relations in the corpus. There are a few words to add information, connect, compare and contrast ideas. Consequently, the reader requires an additional inference to process the text. Verb cohesion is average, revealing a challenging text for the reader. The corpus is high in deep cohesion, reflecting more explicit causal and intentional relationships as needed by the text. It is high in referential cohesion, meaning that there is overlapping in explicit words and ideas between sentences. Word concreteness is very low, suggesting a high volume of abstract words and low imageability, making the reading more difficult. The corpus is low in syntactic simplicity. Thus, sentence structures are complex, increasing the reading difficulty of the product. Narrativity in the corpus is average; academic writing characterizes as being more descriptive and analytical. Some writers employed first-person pronouns, which is not common in formal texts.

Readability formulas provided by Cohmetrix are Flesch reading ease, Flesch-Kincaid grade level, and RDL2, second language readability score (**See appendix 7**). As highest the Flesch reading ease index as easier to read the text, the corpus sites into difficult to read band with a score of 33.75. Flesch-Kincaid grade level works inversely; if the values are high, the texts are more difficult to read. The index is related to the grade level of education in the U.S. required to understand the text.

The score was 15.94 college graduates would readily understand this corpus. The second language readability score shows that the corpus is very difficult to read, with a score of 13.68.

Appendix 8

Results of individual analysis with Cohmetrix

Students	Narrativity	Syntactic Simplicity	Word Concreteness	Referential Cohesion	Deep Cohesion	Verb cohesion	Connectivity	Temporality	Flesch reading ease	Flesch Kincaid	Cohmetrix L2 readability
Student 1	51,6	20,9	2,39	46,81	97,78	51,6	37,45	58,32	42,652	13,412	13,232
Student 2	28,1	4,36	27,76	95,45	59,48	81,59	0,24	28,77	32,066	17,708	16,49
Student 3	31,56	9,51	59,87	85,77	99,04	69,85	0,06	26,11	40,06	14,807	13,313
Student 4	7,93	8,38	35,2	67,36	32,28	50	0,52	34,46	2,232	20,516	7,331
Student 5	55,17	0	37,83	98,64	99,92	92,79	31,21	95,15	18,187	23,392	17,62
Student 6	55,17	7,64	31,21	93,82	77,64	10,2	6,43	1,07	23,21	20,034	11,277
Student 7	77,94	0,16	59,48	99,71	48,4	10,38	0,24	74,22	10,199	25,571	15,334
Student 8	31,92	6,06	35,2	72,57	92,79	59,48	46,02	89,25	31,81	16,543	9,319
Student 9	28,43	17,36	19,77	86,21	65,91	80,51	37,07	52,79	41,74	13,224	10,433
Student 10	37,45	4,46	27,09	82,12	75,8	35,2	13,79	20,9	17,161	19,583	10,913
Student 11	26,43	44,83	37,45	3,59	94,74	35,2	9,18	95,15	44,575	10,644	2,829
Student 12	67	16,6	24,2	97,06	29,12	13,14	15,35	8,69	46,91	13,743	19,251
Student 13	27,09	32,28	36,32	56,36	97,06	67	30,85	19,22	40,43	13,29	18,505
Student 14	33,72	29,46	6,94	93,06	95,45	51,99	0,09	46,02	37,715	13,756	22,593
Student 15	48,01	11,7	27,76	68,79	97,5	64,06	0	73,42	45,339	13,767	15,583
Student 16	58,71	0,18	6,68	69,5	100	95,73	3,92	99,09	35,247	16,216	15,285
Student 17	58,32	0,3	16,11	98,68	79,39	44,83	37,83	77,34	19,437	20,386	22,28
Student 18	32,28	9,68	27,09	65,54	84,61	44,83	16,11	60,26	25,206	18,087	22,019
Student 19	30,85	0,06	64,06	99,81	97,26	63,68	27,09	4,65	13,322	23,295	18,794
Student 20	12,92	38,21	56,36	28,77	65,54	54,38	2,33	26,11	34,946	15,11	6,99
Student 21	31,21	0,25	50	96,41	82,38	68,44	2,94	35,94	12,899	23,167	15,74

Student 22	76,11	0,43	6,43	95,05	99,96	81,86	15,39	99,34	34,205	16,285	20,959
Student 23	38,59	8,23	46,41	70,54	71,23	51,6	18,67	26,76	35,22	16,144	12,344
Student 24	33	1,29	76,42	99,22	30,85	14,69	1,62	88,3	24,107	19,586	10,118
Student 25	35,57	2,22	40,13	89,44	80,51	46,41	60,26	94,41	39,514	14,846	16,282
Student 26	29,81	12,51	19,77	78,52	87,49	38,97	2,81	63,68	27,15	16,014	14,323
Student 27	45,62	19,22	17,88	71,9	95,54	15,39	0,13	93,32	36,222	15,717	10,227
Student 28	11,7	44,43	24,83	65,91	69,5	68,08	17,62	4,27	42,644	13,664	13,388
Student 29	37,83	1,66	47,21	69,5	99,96	68,79	0,21	59,87	30,45	17,438	14,289
Student 30	37,45	36,69	51,6	61,03	99,87	31,21	1,97	0	47,262	12,304	15,14
Student 31	44,04	0,96	42,86	99,31	99,71	16,85	31,92	85,77	5,073	24,166	15,517
Student 32	32,64	5,37	49,2	91,92	72,57	91,31	31,92	70,54	41,327	14,771	16,596
Student 33	64,43	1,79	42,86	57,14	99,92	92,51	0,1	15,62	43,616	17,038	20,385
Student 34	30,5	38,97	8,08	27,43	97,5	43,25	11,51	82,38	43,025	12,42	11,454
Student 35	29,12	17,88	19,22	57,53	99,93	80,23	1,36	93,19	36,421	13,198	13,187
Student 36	50	45,62	29,46	56,75	98,12	31,92	0	6,55	49,625	11,665	17,631
Student 37	53,59	4,55	41,68	95,54	99,82	16,85	5,26	91,62	30,11	16,015	16,755
Student 38	68,08	0,28	35,57	98,75	100	61,03	82,38	98,57	23,474	21,63	21,468
Student 39	16,35	4,09	15,62	46,02	87,9	86,43	17,36	24,51	17,608	18,604	10,876

Elaborated by: Nataly Romero

Source: Cohmetrix results of individual texts analysis

Appendix 9

Statistical Analysis of data

9.1 Anova one way

The following chart describes p-value for the relation between Flesch reading ease, lexical diversity, lexical density and lexical sophistication.

SUMMARY				
<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Flesch reading ease	39	1222,396	31,34348718	155,560663
Lexical diversity	39	16,98	0,435384615	0,00167287
Lexical density	39	20,39	0,522820513	0,00072078
Lexical Sophisitcation	39	5196	133,2307692	467,234818

ANOVA						
<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between groups	463439,574	3	154479,858	992,166894	*0,0000	2,664106703
Within groups	23666,3192	152	155,6994686			
Total	487105,893	155				

*Scientific E Notation:1.4417E-99

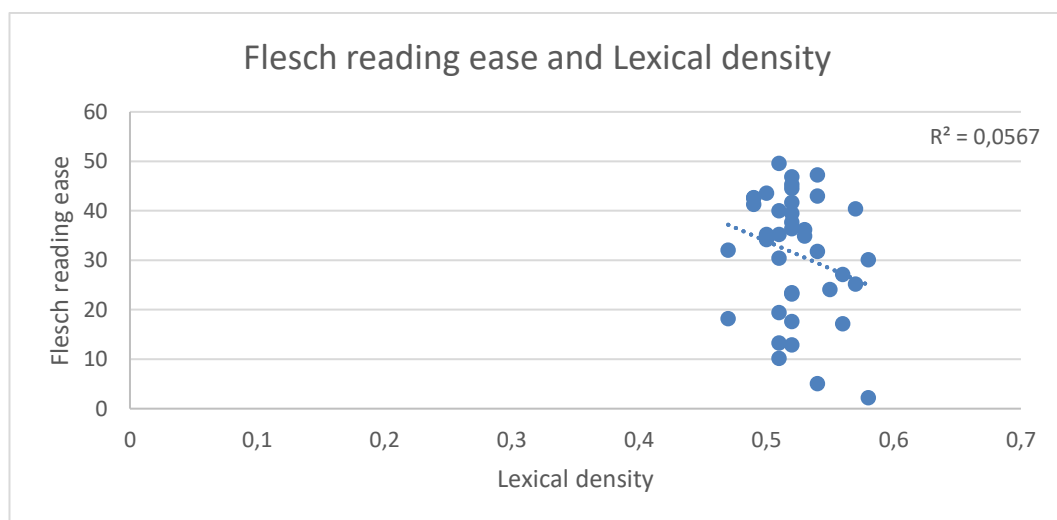
The statistical analysis of data was developed with Microsoft Excel; this software package allowed performing single-factor ANOVA calculus between the variables. Readability, represented by values of Flesch reading ease of texts measured individually. Lexical richness considers some aspects in this, lexical diversity values, lexical density indexes, and the number of off-list words on each text. Results of the comparison answer the main research question; Is lexical richness related to the readability of English degree students' writing? According to the ANOVA one-way analysis results, there is a significant relationship between the components of lexical richness and the readability of English degree students. The significance level is represented with a p-value <0.001.

9.2 t-Test for Flesch reading ease and lexical density

	<i>Flesch reading ease</i>	<i>Lexical density</i>
Mean	31,34348718	0,522820513
Variance	155,5606628	0,000720783
Observations	39	39
Pearson Correlation	-0,238170368	
Hypothesized Mean Difference	0	
Df	38	
T Stat	15,42413784	
P(T<=t) one-tail	*0.0000	
T Critical one-tail	1,68595446	
P(T<=t) two-tails	*0.0000	
T Critical two-tails	2,024394164	

*Scientific E Notation P (T<=t) one-tail 3,0207E-18

*Scientific E Notation P(T<=t) two-tails 6,04139E-18



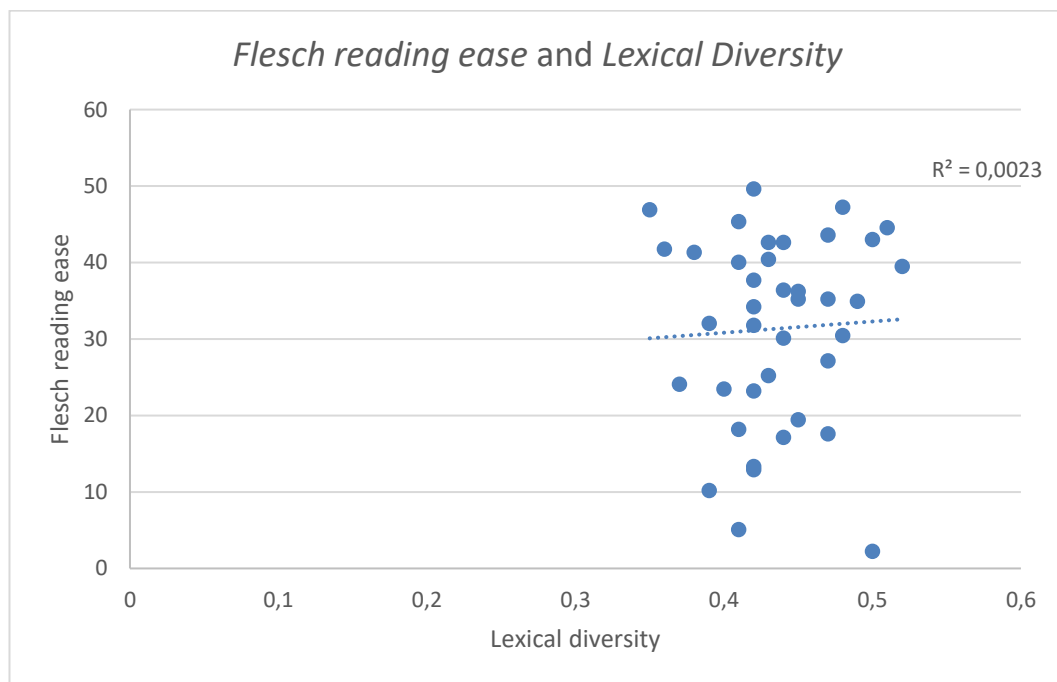
Pearson's correlation analysis supports the understanding of the relationship between these elements. Flesch reading ease (readability score) and lexical density of the texts have a low negative correlation (-0.23) with a p-value < 0.001. So, A small relationship exists between the variables, and it has statistical significance. An increase of lexical density in the text might moderately decrease its readability. The graphic representation releases the R² value, which describes that 5% of readability variation is explained by lexical density in the examined texts.

9.3 t-Test for Flesch reading ease and Lexical Diversity

	<i>Flesch reading ease</i>	<i>Lexical Diversity</i>
Mean	31,34348718	0,435641026
Variance	155,5606628	0,001662078
Observations	39	39
Pearson Correlation	0,034695776	
Hypothesized Mean Difference	0	
Df	38	
T Stat	15,47740302	
P(T<=t) one-tail	*0.0000	
T Critical one-tail	1,68595446	
P(T<=t) two-tails	*0.0000	
T Critical two-tails	2,024394164	

*Scientific E Notation P (T<=t) one-tail 2,6967E-18

*Scientific E Notation P(T<=t) two-tails 5,3934E-18



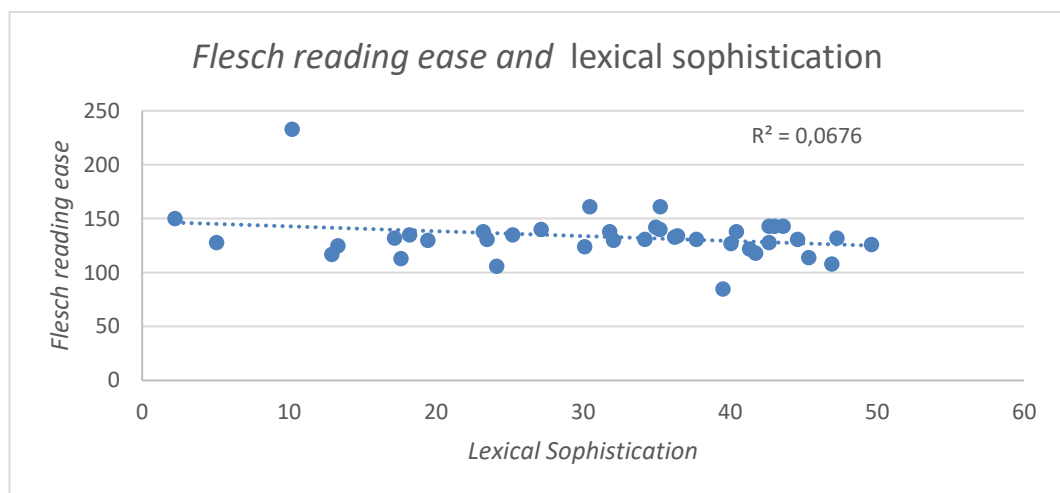
Flesch reading ease and lexical diversity has a very low positive correlation with a significance of p-value < 0.001. Then, the linear correlation between these variables is very low as Pearson's coefficient is near zero (0.03), data is highly significant to suggest that lexical diversity has a slight influence over texts' readability.

9.4 t-Test for Flesch reading ease and Lexical sophistication

	<i>Flesch reading ease</i>	<i>Lexical sophistication</i>
Mean	31,34348718	133,2307692
Variance	155,5606628	467,2348178
Observations	39	39
Pearson Correlation	-0,260034757	
Hypothesized Mean Difference	0	
Df	38	
T Stat	-23,03501297	
P(T<=t) one-tail	*0.0000	
T Critical one-tail	1,68595446	
P(T<=t) two-tails	*0.0000	
T Critical two-tails	2,024394164	

*Scientific E Notation P(T<=t) one-tail 3,13901E-24

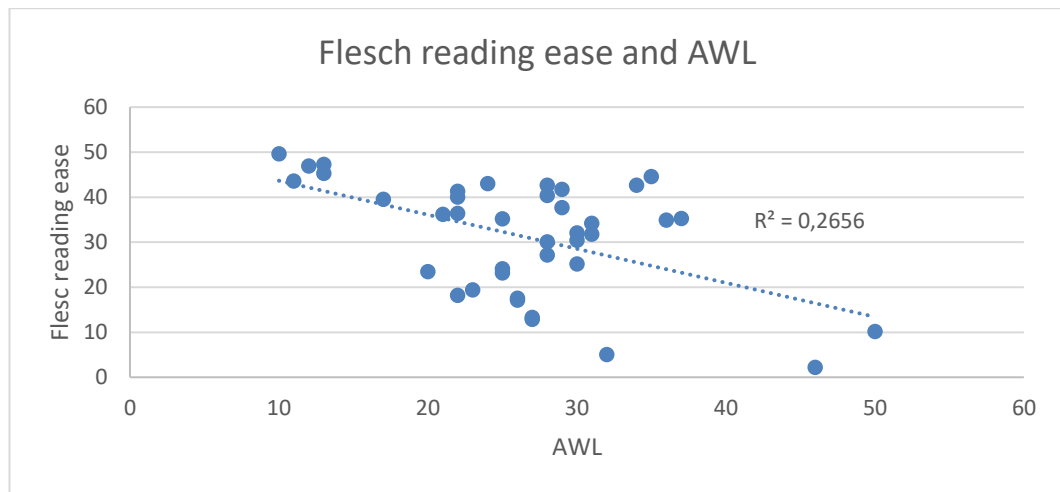
*Scientifi E Notation P (T<=t) two-tails 6,27803E-24



Flesch reading ease and Lexical sophistication display a low negative correlation (-0.26) with a p-value < 0.001. Consequently, data is significant enough to disclose a weak relation between lexical sophistication in the text and its readability. R² value elucidates that at least 7% of readability variation is explained by lexical sophistication.

9.5 t-Test for Flesch reading ease and the number of AWL families in the text

	<i>Flesch reading ease</i>	<i>AWL</i>
Mean	31,34348718	26,2820513
Variance	155,5606628	72,2604588
Observations	39	39
Pearson Correlation	-0,515389061	
Hypothesized Mean Difference	0	
Df	38	
T Stat	1,721559087	
P(T<=t) one-tail	0,05	
T Critical one-tail	1,68595446	
P(T<=t) two-tails	0,09	
T Critical two-tails	2,024394164	



ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	499,554	1	499,5536002	4,385	0,04	3,966759784
Within groups	8657,203	76	113,9105608			
Total	9156,756	77				

As has been noted, lexical richness is not the only variable that influences the readability of pre-service teachers' writing. Other elements also may affect the ease of reading to explain the model. Comparing Flesch reading ease and the number of AWL families in the text, there is moderate negative linear relation (-0.52). The

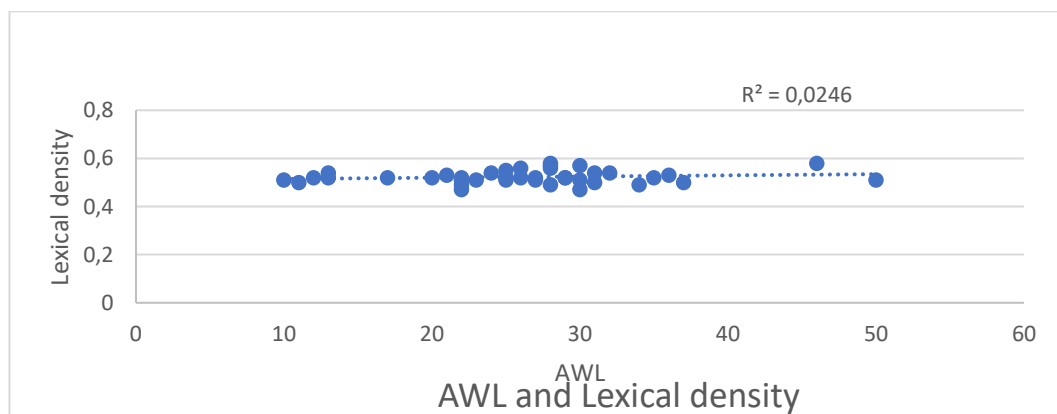
significance level was calculated with an ANOVA analysis to avoid errors achieving a p-value of 0.04, which offers a statistical significance of the relationship. Academic Vocabulary might have a good influence on the readability of texts. R^2 suggests that 27% of readability variation is explained by the number of families from the text's academic word list.

9.6 t-Test for the number of Academic word list families in the text and lexical density

	<i>AWL</i>	<i>Lexical density</i>
Mean	26,28205128	0,522820513
Variance	72,26045884	0,000720783
Observations	39	39
Pearson Correlation	0,16	
Hypothesized Mean Difference	0	
Df	38	
T Stat	18,93336081	
P(T<=t) one-tail	*0.0000	
T Critical one-tail	1,68595446	
P(T<=t) two-tails	*0.0000	
T Critical two-tails	2,024394164	

*Scientific E Notation P(T<=t) one-tail 3,01123E-21

*Scientific E Notation P(T<=t) two-tails 6,02245E-21



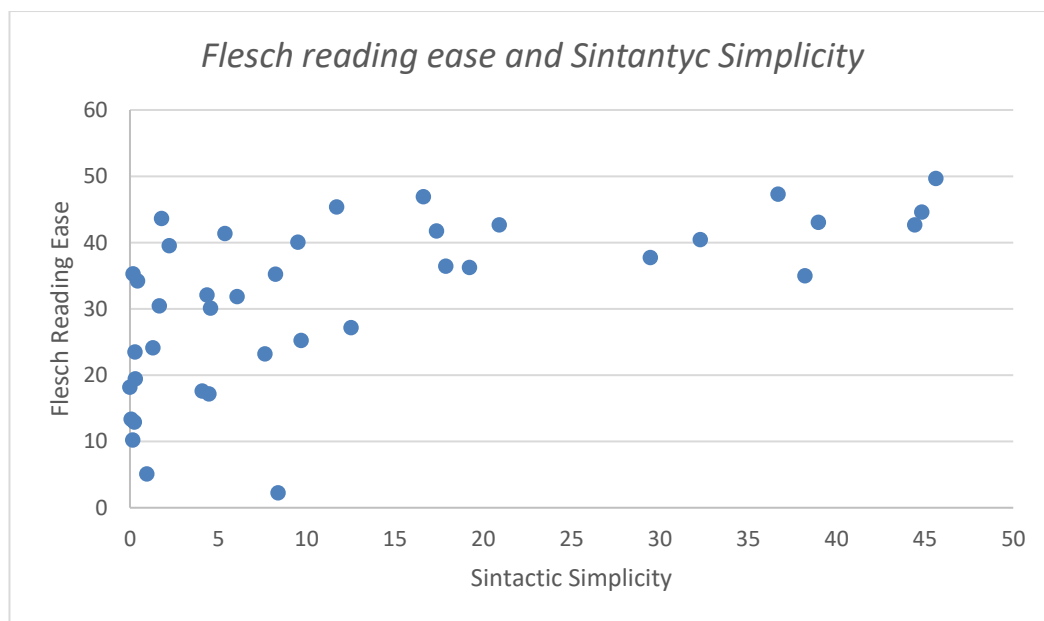
In the correlation between lexical density and academic vocabulary, the relationship is weaker (0.16) but, in this case, positive. P-value < 0.001 reveals statistical significance in the analyzed data. Academic words also show to be linked with lexical density, thus, to readability in a shallow linear degree.

9.6 t-Test of Flesch reading ease and Syntactic Simplicity index of the text

	<i>Syntactic Simplicity</i>	<i>Flesch reading ease</i>
Mean	13,04025641	31,34348718
Variance	215,797671	155,5606628
Observations	39	39
Pearson Correlation	0,608020634	
Hypothesized Mean Difference	0	
Df	38	
T Stat	-9,378132089	
P(T<=t) one-tail	*0.0000	
T Critical one-tail	1,68595446	
P(T<=t) two-tails	*0.0000	
T Critical two-tails	2,024394164	

*Scientific E Notation P(T<=t) one-tail 9,85669E-12

*Scientific E Notation P(T<=t) two-tails 1,97134E-11



Finally, an element that displays a moderate positive correlation with the readability of texts is syntactic simplicity. Pearson's coefficient (0.61) indicates a significant relation with a p-value < 0.001.

9.7 Multiple regression between readability and lexical density, lexical sophistication, syntactic simplicity.

<i>Regression Statistics</i>	
Multiple R	0,71847072
R Square	0,51620017
Adjusted R Square	0,47473161
Standard Error	9,03941913
Observations	39

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	3051,41675	1017,13892	12,44798979	*0,0000
Residual	35	2859,88844	81,7110982		
Total	38	5911,30519			

*Scientific E Notation 1,07594E-05

	<i>Coefficients</i>	<i>Standard Error</i>	<i>T Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	117,72	30,46	3,86	0,0005	55,89	179,56	55,89	179,56
Lexical density	-142,28	54,90	-2,59	0,0138	-253,73	-30,83	-253,73	-30,83
Lexical sophistication	-0,14	0,07	-2,09	0,0442	-0,28	0,00	-0,28	0,00
Syntactic Simplicity	0,53	0,10	5,28	*0,0000	0,33	0,73	0,33	0,73

*Scientific E Notation 6,87E-06

Although this feature is not related to the lexical richness, the author considered its influence for instruction purposes. In a multiple regression analysis between the dependent variable, readability, and some independent variables such as lexical density, lexical sophistication, and syntactic complexity, there was a strong positive correlation (0.71). The significance given by p-values under 0.05 in all the variables is acceptable, validating the model. Adjusted R² shows that these independent variables explain 47% of readability variation. Therefore, readability may be improved by enhancing these elements in a text.

Appendix 10

Answers to the activities in the booklet

Answers:

Pre-task 1

Sentences a,b,d,f require improvement. These sentences present poor academic style in informality, first-person writing, vagueness, contraction use, and wordiness.

Pre-task 2

Teacher's feedback

Pre-task 3

1. While, whereas
2. Both
3. Differs from
4. Whereas, While
5. On the other hand
6. Likewise

Pre-task 4

1. Sufficient
2. Together
3. Of the fact
4. Almost
5. Forward

Pre-task 5

Teacher's feedback

Pre-task 6.1

a. theories b. topic c. opportunity d. culture e. theme

Pre-task 6.2

Nouns		Verbs	
area	field	accelerate	speed up
authority	source	achieve	reach
behavior	conduct	alter	change
beliefs	ethics	analyze	take apart
benefit	advantage	assist	help
category	type	attach	join
component	part	challenge	question
concept	idea	claim	insist
controversy	argument	clarify	explain
drawback	disadvantage	concentrate on	focus on
expansion	increase	confine	limit
feeling	emotion	develop	evolve
framework	structure	eliminate	remove
goal	target	evaluate	examine
hypothesis	theory	found	establish

interpretation	explanation	maintain	keep
issue	topic	predict	forecast
method	system	prohibit	ban
option	possibility	quote	cite
quotation	citation	raise	increase
results	findings	reduce	decrease
statistics	figures	respond	reply
study	research	retain	suggestion
trend	tendency	show	demonstrate
output	production	strengthen	reinforce

Pre-task 7

- a. significant b.alternative c.consistent d.possible

Pre-task 8.1

- There was a significant increment in the number of people learning English as a second language in the 15th century.
- The application of teaching strategies revealed the enhancement of oral production.
- The coverage of lexical sophistication in the texts is irregular.
- The increment of lexical density by 5% in the text caused a diminution of readability.

Pre-task 8.2

Teacher's feedback

Pre-task 8.3

Teacher's feedback

Evaluation

Teacher's feedback

Appendix 11

Academic Words to be used with the strategies

List 1

accurate adjacent advocate allocate alter ambiguous analogous annual

List 2

appendix arbitrary assemble assure attach behalf bias bulk

List 3 cease core criterion currency decline deduct denote deviate

List 4

diminish displace dispose distinct empirical fee fluctuate

forthcoming founded fund grant guideline hierarchical ideology

implicate incline income induce infer inherent

List 5

injure input intermediate invoke isolate layer offset ongoing outcome

overall overlap overseas paradigm paragraph

parameter partner passive perceive percent regime

List 6

straightforward subordinate subsidiary substitute

successive summary supplement survive suspend sustain thereby

trace transition ultimate undergo underlie undertake welfare

whereby widespread

Appendix 12

Experts' Evaluation

TECHNICAL UNIVERSITY OF COTOPAXI

GRADUATE DEPARTMENT

Master's degree in Applied Linguistics to

Teaching English as a Foreign Language

PROPOSAL VALIDATION

1. Research proposal data:

Author: Romero Mayorga Nataly Monserrat

Topic: Booklet of strategies for academic writing improvement based on lexical and readability analysis of written corpus.

Objective: To design a booklet with strategies for pre-service English teachers based on lexical and readability analysis of written corpus.

2. Evaluator's information

Evaluator's name:	
ID number:	
Academic degree:	
Senescyt registration number	
Current job:	
Phone number:	
e-mail:	

Evaluation

Place an X on the square.

Criteria	Excellent	Good	Terrible
a) The teaching material constitutes a valid, current and relevant contribution related to the field.			
b) The teaching material is the result of an advanced research process, its content is the product of a complete conceptual development and critical contrast with other related researches.			
c) It is properly structured and argued (statement of the problem, methodology and results) in relation to the topic.			
d) The originality of the contributions and reflections of the author give added value to the proposal.			
e) The references are relevant and up-to-date.			
f) The research topic is appropriate.			
g) The proposal has the expected qualities for the level of training (appropriate language, vocabulary, spelling, etc.)			
h) The graphic illustrations that are in the dissertation (all kind of images and tables) are relevant, they clarify and provide significance.			
i) The dissertation embraces a clear and precise introduction on the objectives and issues along dissertation.			
j) The length of the dissertation is appropriate depending on the complexity of the topic, the objectives and the readers.			
k) The dissertation provides contributions regarding methodological proposals, approach, and conceptualization.			
l) The objectives in the introduction are met, so that there is harmony between objectives and results.			

Comment on the following statements, please.

1. TEMPORALITY: Is the proposal the result of an advanced research process, which means that it shows a methodological structure (problem, methodology and application)?
2. CONTENT: The content of the proposal is structured and written in an appropriate way to be understood and discussed by the educational community, and researchers?
3. SELECTIVITY: Can this proposal be considered a valid and significant contribution related to the field?

4. Impact. What is the impact of this research? (Place an X on the square)

Local	<input type="checkbox"/>
Regional	<input type="checkbox"/>
Nacional	<input type="checkbox"/>
Internacional	<input type="checkbox"/>

5. General comments and recommendations for the Author.

Evaluator's signature

I.D.....

TECHNICAL UNIVERSITY OF COTOPAXI
GRADUATE DEPARTMENT
Master's degree in Applied Linguistics to
Teaching English as a Foreign Language
PROPOSAL VALIDATION

1. Research proposal data:

Author: Romero Mayorga Nataly Monserrat

Topic: Booklet of strategies for academic writing improvement based on lexical and readability analysis of written corpus.

Objective: To design a booklet with strategies for pre-service English teachers based on lexical and readability analysis of written corpus.

2. Evaluator's information

Evaluator's name:	Patricia Marcela Chacón Porras
ID number:	0502211196
Academic degree:	Magister en Lingüística Aplicada a la Enseñanza Bilingüe Español – Inglés
Senescyt registration number	1027-15-86062861
Current job:	Universidad Técnica de Cotopaxi
Phone number:	0999844686
e-mail:	patricia.chacon@utc.edu.ec

Evaluation

Place an X on the square.

Criteria	Excellent	Good	Terrible
a) The teaching material constitutes a valid, current and relevant contribution related to the field.	X		
b) The teaching material is the result of an advanced research process; its content is the product of a complete conceptual development and critical contrast with other related researches.	X		

c) It is properly structured and argued (statement of the problem, methodology and results) in relation to the topic.	X		
d) The originality of the contributions and reflections of the author give added value to the proposal.	X		
e) The references are relevant and up-to-date.	X		
f) The research topic is appropriate.	X		
g) The proposal has the expected qualities for the level of training (appropriate language, vocabulary, spelling, etc.)	X		
h) The graphic illustrations that are in the dissertation (all kind of images and tables) are relevant, they clarify and provide significance.	X		
i) The dissertation embraces a clear and precise introduction on the objectives and issues along dissertation.	X		
j) The length of the dissertation is appropriate depending on the complexity of the topic, the objectives and the readers.	X		
k) The dissertation provides contributions regarding methodological proposals, approach, and conceptualization.	X		
l) The objectives in the introduction are met, so that there is harmony between objectives and results.	X		

Comment on the following statements, please.

1. TEMPORALITY: Is the proposal the result of an advanced research process, which means that it shows a methodological structure (problem, methodology and application)?

The proposal shows a good methodological structure, easy to use and with useful activities to develop an appropriate academic writing.

2. CONTENT: The content of the proposal is structured and written in an appropriate way to be understood and discussed by the educational community, and researchers?

The content of the booklet is understandable and relevant since it uses terminology according to the proposal.

3. SELECTIVITY: Can this proposal be considered a valid and significant contribution related to the field?

It is clear that the proposal contributes significantly to the vocabulary acquisition through appropriate learning strategies that allow learners to improve their academic writing.

4. Impact. What is the impact of this research? (Place an X on the square)

Local	X
Regional	
Nacional	
Internacional	

5. General comments and recommendations for the Author.

It is necessary to take into account the importance of vocabulary in academic writing especially when the objective of choosing descriptive words will help readers envision what researchers are describing. Therefore, it is recommended to disseminate and implement this useful booklet to English language learning programs and thus make the acquisition of new and less-frequent vocabulary effective while performing activities of academic writing.



Mg. Patricia Marcela Chacón Porras
DOCENTE CENTRO DE IDIOMAS
 Evaluator's signature

TECHNICAL UNIVERSITY OF COTOPAXI
GRADUATE DEPARTMENT
Master's degree in Applied Linguistics to
Teaching English as a Foreign Language
PROPOSAL VALIDATION

1. Research proposal data:

Author: Romero Mayorga Nataly Monserrat

Topic: Booklet of strategies for academic writing improvement based on lexical and readability analysis of written corpus.

Objective: To design a booklet with strategies for pre-service English teachers based on lexical and readability analysis of written corpus.

2. Evaluator's information

Evaluator's name:	MSC. Miryan Salazar Tobar
ID number:	1802840833
Academic degree:	Master's degree in teaching English as a foreign language
Senescyt registration number	1004-09-695471
Current job:	English teacher at Universidad Técnica de Ambato
Phone number:	0998321151
e-mail:	mc.salazar@uta.edu.ec

Evaluation

Place an X on the square.

Criteria	Excellent	Good	Terrible
a) The teaching material constitutes a valid, current and relevant contribution related to the field.	X		

b) The teaching material is the result of an advanced research process, its content is the product of a complete conceptual development and critical contrast with other related researches.	X		
c) It is properly structured and argued (statement of the problem, methodology and results) in relation to the topic.	X		
d) The originality of the contributions and reflections of the author give added value to the proposal.	X		
e) The references are relevant and up-to-date.	X		
f) The research topic is appropriate.	X		
g) The proposal has the expected qualities for the level of training (appropriate language, vocabulary, spelling, etc.)		X	
h) The graphic illustrations that are in the dissertation (all kind of images and tables) are relevant, they clarify and provide significance.	X		
i) The dissertation embraces a clear and precise introduction on the objectives and issues along dissertation.	X		
j) The length of the dissertation is appropriate depending on the complexity of the topic, the objectives and the readers.	X		
k) The dissertation provides contributions regarding methodological proposals, approach, and conceptualization.	X		
l) The objectives in the introduction are met, so that there is harmony between objectives and results.	X		

Comment on the following statements, please.

1. TEMPORALITY: Is the proposal the result of an advanced research process, which means that it shows a methodological structure (problem, methodology and application)?

<p>The proposal is developed according to a methodological structure so it's easy to follow and apply as it is well. In addition, the strategies presented in this booklet supports learners in developing academic writing</p>
<p>2. CONTENT: The content of the proposal is structured and written in an appropriate way to be understood and discussed by the educational community, and researchers?</p>
<p>The content of the proposal is structured and written in a very appropriate way so it can be applied for teachers as a useful guide to improve their teaching practice and in the development of their academic writing skills. Furthermore, teachers will be trained to develop different strategies that help learners to write better based on lexical and readability analysis of written corpus.</p>
<p>3. SELECTIVITY: Can this proposal be considered a valid and significant contribution related to the field?</p>
<p>Yes, this proposal has great benefits for educators as it great feasible tool to implement as strategic approach to improve performance in their academic writing skills.</p>

4. Impact. What is the impact of this research? (Place an X on the square)

Local	X
Regional	
Nacional	
Internacional	

5. General comments and recommendations for the Author.

All the strategies are designed to support teacher in improvement their students' writing academic skills.

A few grammar, spelling and capitalization mistakes need to be checked in the content of the proposal.

Evaluator's signature



Firmado digitalmente por:
MIRYAN CONSUELO
SALAZAR TOBAR

TECHNICAL UNIVERSITY OF COTOPAXI
GRADUATE DEPARTMENT
Master's degree in Applied Linguistics to
Teaching English as a Foreign Language
PROPOSAL VALIDATION

1. Research proposal data:

Author: Romero Mayorga Nataly Monserrat

Topic: Booklet of strategies for academic writing improvement based on lexical and readability analysis of written corpus.

Objective: To design a booklet with strategies for pre-service English teachers based on lexical and readability analysis of written corpus.

2. Evaluator's information

Evaluator's name:	Nelly Patricia Galora Moya
ID number:	1803104601
Academic degree:	Master's degree in teaching English as a Foreign Language
Senescyt registration number	1010-2018-1990762
Current job:	English teacher at Universidad Técnica de Ambato
Phone number:	0996172938
e-mail:	np.galora@uta.edu.ec

Evaluation

Place an X on the square.

Criteria	Excellent	Good	Terrible
a) The teaching material constitutes a valid, current and relevant contribution related to the field.	X		

b) The teaching material is the result of an advanced research process, its content is the product of a complete conceptual development and critical contrast with other related researches.	X		
c) It is properly structured and argued (statement of the problem, methodology and results) in relation to the topic.	X		
d) The originality of the contributions and reflections of the author give added value to the proposal.	X		
e) The references are relevant and up-to-date.	X		
f) The research topic is appropriate.	X		
g) The proposal has the expected qualities for the level of training (appropriate language, vocabulary, spelling, etc.)	X		
h) The graphic illustrations that are in the dissertation (all kind of images and tables) are relevant, they clarify and provide significance.	X		
i) The dissertation embraces a clear and precise introduction on the objectives and issues along dissertation.	X		
j) The length of the dissertation is appropriate depending on the complexity of the topic, the objectives and the readers.	X		
k) The dissertation provides contributions regarding methodological proposals, approach, and conceptualization.	X		
l) The objectives in the introduction are met, so that there is harmony between objectives and results.	X		

Comment on the following statements, please.

1. TEMPORALITY: Is the proposal the result of an advanced research process, which means that it shows a methodological structure (problem, methodology and application)?

Yes, this proposal shows a well-developed methodology which will guarantee its application

2. CONTENT: The content of the proposal is structured and written in an appropriate way to be understood and discussed by the educational community, and researchers?

The content of this proposal is easy to understand because it explains what is planned to do to solve the problem, why and how the researcher is going to design it.

3. SELECTIVITY: Can this proposal be considered a valid and significant contribution related to the field?

Yes, this proposal will be a contribution for English language learning and well as for English language teaching.

4. Impact. What is the impact of this research? (Place an X on the square)

Local	X
Regional	
Nacional	
Internacional	

5. General comments and recommendations for the Author.

I recommend you to check some spelling , punctuation and Capitalization suggested in the proposal.



Formado electrónicamente por:
**NELLY
PATRICIA
GALORA MOYA**

Evaluator's signature

ID 1803104601

Appendix 13

Users Evaluation

TECHNICAL UNIVERSITY OF COTOPAXI

GRADUATE DEPARTMENT

Master's degree in Applied Linguistics to

Teaching English as a Foreign Language

PROPOSAL VALIDATION

3. Research proposal data:

Author: Romero Mayorga Nataly Monserrat

Topic: Booklet of strategies for academic writing improvement based on lexical and readability analysis of written corpus.

Objective: To design a booklet with strategies for pre-service English teachers based on lexical and readability analysis of written corpus.

4. Evaluator's information

Evaluator's name:	
ID number:	
Academic degree:	
Senescyt registration number	
Current job:	
Phone number:	
e-mail:	

Evaluation

Place

an X on the square.

Criteria	Excellent	Good	Terrible
a) The teaching material constitutes a valid, current and relevant contribution related to the field.			
b) The teaching material is the result of an advanced research process, its content is the product of a complete conceptual development and critical contrast with other related researches.			
c) It is properly structured and argued (statement of the problem, methodology and results) in relation to the topic.			
d) The originality of the contributions and reflections of the author give added value to the proposal.			
e) The references are relevant and up-to-date.			
f) The research topic is appropriate.			
g) The proposal has the expected qualities for the level of training (appropriate language, vocabulary, spelling, etc.)			
h) The graphic illustrations that are in the dissertation (all kind of images and tables) are relevant, they clarify and provide significance.			
i) The dissertation embraces a clear and precise introduction on the objectives and issues along dissertation.			
j) The length of the dissertation is appropriate depending on the complexity of the topic, the objectives and the readers.			
k) The dissertation provides contributions regarding methodological proposals, approach, and conceptualization.			
l) The objectives in the introduction are met, so that there is harmony between objectives and results.			

Comment on the following statements, please.

6. TEMPORALITY: Is the proposal the result of an advanced research process, which means that it shows a methodological structure (problem, methodology and application)?
7. CONTENT: The content of the proposal is structured and written in an appropriate way to be understood and discussed by the educational community, and researchers?
8. SELECTIVITY: Can this proposal be considered a valid and significant contribution related to the field?

9. Impact. What is the impact of this research? (Place an X on the square)

Local	<input type="checkbox"/>
Regional	<input type="checkbox"/>
Nacional	<input type="checkbox"/>
Internacional	<input type="checkbox"/>

10. General comments and recommendations for the Author.

Evaluator's signature

I.D.....

Appendix 14

Results from users evaluation

Criteria	Excellent	Acceptable	Deficient	Total
a) The teaching material constitutes a valid, current and relevant contribution related to the field.	26	4	0	30
b) The teaching material is the result of an advanced research process; its content is the product of a complete conceptual development and critical contrast with other related researches.	27	3	0	30
c) The originality of the contributions and reflections of the author give added value to the proposal.	22	8	0	30
d) The proposal has the expected qualities for the level of training (appropriate language, vocabulary, spelling, etc.)	29	1	0	30
e) The graphic illustrations that are in the dissertation (all kind of images and tables) are relevant, they clarify and provide significance.	27	3	0	30
f) The objectives in the proposal are met, so that there is harmony between objectives and results.	30	0	0	30
g) The proposal is substantial with well-structured reflections and ideas.	27	3	0	30

Elaborated by: Nataly Romero

Source: Users' Evaluations

Appendix 15

Results from the activities developed in the workshop

	Strategy	Technique	Questions	# Correct	# Incorrect	Total Answers	Percentage
1	Guess meaning from context first set	Synonyms	The word "disorientation" in paragraph 1 is closest in meaning to:	21	9	30	70%
2	Guess meaning from context first set	General knowledge	The word "poppies" in paragraph 1 is closest in meaning to:	20	8	28	71%
3	Guess meaning from context first set	Parts of speech	The word "synthetic" in paragraph 2 is closest in meaning to:	23	7	30	77%
4	Guess meaning from context second set	Definition	Try to guess the meaning of the highlighted word. Please, write what you think it means. c. Once people attune themselves to new concepts, they realize changes are not as difficult as they think.	15	3	18	83%

5	Guess meaning from context second set	Antonym and contrast	Try to guess the meaning of the highlighted word. Please, write what you think it means. d. Many authors consider that vocabulary is the linchpin of second language acquisition.	20	0	20	100%
6	Guess meaning from context second set	Parts of speech	Try to guess the meaning of the highlighted word. Please, write what you think it means. c. Once people attune themselves to new concepts, they realize changes are not as difficult as they think.	18	4	22	82%
7	Use of dictionaries		The evidence produced an <i>argument</i> between those who followed Chomsky's ideas and those who supported Skinner's believes.	18	5	23	78%
8	Use of dictionaries		The Ecuadorian ministry of education <i>establishes</i> academic standards through the national curriculum guidelines.	18	8	26	69%
9	Use of dictionaries		Academic papers pass through a strict <i>review</i> process before being published in a journal.	24	5	29	83%
10	Use of dictionaries		The budget reduction for scholarships in Ecuador stopped many teachers from achieving a master's <i>degree</i> .	27	0	27	100%

Elaborated by: Nataly Romero

Source: Answers to the activities developed in the booklet's workshop retrieved from Quizizz: Free quizzes for every student", quizizz.com.

Note: The time was limited so, some students could not finish some tasks. The table only shows the answers registered on time.

Appendix 16











Urkund Report



Document Information

Analyzed document	RESEARCH PROJECT 1 (1).docx (D105625790)
Submitted	5/20/2021 12:06:00 AM
Submitted by	Rodrigo Tovar
Submitter email	rodrigo.tovar@utc.edu.ec
Similarity	7%
Analysis address	rodrigo.tovar.utc@analysis.orkund.com

Sources included in the report

SA	Proposal - Romero MAyorca - OMrev.pdf Document Proposal - Romero MAyorca - OMrev.pdf (D82548602)		17
W	URL: https://idoc.pub/documents/learning-vocabulary-in-another-language-vlr03djk9xlz Fetched: 12/6/2020 3:45:16 PM		12
W	URL: http://grad.litu.tu.ac.th/assets/public/kcfinder/upload_grad_web/public/MA-ELT_The ... Fetched: 1/8/2021 6:01:17 PM		1
W	URL: https://epdf.pub/vocabulary-studies-in-first-and-second-language-acquisition-the-i ... Fetched: 9/24/2019 2:39:17 PM		1
W	URL: https://stars.library.ucf.edu/cgi/viewcontent.cgi?article=3633&context=etd Fetched: 5/12/2020 11:58:07 AM		1
W	URL: https://uis.brage.unit.no/uis-xmlui/bitstream/handle/11250/2603032/Onyszko_Martyna ... Fetched: 10/21/2019 9:03:13 AM		1
W	URL: https://www.researchgate.net/publication/262084862_The_Relationship_of_Lexical_Ric ... Fetched: 12/12/2019 11:05:25 PM		1
SA	Zoachina Chapter 2 PhD.docx Document Zoachina Chapter 2 PhD.docx (D54682336)		1
W	URL: https://revistaseug.ugr.es/index.php/portalin/article/download/13677/11836 Fetched: 6/2/2020 4:20:27 AM		1
W	URL: http://www.rte.espol.edu.ec/index.php/tecnologica/article/download/628/377/1864 Fetched: 11/27/2020 8:24:03 AM		2