

**PONTIFICAL CATHOLIC UNIVERSITY OF ECUADOR IN
ESMERALDAS**



APPLIED LINGUISTICS SCHOOL

THESIS REPORT:

ANALYSIS OF A GENDER-BASED CORPUS OF 6TH LEVEL
ENGLISH LANGUAGE LEARNERS AT PUCE ESMERALDAS,
2017

TESIS DE GRADO:

ANÁLISIS DE UN CORPUS DE APRENDICES DE INGLÉS CON
ENFOQUE DE GÉNERO EN LOS ESTUDIANTES DE 6TO NIVEL
DE INGLÉS GENERAL EN LA PUCE ESMERALDAS, 2017

PRIOR TO THE ACADEMIC DEGREE IN:

APPLIED LINGUISTICS IN ENGLISH TEACHING

AUTHOR:

María Gabriela Ballesteros Chica

ADVISOR:

MSc. Javier Fernández Cruz

Esmeraldas, January 2018

Dissertation committee

As members of the Dissertation Committee at PUCE Esmeraldas, we certify that we have read the dissertation prepared by MARÍA GABRIELA BALLESTEROS CHICA entitled **Analysis of a gender-based corpus of 6th level English language learners at PUCE Esmeraldas, 2017** and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Bachelor of Applied Linguistics in E.L.T.

Reader 1

Reader 2

School Director

Dissertation supervisor

Final approval and acceptance of this dissertation are contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College. I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation supervisor

Statement by the author

I, María Gabriela Ballesteros Chica, affirm that the investigation in the present thesis report is totally unique, authentic, and personal. The content of this research is legally and academically a responsibility of the author and PUCE Esmeraldas.

María Gabriela Ballesteros Chica

C.I. 080290985-3

AGNOWLEDGMENTS

Foremost, I would like to thank God who gave me every opportunity with the pass of days, specially, health and strength to not give up.

I also want to thanks to my dad and my grandma Aida Saltos to supported me in my educational process, mainly spent a big part of their salaries in my university studies. Furthermore, I want to thank them for being unconditional in spite of the difficulties.

To my uncles, Liliana and Mario Ballesteros for giving to their sincerely support during all my life and to help me in all that I needed.

Lastly, to my dearest teachers PhD. Haydee Ramirez and MSc. Javier Fernandez who have been excellent tutors, because they guided me to complete the investigation process with patience and wisdom.

TABLE OF CONTENTS

Cover.....	i
Dissertation committee	ii
Statement by the author	iii
AGNOWLEDGMENTS	iv
List of Figures	vii
ABSTRACT.....	viii
RESUMEN	ix
1. INTRODUCTION	10
JUSTIFICATION	11
1.1 OBJECTIVES	12
1.1.1 General Objective	12
1.1.2 Specific Objectives	12
2. THEORETICAL FRAMEWORK.....	12
2.1. What is a Corpus?.....	12
2.1.1. Corpus Linguistics	13
2.1.2. History of Corpus Based Approach	15
2.2. Variation in Grammar Mistakes According to Gender	16
2.3. Writing Errors in Teenagers	17
2.4. Previous Studies.....	18
2.5. Legal Foundation	20
3. METHODOLOGY	20
3.1. Type of Investigation	20
3.2 Sample and Population	21
3.3. Methods	21
3.4. Techniques and Instruments	22

3.4.1 Test.....	22
2.4.2 Corpus.....	22
3.4.3 Concordancer.....	22
3.4.4. Data Processing.....	23
4. RESULTS.....	24
4.1. Verb Errors.....	24
4.2. Article / Determiner Errors.....	26
4.3. Noun Use.....	27
4.4. Pronoun Errors.....	28
4.5. Wrong Collocations, Idioms and Prepositions.....	28
4.6. Punctuation Errors.....	29
4.7. Syntax Errors.....	30
4.8. Transitions, Mechanics and Redundancy.....	31
5. DISCUSSION.....	34
6. CONCLUSIONS AND RECOMMENDATIONS.....	35
6.1. CONCLUSIONS.....	35
6.2. RECOMMENDATIONS.....	36
7. REFERENCES.....	37
8. APPENDICES.....	42
Appendix 1.....	42
Appendix 2.....	43
Appendix 3.....	44
Appendix 5.....	46
Appendix 6.....	46
Appendix 7.....	47
Appendix 8.....	48

List of Figures

Table 1 Sample details	21
Table 2 Tags for error annotation	23
Figure 1: Verbs	24
Figure 2: Subject Verb Agreement	25
Figure 3: Article or Determiner	26
Figure 4: Nouns	27
Figure 5: Pronouns	28
Figure 6: Runnons- Comma Splice.....	29
Figure 7: Word Order	30
Figure 8: Transitions	31
Figure 9: Mechanincs.....	31
Figure 10: Local Redundancy.....	32

ABSTRACT

The study of English as a foreign language has been presenting difficulties in the different educational levels, taking into account the high impact of the internal and external factors that influence the teaching-learning process to people who want to improve Linguistics skills. This research focuses on the study of gender as a consequence of a high or low level in university performance with respect to ESOL at the Catholic University of Ecuador in Esmeraldas. After the compilation of a learner corpus from sample texts which were compiled through Google forms, linguistic analysis was carried out by means of the application of quantitative and qualitative methods in order to obtain a systematic summary of errors in the learner's corpus while taking several parameters into account, such as age, years of training and, especially, gender. Statistical data were collected with AntConc and, by means of a comparison of the results obtained, it revealed that errors were more prevalent in male students than on female at the time of composing general domain texts in English. In addition, we present a summary of the most frequent errors that University students commit while using English as a foreign language.

Keywords: learner's corpus, gender, English as a Second Language, Ecuador.

RESUMEN

El estudio del inglés como lengua extranjera ha venido presentando dificultades en los distintos niveles educativos, teniendo en cuenta el alto impacto de los factores internos y externos que influyen en el proceso de enseñanza-aprendizaje a quienes buscan mejorar sus habilidades Lingüísticas. Esta investigación se centra en estudiar el género como consecuencia de un alto o bajo nivel en el rendimiento universitario con respecto a la materia de inglés en la PUCE Esmeraldas, Ecuador. Por medio de la recopilación de textos escritos luego de aplicar la encuesta en la página virtual de google drive, se realizaron análisis lingüísticos mediante la aplicación del método cuantitativo para conocer el margen de error en el corpus de los estudiantes matriculados en la materia de inglés general, teniendo en cuenta las diferentes carreras en las que se encontraba cada uno. Mediante una comparación de los resultados obtenidos a través de software AntConc, se dio a conocer el que el género masculino fue el que cometió más errores a la hora de redactar un texto en Inglés, también se pudo visualizar los tipos de errores más comunes que los estudiantes universitarios cometen hoy en día en la escritura del Inglés como lengua extranjera.

Palabras clave: corpus de aprendices, género, inglés como segunda lengua, Ecuador.

1. INTRODUCTION

What are the most frequent problems that students in 6th level of general English from PUCE Esmeraldas face when they are learning English? Many factors are influential in the process of learning English as a Second Other Language (hereafter, ESOL) such as economic factors, environment, motivation, age and gender. If we want to deepen into the most frequent problems in the English learning process, first it is necessary to analyze the factors by which development can be interfered. In this research, gender is going to be focus on future researches.

Through the ages, Linguistics have developed different approaches to study different systems of language. Corpus based approaches allows to perform an empirical-based study on which a large compilation of electronic texts containing real samples permits us to study language as used by its speakers. Before the widespread use of computers, analyses of this nature were done by hand. With the advent of computational linguistics, we can obtain accurate empirical results through automatic processes.

Kennedy (1998) stated that there are meant types of corpus design and corpus-based approaches which can be used for various kinds of analysis. In this case, this research is focused on the grammatical analysis of texts written by students in 6th level of general English at PUCE Esmeraldas while considering the gender differences of the students. A learner's corpus can be defined as an electronic collection of language data produced by foreign-language learners. This resource is of great relevance for teaching and learning English as a foreign-language (Granger, 2012). The problem of investigation is: Which are the most writing common errors among male and female gender in Students from 6th level of General English at PUCE Esmeraldas?

To sum up, the problem leads to solve the following questions:

- What are the theories that support linguistics analysis of language corpus?
- What are the principal errors in writing according to the age and gender in students from 6th level of General English at PUCE Esmeraldas?

- What proposals can we make in order to improve these problems?

All in all, this research has six sections. Section 1, the theoretical framework and literature review can be found, previous studies and legal foundation. In section 2 is written about methodology focused on techniques and instruments that were used, universe, sample, population and the compilation of information. Section 3 is structured by the analysis and interpretation of the results. Section 4 refers to the discussion of the investigation process. Finally, Section 5 and 6 show the conclusions and recommendations after having analyzed the results.

JUSTIFICATION

Linguistics is an important science which studies the language processes in all its aspects, changing and improving through time. Nowadays, at a time where people live in a global world, there are diverse ways to study linguistics and one of them is through 21st century skills making use of the language corpus that permit to know the principle of total accountability. Also, the inspection and structure of a word; so real language provides lexical, morphosyntactic, semantic and pragmatic information. Moreover, corpus shows an effective way to analyze the students' errors in the writing skills by means of a whole step, using this research approach in our computers, taking into account the branches of linguistics mentioned before. This resource will make an analysis about common writing mistakes in the students form 6th level of General English at PUCE Esmeraldas, through language corpus as a relevant tool that teachers will be able to use.

The importance of this research is based on the fact that it is the first time that a study of this type has been carried out in Esmeraldas city. Likewise, it shows corpora as a relevance tool to analyze the real students' writing by which teachers can detect difficulties that students present about grammar and writing. Finally, it helps to find common errors according to the writing error category.

1.1 OBJECTIVES

1.1.1 General Objective

- To make a diagnosis of the English writing errors in students from 6th level of General English at PUCE Esmeraldas through the use of computational tools.

1.1.2 Specific Objectives

- To explore the required theoretical background and a set of corpus-based techniques that permit us to analyze the students' production.
- To find common errors among male and female gender in Students from 6th level of General English at PUCE Esmeraldas.
- To classify these errors during the language learning process according to gender through computational techniques.
- To build a specific learner corpus from the students' production while learning English.

2. THEORETICAL FRAMEWORK

2.1. What is a Corpus?

Learner corpora, or searchable collections of the written or spoken production of language learners, have become increasingly vital in the field of applied linguistics. Through the years, a corpus can be created in an electronic way with the capacity of collect multimillion words. Sinclair (2004) said that it is defined as a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research. Likewise, Hunston (2002) states that Linguistics has used the word corpus to describe a collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written texts or tape recordings, which have been collected for linguistic studies.

After that, making reference with the electronic corpus, this research is characterized as a design of an annotated corpus or corpus mark-up which provides great advantages in terms of linguistic research studies, because the information added to the text permits to do an analysis of the grammar class of words (Lopez 1998). Besides, tagging can be automatic (the computer program tags the corpus) or manual (there is human intervention in the tagging with the aid of a computer software). Hunston (2002: 18) points out the level of accuracy achieved depending on the automatic or manual tagging, and states that even though in the first case a corpus can be totally tagged, the level of accuracy is lower than that obtained through manual tagging, however this is only feasible when the corpus is small.

To talk about Sinclair is to talk about an important character of corpus, because Professor Sinclair was the author of the creation of the first electronic corpus, and was instrumental in developing the tools needed to analyze the data. Having corpus data allowed Professor Sinclair and his team to find out how people really use the English language and to develop new ways of structuring dictionary entries. The corpus also can study phrases through collocation. As Bennett (2010) argues that collocation is the statistical tendency of words to co-occur. This means that when one word is used, there is a high statistical probability that a certain word or words will occur alongside of it. For example, the noun form of the word deal. The words, 'big', 'good', and 'great' are collocations of 'deal' as a noun, meaning that when people use 'deal' as a noun, they often refers to 'a big deal', 'a good deal', and/or 'great deal'.

2.1.1. Corpus Linguistics

The soonest and most noteworthy effect that corpus phonetics had on dialect instructing can be found in lexicography. The compilation of the Collins Birmingham University International Language Database (COBUILD) in the 1980s, initiated and organized by John Sinclair, led to the first corpus-based dictionary and the first edition of the Collins COBUILD English Language Dictionary (1987), which depended on a corpus of 20 million words. In the 1990s, COBUILD was extended to frame the Bank of English, a dynamic corpus to which new texts have been continually included with the goal that someday its size is around 500 million words. (Mukherjee, 2006)

Furthermore, COBUILD set new principles in lexicography on the grounds that the corpus-based depiction of English and its corpus-based codification has brought about another age of word references that incorporate data that had not been accessible in customary pre-corpus dictionaries. The following entry for the noun assumption, which is taken from the corpus-based Macmillan English Dictionary (2002), exemplifies lexicographical information that can only be derived from corpus data:

In the same way, O’Keeffe (2007) determined that corpora are storages of written and/or spoken texts in a computer with a detail design matrix addressed to a determine audience. That is, like Ganger, to pay attention to specific characteristics such as varieties of age, gender, location, school, level, teacher, class, nationalities, etc.

On the other hand, Xiao (2008) stated about a lot of benefits of corpus based approach. First, a corpus can be more comprehensive and reliable, so a corpus shows linguistic intuitions of a range of language speakers, which helps in intuitions of individual speakers, also it can provides us a high number of examples in real communication context. Second, a corpus can show us what is common and typical, it refers to learners of a foreign language can search about specific terms to know the high or low range by which people used any words. Moreover, it can find differences that intuitions alone cannot perceive, for example synonyms of the word “totally”, so could be absolutely, utterly, completely or entirely. Third, corpora provide accurate statistics because we can know about the exact quantity of collocates or repeated words to be compared in linguistics analysis, for example the percent of errors among men and women as in this research. After that, corpus data is more natural because it is used in real communications instead of being invented specifically for linguistics analysis. Besides, a constantly updated corpus can reflect even recent changes in the language, so it is important because we can learn about the new language tendencies to be applied in our speaking.

Additionally, some authors like McEnery, Xiao and Tono (2007) described different kinds of corpus in their book about corpus based language studies. They have classified it in five. First, *specialized corpora* are collection of texts of a type, such as newspaper editorials, geography textbooks, academic articles in a special subject, lectures, casual conversations or essays written by students. After that, it is used to investigate a specific type of language. Second, a general corpus which is considered as a frequent type of

corpus because contains texts just in one language. The corpus is usually tagged for parts of speech and is used by a wide range of users for various tasks from highly practical ones. Third, a *parallel corpus* refers to those that contain two monolingual corpora, being one corpus the translation of the other. For example: a novel and its translations. Fourth, a *learner's corpus* is characterized as a group of texts produced by language learners. This kind of corpus is used to study mistakes and problems learners have when they are learning a foreign language. Fifth, a *diachronic corpus* contains texts from different periods and it is used to study the development or change in language. In addition, there is a specialized diachronic feature referred as trends, which identifies words whose usage changes most of the selected period of time.

2.1.2. History of Corpus Based Approach

Sinclair was the progenitor of the original of present day corpus semantics through the advancement of the COBUILD venture by the University of Birmingham. This present undertaking's point was to construct corpus-driven vocabularies for outside students of English. Therefore, he turned into the central consultant of the Collins' COBUILD *English Language Dictionary*, which its first version was distributed in 1987. Certainly, corpora have been engaged with our lives since long time back. For example, lexicons and present day lexicographic materials are at present in light of corpus work. From the 1990s on, word reference producers utilize corpora information keeping in mind the end goal to make true and genuine cases of dialect (McEnery & Xiao, 2010).

Nowadays, corpus is seen as one of the most valuable linguistic tools as it passed from being an enterprising idea of a few research groups to become one of the most valuable tools for both, Linguistics and computer science (Granger, 2003). Besides, thanks to the advantages of technology, corpora collect enormous multimillion-word texts from many different sources to be published online and ready to be analyzed by linguists around the world. Written and spoken texts, such as a set of newspapers or radio recordings, are transcribed (if necessary) and labeled. Some years ago, people had to compile the texts writing them by hand, for example the first concordance of a Bible made by around 500 monks in the 13th century (Tribble & Jones, 1990, named in Kindt & Wright, 2001). This indeed demanded too much effort and time. But technology advances are so fast

today; it cannot only has texts in a corpus but also multimedia elements such as videos, scans and recordings in a way that they can appear in the computer screen in matter of seconds when someone queries them.

Moreover, O'Keeffe (2007) also points out the fact that corpora give the opportunity to do a quantitative and qualitative analysis. In terms of quantitative results, a corpus can show statistical data about the frequency of occurrence of a certain word and it can be compared with other sources.

2.2. Variation in Grammar Mistakes According to Gender

This research is focused on gender differences. Males and females clearly differ in various perceptual, motor and cognitive domains, both in humans and in other animals (Halpern, 2000; Kimura, 1999). The use of foreign language in the EFL classrooms can be shaped by many factors, among which gender has a leading role (Ali, 2016) Moreover, gender has been viewed as an important factor that assumes a particular part and impacts in the foreign language learning and there are a few differences in the way of producing language between men and ladies, and no instruction or social molding can entirely eradicate these distinctions.

In language, however, sex differences have not been so evident (Halpern, 2000). Such girls and boys tend to use different areas of their brain while learning some important parts of grammar. To support, authors as Zoghi, Kazemi, and Kalani (2013) in their studies about the effects of gender in EFL, state that while both boys and girls have improved their performances, girls achieved higher marks than boys in EFL learning.

In addition, this study, Zoghi and his colleagues (2013) studied how students were learning EFL by a group of 50 boys and 50 girls and age about 12-14, selected from four different classes. After that, from that point onward, the scores of the students in the English accomplishment test which managed toward the finish of second semester were used as instrument in their examination. The test paper included four sections: vocabulary, language structure, sentence capacity, and perusing appreciation. The items contained fill in the clear, coordinating, various decision, and correspondence which proposed to evaluate general capacity in students' production.

All items considered, shockingly, the consequences of the investigation showed that ladies were better in EFL. These discoveries however uncovered a noteworthy cooperation impact of gender on students' accomplishment test. To sum up, there is a noteworthy contrast amongst male and female performance, the magnitude of the difference and the strength of association between the total males and the total females is relatively small.

2.3. Writing Errors in Teenagers

Writing in a foreign language is not simple for learners. It is muddled to write in other language and sometimes most of the understudies try to decipher or translate words, expressions, and sentences from the first language to English getting awful outcomes. Many examinations demonstrate for the starting English Foreign Language (EFL) learners, there has a tendency to be obstruction from their first dialect during the time spent writing in English (Benson, 2002).

The difficulty of teaching is to discover strategies to initiate definitively the inactive learning that learners have as far as the written work aptitude, and to enable the students to end up plainly more capable while attempting to dispense with some of their normal blunders. At that point, it is basic here to show a difference between errors and mistakes. Both, Corder (1967, 1971) and James (1988) reveal a criterion that helps us to do so: a mistake can be self-corrected, but an error cannot. Errors are “systematic,” i.e. likely to happen regularly and not recognized by the learner. Hence, only the teacher or researcher would locate them, the learner would not (Gass & Selinker, 1994).

Moreover, as Richards & Renandya (2002:303) claim that writing is the most troublesome aptitude for EFL students to ace. The trouble lies in creating and arranging thoughts, as well as in making an interpretation of these ideas into understandable content.

2.4. Previous Studies

Corpus linguistics is one of the technology-based tools that could be very useful in teaching but still has not been widely used or tested. Nevertheless, in the last 30 years, the use of corpora in classrooms has started to develop (Varley, 2008). One of these researchers as Díez (2001, 2003) in the first studies based on the Madrid corpus analyzed the students' problems with coherence and cohesion. The use of the Madrid Corpus exposed the study of various measures of lexical complexity, fluency, syntactic complexity, information-structure and the use of connectors, conjuncts and conjunctions.

However, this investigation applied by seven Spanish Research Groups was done to improve the teaching materials offered to the students or to describe their interlanguage. The wealth of information provided by researchers using learner corpora by Spanish students of English at various levels confirms the vitality of the field in Spain. Nevertheless, further research is encouraged so that a better understanding of the acquisition process of the foreign language is gained, interlanguage patterns are described, difficult aspects of the foreign language at different proficiency levels are highlighted, etc. With the scientific data obtained from rigorous learner corpus-based studies, teaching materials (i.e. textbooks, dictionaries, etc.) which meet Spanish students' real needs can be designed and used in our classes (Díez, 2001, 2003).

Nowadays, corpus is considered as an important implement in English classroom, so other researchers as Saeed and Waly (2009) in their investigation about Corpus as tool for Material Design in classrooms obtains that corpora can be used in the English learning process to improve language skills providing a technological tool. The author obtains that if teachers are trained on how to design suitable corpus-based tasks. Teachers can help their students be exposed to a broader framework of how English is used for communication by native speakers. Certain points to bear in mind while designing corpus-based tasks include the ability level of the students, cultural and educational backgrounds and the age group of the students. It is also recommended to use online corpora as they are available for anybody at any given time. In class a specialized corpus is needed, EFL instructors and researchers need to make sure it is available at their institutions (Saeed & Waly, 2009).

In Esmeraldas, Castillo (2016) stated in her studies about corpora as a teaching learning material that the use of corpora is not the definitive solution for improving vocabulary level but, it is an effective optional strategy that teachers should use when they create materials for their classes. After that, the author in the results of the survey suggests that working with corpus can provide some benefits such as exposure to a more real language, challenges students, and increases interest. On the other hand, the study conducted by Darus et. al. (2009, p. 493) by concluding the result as the errors that participants committed were basically grammatical. Further, emphasis indicated that the problems of the participants had in acquiring normal grammatical rules. Grammar is also considered as one of the most important factor in language teaching and learning. This is also enlightened in the study conducted by Caroline Mei Lin Ho (2003).

Contributing with my research, Boroomand & Rostami Abusaeedi (2013) presented a study conducted on 100 Iranian advanced EFL learners' written errors (50 male learners and 50 female learners), presents different classifications and subdivisions of errors, and carries out an analysis on these errors. Detecting the most committed errors in each classification, findings reveal significant differences between error frequencies of the two male and female groups (more error frequency in female written productions).

Supporting Boroomand, & Rostami Abusaeedi, investigation, Pearson Chi-Square test, contrasted the frequencies of different classifications of errors in the two groups of male and female learners. Regarding the p-values for processing errors and syntactic errors in the two groups which were less than the significance level ($\alpha= 0.05$), it is concluded that there are significant differences between the two groups of males and females, considering processing errors and syntactic errors. Despite some researches done by different researchers which reveal female language learners commit less writing errors, show more improvement in writing over time or gain higher writing scores than male learners (Chen, 1996, Abu-Jarad, 2008, Chiu, 2008, Al-Nawas, 2009), the present study reveals more error commitment by female learners.

2.5. Legal Foundation

This work is written according to the Constitution of Ecuador and general Education under Title VII “Plan Nacional del Buen Vivir” (2012) in the Article 344 and Article 350.

Art. 344.- El sistema nacional de educación tendrá como finalidad el desarrollo de capacidades y potencialidades individuales y colectivas de la población, que posibiliten el aprendizaje, y la generación y utilización de conocimientos, técnicas, saberes, artes y cultura. El sistema tendrá como centro al sujeto que aprende, y funcionará de manera flexible y dinámica, incluyente, eficaz y eficiente. El sistema nacional de educación integrará una visión intercultural acorde con la diversidad geográfica, cultural y lingüística del país, y el respeto a los derechos de las comunidades, pueblos y nacionalidades.

[...]

Art.350.- El sistema de educación superior tiene como finalidad la formación académica y profesional con visión científica y humanista; la investigación científica y tecnológica; la innovación, promoción, desarrollo y difusión de los saberes y las culturas; la construcción de soluciones para los problemas del país, en relación con los objetivos del régimen de desarrollo.

Constitution of the Republic of Ecuador (2008)

3. METHODOLOGY

3.1. Type of Investigation

A corpus-based approach was followed as a learner’s corpus designed in order to identify common mistakes and vocabulary used in writing by 6th level of English students at PUCE Esmeraldas. When regarding the levels of measuring, it is quantitative research, as it uses both a survey and corpus-based analysis of features in order to describe the problems that students from PUCE Esmeraldas have when they are writing in the foreign language.

3.2 Sample and Population

PUCE Esmeraldas is the main beneficiary of this research. The population of this study was a class of 6th level General English at PUCE Esmeraldas (similar to level B1 ECFL) which include students from different programs, being the majority of them undergraduate students of Environmental Management. Our intentional sample included all of the 24 students that regularly attend the class. There were 11 male and 13 female students.

As results of the initial corpus, codenamed *Esmeraldas Learner's Corpus*, it includes 3172 tokens, originated from an in-class writing exercise. This corpus was subsequently error-annotated by hand, in order to be processed manually.

English learners' corpus	
Types:	660
Tokens:	3172
Male:	11
Female:	13
Ages:	21-25: 17 students 19-21: 7 students
Degree:	Administration: 1 International Business: 1 Environmental Management: 22

Table 1 Sample details

3.3. Methods

This research followed both inductive and deductive methodologies, because from the particularity of the students' errors, general rules can be identified. As a result, other researchers may use this general information in their lessons. The method of analysis and synthesis was used to analyze theories and the results to synthesize the information. To conduct this research, a five-step methodology was followed: (1) Compilation of writing output through an online formulary, (2) manual detection and annotation of

errors, (3) data extraction, XML labeling and final corpus compilation, (4) automatic calculation of error statistics, (5) hand processing in order to detect exceptions and other relevant features.

3.4. Techniques and Instruments

3.4.1 Test

An activity was designed in Google Forms in order to collect and compile output from students in the campus' language lab. The activity that was proposed to students contained one question that consisted on writing a paragraph about travelling, along with some extra metadata such as age, gender and degree. This way, information will be collected in order to create a learner corpus in which the errors and common vocabulary can be defined, to observe the differences between men and women. The survey can be seen in Appendix 3.

2.4.2 Corpus

To create an annotated learners' corpus through the extraction of the text in the cells from the test spreadsheet to individual UTF8 text files, and its metatags were labeled in XML code. As a result, we obtained a small corpus that included 3732 tokens and 660 types. The errors committed by the students in this corpus were manually annotated through an error tag code, similar to the provided by Dahlmeier, Ng and Wu (2013). A concise list of error tags is found in Table 2 and the learners' corpus can be found in the CD attached to this copy.

3.4.3 Concordancer

Concordances were found through AntConc, a freeware, multiplatform tool for carrying out corpus linguistics research and data-driven learning. It runs on any computer running Microsoft Windows, Macintosh, and Linux. It is developed in Perl using various compilers to generate executable for the different operating systems. AntConc does not require installation, as it is a stand-alone program (Laurence, 2014)

3.4.4. Data Processing

After the corpus is extracted from the Google Spreadsheet created ad hoc, it was converted into individual .txt files and compiled into a corpus, no further lemmatization not tokenization was performed. The resulting corpus was analyzed with the AntConc concordance in order to obtain statistic results. The resulting digital information was tabulated with Microsoft Excel.

ERROR TAG / ERROR CATEGORIES			
VERBS		PRONOUNS	
VT	VERB TENSE	Pform	PRONOUN FORM
VFORM	VERB FORM	Pref	PRONOUN REFERENCE
		Wcip	WRONG COLLOCATIONS, IDIOMS AND PREPOSITIONS
SUBJECT-VERB-AGREEMENT		SENTENCE STRUCTURE	
SVA	SUBJECT-VERB- AGREEMENT	Srun	RUNNONS, COMMA SPLICE
ARTICLES/DETERMINERS		WORD ORDER	
ArtOrDet	ARTICLE OR DETERMINER	Woinc	INCORRECT SENTENCE FORM
		Woadv	ADVERB/ADJECTIVE POSITION
NOUNS		TRANSITIONS	
Nn	NOUN NUMBER	Trans	LINK WORDS/PHRASES
MECHANICS		REDUNDANCY	
Mec	PUNCTUATION, CAPITALIZATION, SPELLING AND TYPOS	Rloc	LOCAL REDUNDANCY

Table 2 Tags for error annotation

4. RESULTS

4.1. Verb Errors

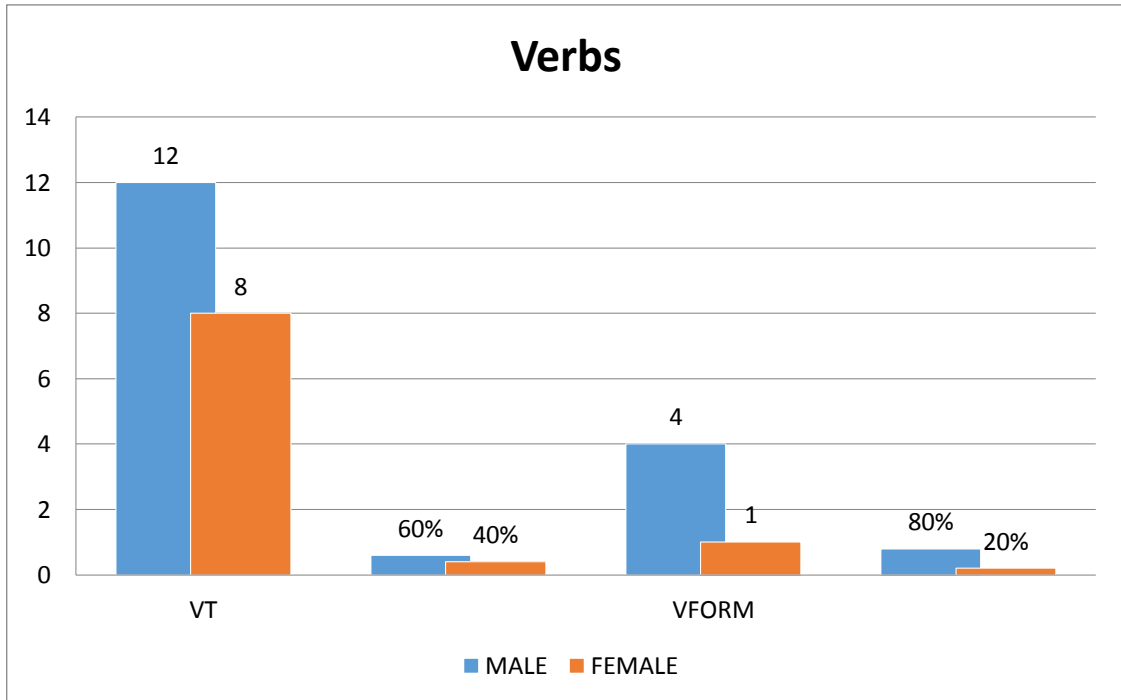


Figure 1: Verbs

In this section 11 of 24 students used incorrectly the verb selection and tense, so the 60% of errors equivalent to male gender who wrote some incorrect sentences. For example:

- Within (Wcip) Colombia I know (**vt**) (vform) Cali, Pasto, (trans) Ipiales.
- My best vacation was go to yasuni itt where we were 1 weekend and visit (**vt**) many places interesting (WOadv).
- To spend the day at the beach with friends or family make (**vt**) games and eat(**vt**).

The 80% of errors belong to male gender, being just 4 of 24 students who wrote sentences without respect the verb form. After that, the majority of the students did not commit any errors, but out of the 5 individuals that had writing incorrections, 4 were men. For example:

- Countries is (vform) impresionated.
- I like (vform) travel around the world
- I like (Vform) watch tv in general sports

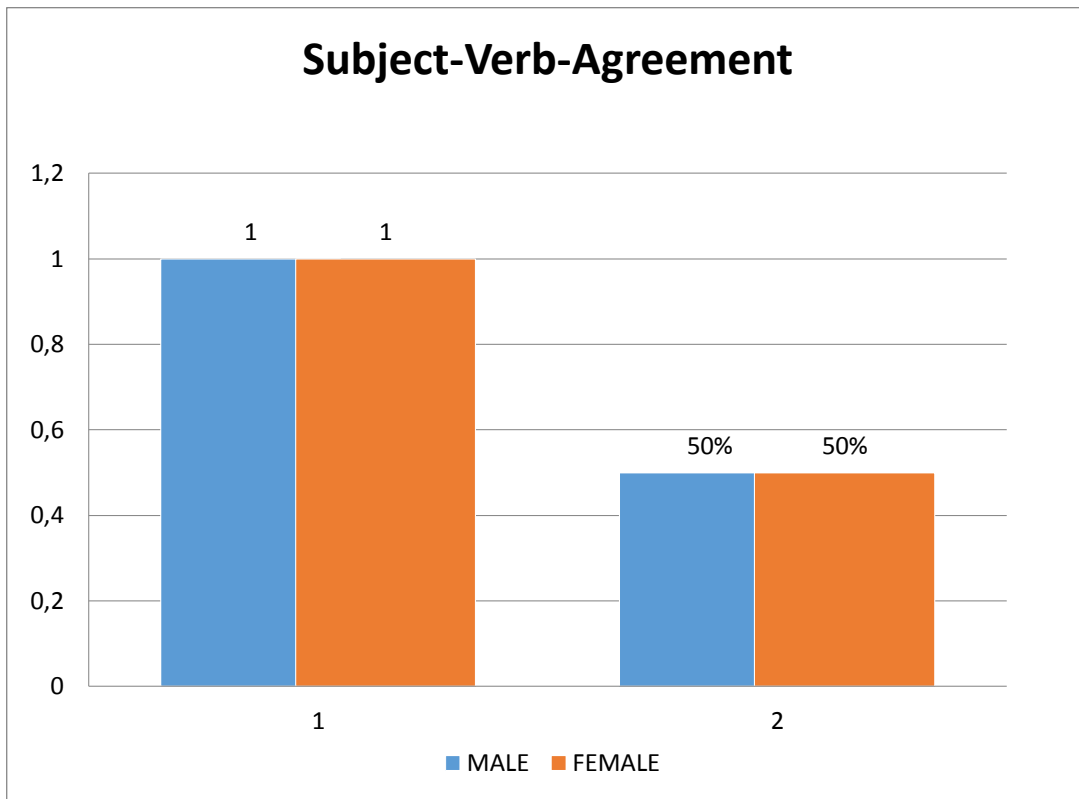


Figure 2: Subject Verb Agreement

In Figure 2, we can observe that only 2 out of 24 (1 M, 1 F) students used an incorrect subject-verb agreement, so very few learners have problems at the moment of using third person singular:

- For me Cuba have (sva) a (ArtOrDet) culture (WOadv) interesting.

4.2. Article / Determiner Errors

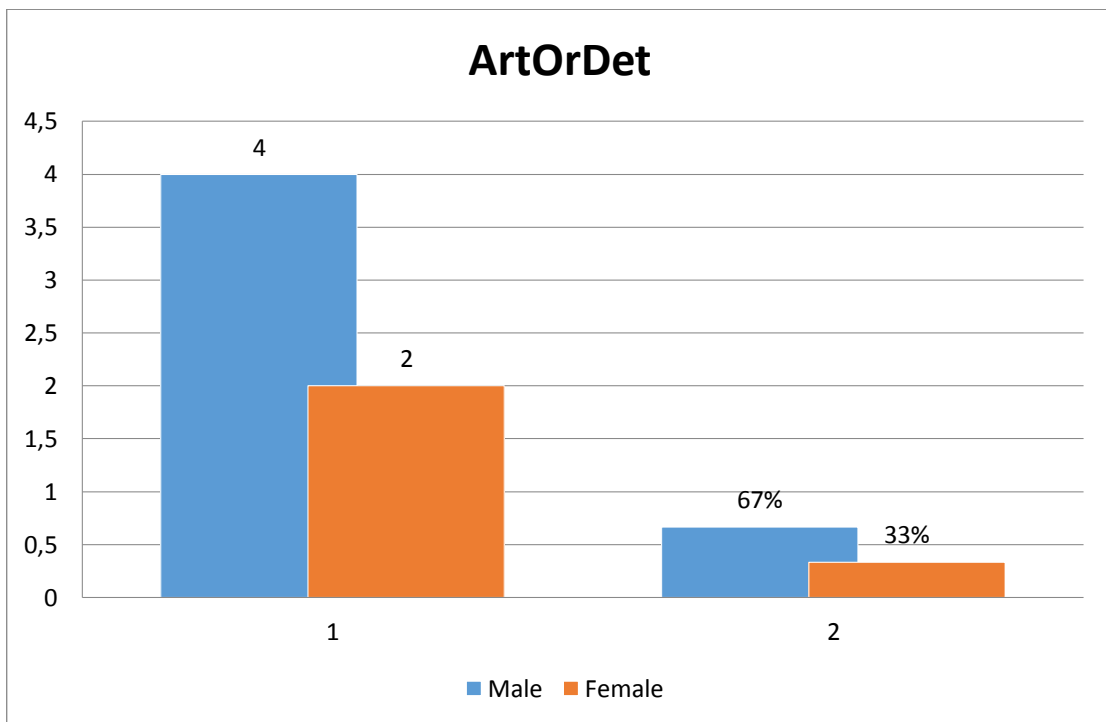


Figure 3: Article or Determiner

Some learners present problems at the moment of using articles and determiners. The 67% of the errors corresponded to males while the 33% corresponded to females. For all, this figure shows to 6 of 24 students had difficulties in this case. As result, most of the students are good using articles and determiners. In the following sentences you can observe a short sample of their errors.

- (mec) have (ArtOrDet) good experience.
- Sunday at noon we usually go to (ArtOrDet) church with my family

4.3. Noun Use

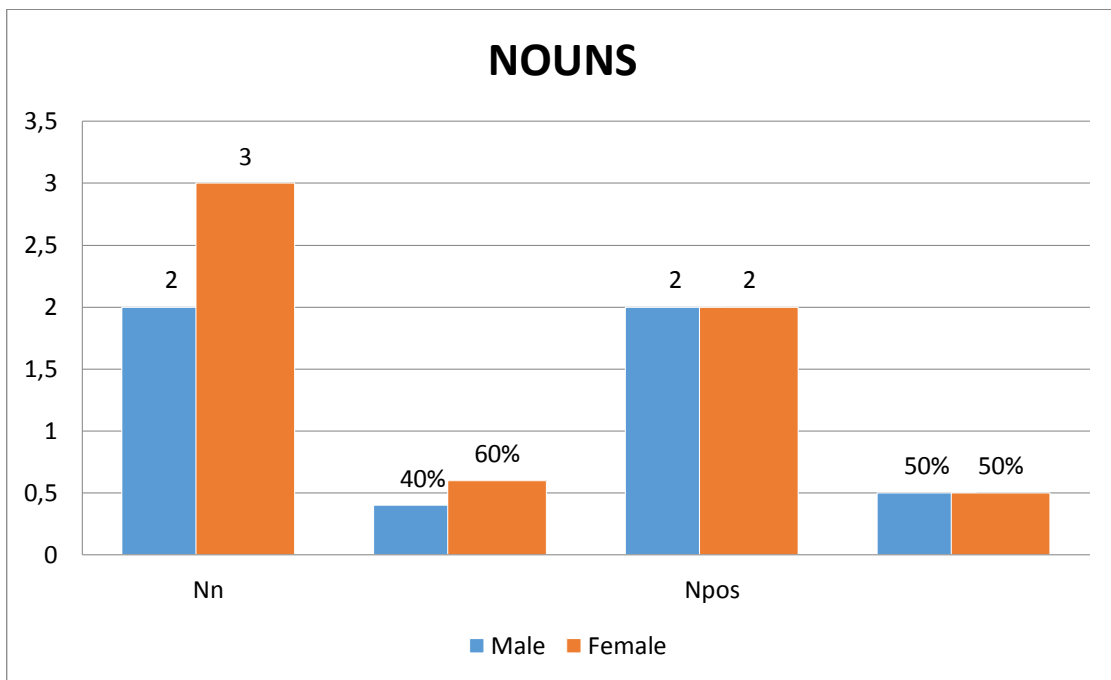


Figure 4: Nouns

Students in the English class demonstrated good management in the use of nouns in sentences. Out of the 5 students that committed errors in this category, only 40% of the errors were committed by men. For example:

- This (nn) countries is (vform) impresionated.
- Sports and health life are my other hobbie (nn).

Learners had to use noun possessive to write about their travels in the paragraph. They did not present a lot of problems in this part because the high percent of them use correctly the apostrophe to indicate possession, but a low percent had problems with this rule. In fact, 4 students (2M, 2F) have forgotten the use of apostrophes as in the following examples:

- I need a capacitation about English by Language student's. (Npos)
- I found interesting places such as Carolina (Npos) Park.

4.4. Pronoun Errors

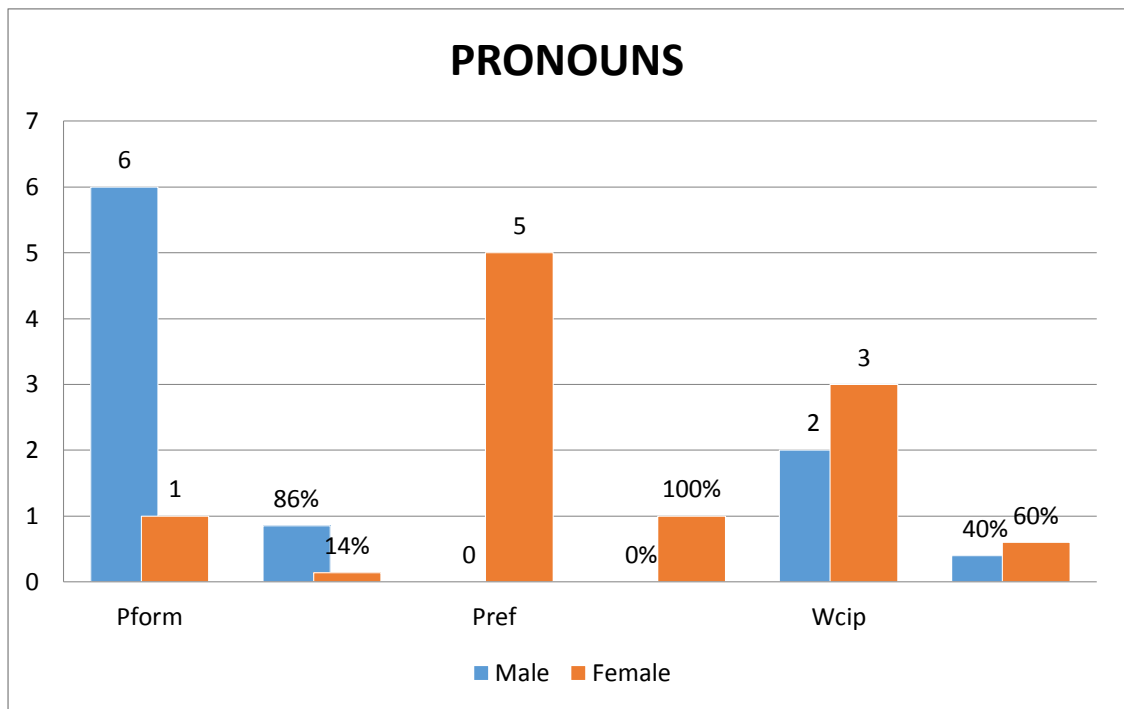


Figure 5: Pronouns

Figure five shows 86% of errors correspond to male gender who committed approximately 6 errors while the 14% of errors was by a woman who committed just one error. E.g. for my (Pform) the most exciting was Acapulco.

In this case male gender did not commit any error, but female gender present 5 errors. Besides, 2 of 24 students use incorrectly the pronoun reference so it shows that women are more probably to commit errors than men. E.g. I like to learn about the cities, their (pref) religion and their (pref) customs.

4.5. Wrong Collocations, Idioms and Prepositions

Corpus results reflect just 5 errors among all the 24 paragraphs. It shows that students have good management of prepositions, and just a small part of them were confused with this rule and 60% were women. For instance, women had more problems than men during the use of prepositions as in these examples:

- I went (Wcip) a diferent (mec) city.
- beach of the monks of (Wcip) Manabi.

4.6. Punctuation Errors

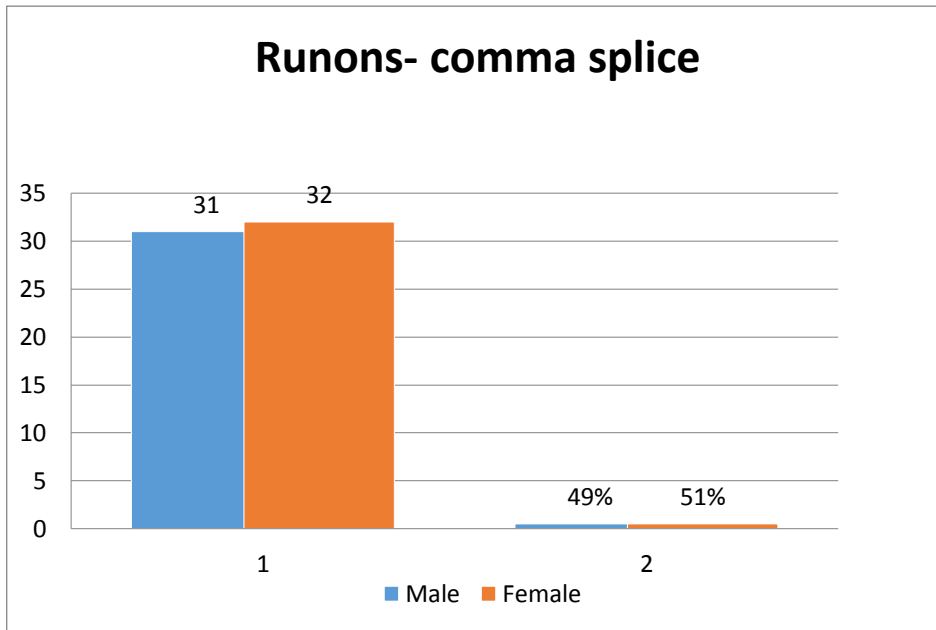


Figure 6: Runnons- Comma Splice

The correct use of punctuation in a sentence was difficult for all the students because a high number of them, exactly 21 of 24 students did not respect periods and commas. This is the main problem detected in this learners' corpus, as it reflected 63 errors of punctuation and commas in just 24 paragraphs. Moreover, 51% of errors were committed by women and the other 49% were committed by men. Concluding with a small difference of percentage, female adolescents are more probably to commit errors than male.

4.7. Syntax Errors

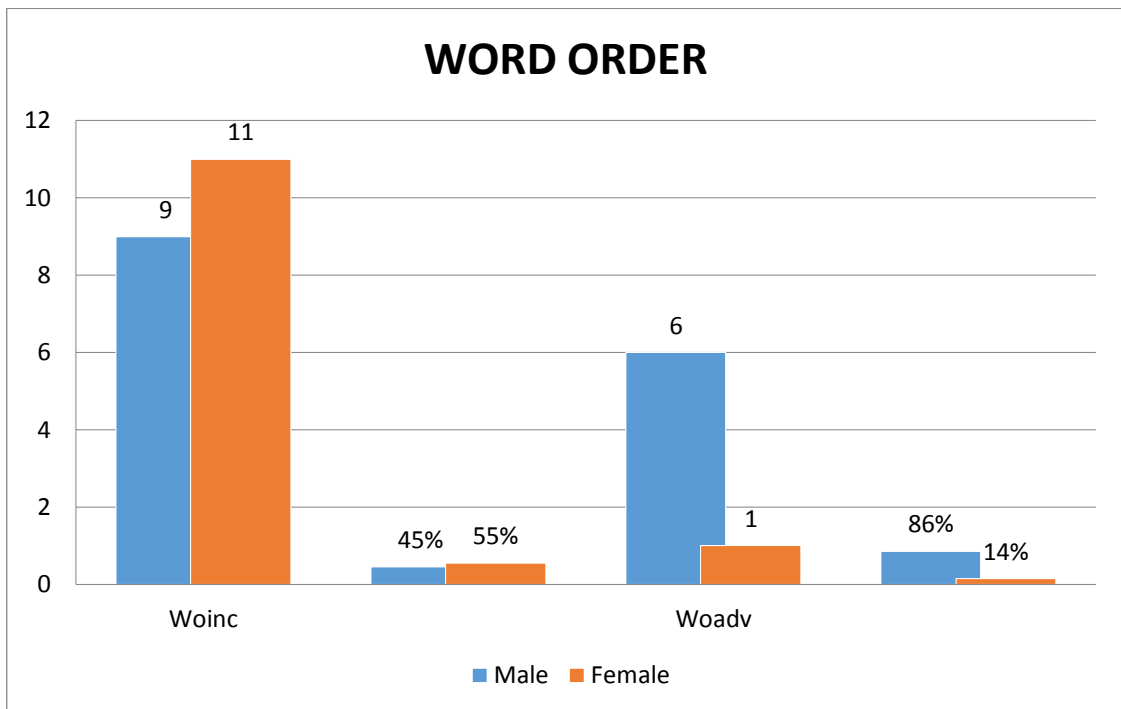


Figure 7: Word Order

Learners generally used an incorrect word order in sentences. As consequence, corpus presents 20 errors in this category so female gender committed the 55% of errors and males 45%. This result denotes that men are less susceptible to commit word order errors than women. In the sentences below, there are some errors to observe.

- I also (WOinc) like going to the disco to dance, (srun) although it is not often, but I do like it (WOinc).
- I love them very much and have I (WOinc) a lot of fun with them.

Word order is a common problem in students because of Spanish interference and did not change the adjective position. This figure shows to 4 students who wrote incorrectly the adjective position. It corresponds to: 86% of errors were committed by men and the other 14% were committed by women.

4.8. Transitions, Mechanics and Redundancy

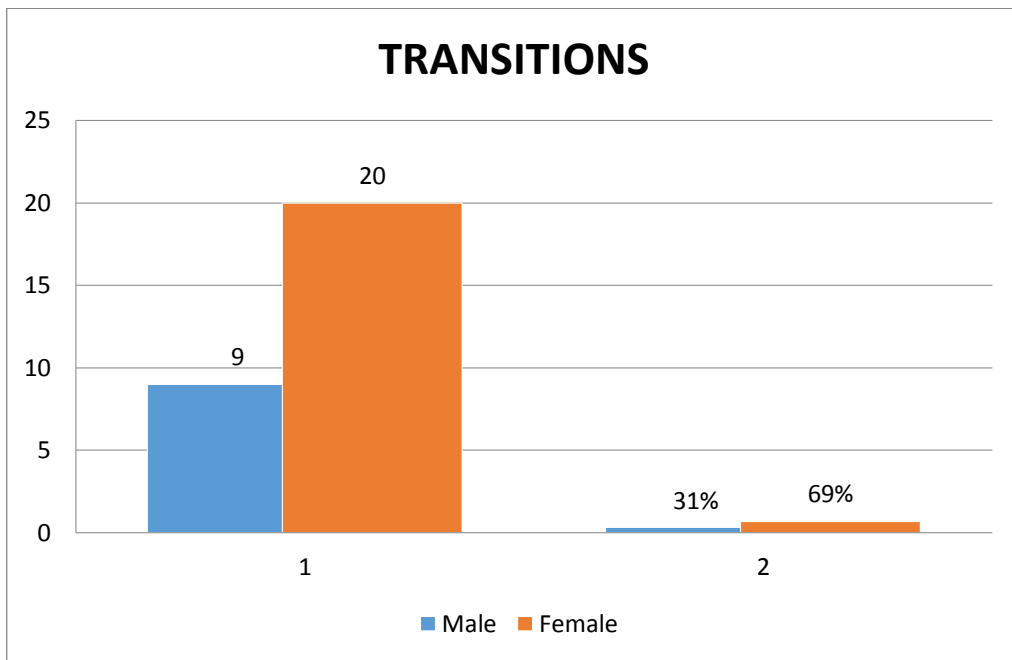


Figure 8: Transitions

Around 16 students committed 29 errors in this category, so 69% of errors belong to female gender and the other 31% for male gender. As result of this analysis, the majority of the students have problems to use linking words because of their limited vocabulary.

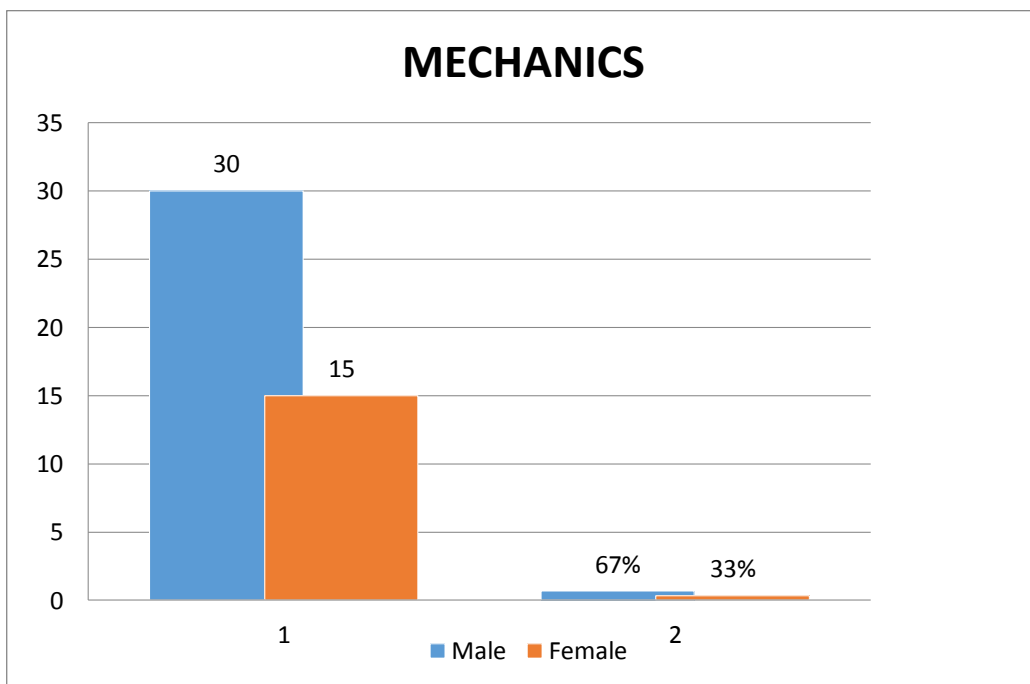


Figure 9: Mechanics

AntConc showed in the screen 45 errors in 18 paragraphs, so it reflects that students have difficulties with this error category. They presented problems in capital letters, translation words, missing words and with correct writing of some words. 67% of errors were done by male gender and 33% were done by female gender. In this case women are better than men with translation words and remembering its writing.

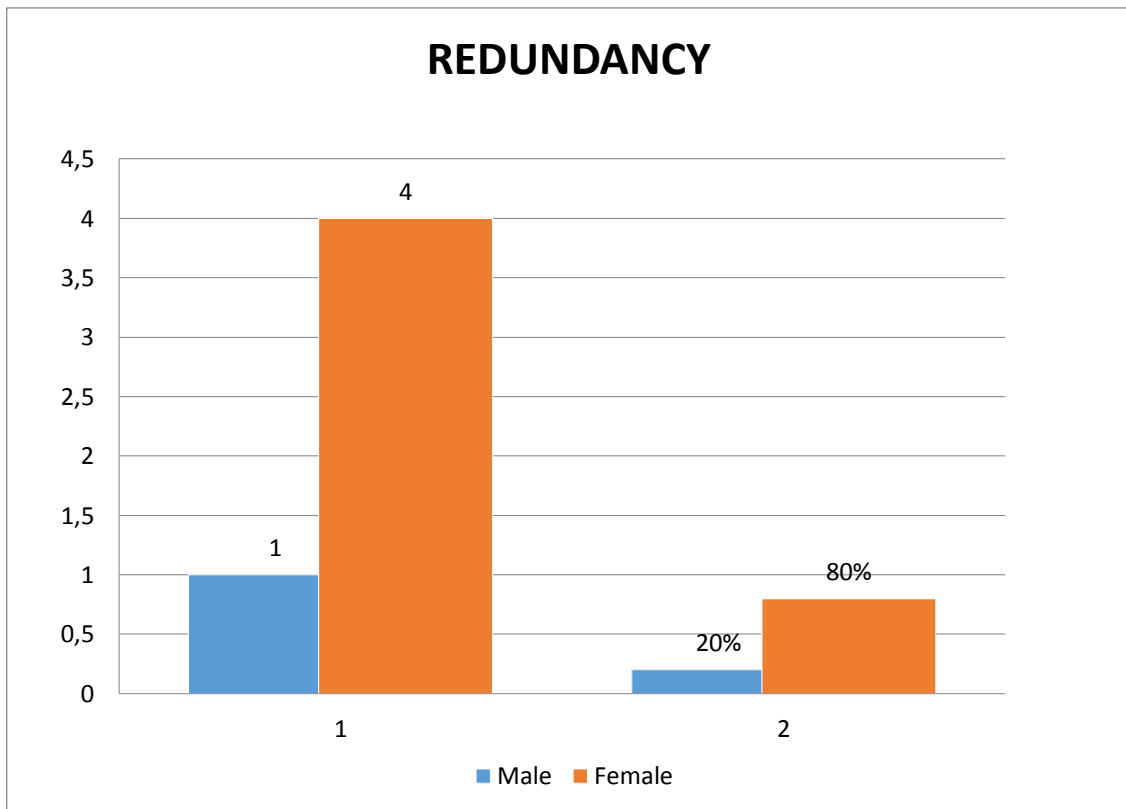


Figure 10: Local Redundancy

Fortunately, just 5 errors were counted in this category. 80% correspond to female gender and the other 20% correspond to male. Students are not so vulnerably to committee errors in redundancy category, but some of them had a limited vocabulary and are more susceptible to repeat some words when they want to number features, descriptions or actions. For example:

- On our weekends we always try to make each one different, We go for a walk, we (Rloc) play, we do different meals, we talk, we laugh and we love to spend together.
- I prefer to go to some place for to (Rloc) eat.

To sum up, male students committed a total sum of 119 errors (averaging 10,82 per individual). Meanwhile, women, being more abundant, only summed up to 109 errors (an average of 8.38 per individual). Males were especially outperformed in noun and article use and spelling (mechanics), while women tended to be more redundant and did especially worse at the use of transitional connectors. A complete chart of figures can be found in Appendix 4

5. DISCUSSION

Writing skills is important to know how people can express their thoughts through this language skill. This investigation was aimed to analyze the main difficulties by students from 6th level of General English at PUCE Esmeraldas to identify the most common errors in grammar establishing differences in range of errors among male and female learners. Result figures show that students committed around 228 errors in total. Out of them, the highest percentage of errors corresponded to males; while with a minimal difference the rest of errors were committed by females (see Appendix 4 for more details).

In a similar fashion, research performed by different researchers reveal that female language learners commit less writing errors, show more improvement in writing over time or gain higher writing scores than male learners (Chen, 1996, Abu-Jarad, 2008, Chiu, 2008, Al-Nawas, 2009). However, other studies about gender variation in grammatical errors, some researchers as Boroomand. & Rostami Abusaeedi (2013) presented a study conducted on 100 Iranian advanced EFL learners' written errors (50 male learners and 50 female learners). In this case, they offered different results stating that more error frequency appeared in female written productions. However, they denoted different classifications and subdivisions of errors, so the authors detected the most committed errors in each classification.

Moreover, the study conducted by Darus et. al (2009, p. 493) provide results, presenting that grammatical errors are the principal problem detected on writing skill. The author concluded stating that the errors committed by participants were basically grammatical. Further, emphasis indicated that the problems of the participants had in acquiring normal grammatical rules. Grammar is also considered as one of the most crucial factor in language teaching and learning.

As of local precedents, recent observations made by López (2017) on students' grammar mistakes concluded that they also made mistakes in the use of the form FOR+ gerund for example: correct way (for studying) and incorrect way (for study) and many of them wrote them incorrectly. Apart from this, they also showed mistakes in the use of the verb to be, some of them did not use properly that grammatical pattern, which is important and relevant in the development of students' grammar. In similar terms to us,

learners are not clear with simple English grammar, like the verbal use of “to be”; grammar use is concerning as it causes an important part of grammar errors.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

- Women perform better than man when they are writing in EFL. The highest percentage of concern male individuals while the rest were committed by female, so students presented more problems in writing errors referring to punctuation, followed by mechanic errors, such as capitalization and spelling. Finally, the third high range of writing errors were done in the use of transition words.
- Corpus provides an interesting and easy way of observing the English learning process as it provides relevant information, while data are considered fast to obtain, realistic, aseptic and confidential. It facilitated the analysis of a small 3172 token corpus in few minutes, which allowed to spend more time observing language use than obtaining data.
- In the current research students presented a high percentage of errors in the location of periods and commas, but in other studies they presented more problems in subject verb agreement referring to singular or plural verb as Mohammed (2008). In contrast, other researcher found the majority of difficulties in verb tense referring to the use of infinitive (López, B. 2017).
- In relation to previous research, results were unexpected as other studies showed a higher level of grammar errors in verb tense or adjective position, but in this case, students demonstrated an elevated level of English. Surprisingly, most of the errors were related to punctuation.

- This is one of the first approaches to corpus studies on learner's output in Esmeraldas. Even if it has been a minimal scale analysis, the experience and results are encouraging. As Sinclair stated, 'bigger is always better', so further corpus-based analyses will be possible in the future by expanding the sample and population.

6.2. RECOMMENDATIONS

- To initiate a program in which English teachers reinforce writing skills to students in classroom emphasizing grammatical rules as punctuation, spelling and the use of transition words without forget the practice of topics as present simple tense, verb to be, present progressive, simple past tense and others.
- PUCE Esmeraldas should provide training in corpus-based analysis to be applied as a tool in EFL classes which would provide a powerful empirical tool for teacher's research, in order to obtain large amounts of language samples from real situations
- This investigation should be continued with further pedagogic, linguistic and extra linguistic researches. Then, with a corpus design training, to obtain more observations about learners to contribute with more quantitative real language studies in different levels in Esmeraldas in order to determine the sociocultural causes of the higher prevalence of errors in men in contrast with women of the same level.
- To continue with the development of this learner corpus with the widening of the information through the extension of this corpus during the years to come.

7. REFERENCES

- Abu-Jarad, H. A. (2008). Evaluating grammar development through longitude error analysis of English major students. *Al-Aqsa University Newsletter*, 12.
- Aijmer, K. (Ed.). (2009). *Corpora and language teaching* (Vol. 33). John Benjamins Publishing. Retrieved from:
https://books.google.com.ec/books?hl=es&lr=&id=Il-hYaUJLu0C&oi=fnd&pg=PR1&dq=corpora+in+language+teaching&ots=Wbf1P17yfH&sig=IXzpu0aNAkFyGdE35BbHv86MgTI&redir_esc=y#v=onepage&q=corpora%20in%20language%20teaching&f=false
- Al Saeed, N., & Waly, S. (2009, May). Corpus for classrooms: Ideas for material design. In *Proceedings of the 10th METU ELT Convention*. Ankara, Turkey.
- Ali, H. O. (2016). Gender differences in using language in the EFL classes: From teachers' views. *International Journal of Humanities and Cultural Studies (IJHCS) ISSN 2356-5926*, 2(4), 73-91.
- Anthony, L. (2014). AntConc [Software]. *Laurence Anthony's*.
- Arppe, A., Gilquin, G., Glynn, D., Hilpert, M., & Zeschel, A. (2010). Cognitive corpus linguistics: Five points of debate on current theory and methodology. *Corpora*, 5(1), 1-27. Retrieved from:
<http://www.eupublishing.com/doi/abs/10.3366/cor.2010.0001>
- Bennett, G. R. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers*. University of Michigan Press.
- Benson, C. J. (2002). Real and potential benefits of bilingual programmes in developing countries. *International Journal of Bilingual Education and Bilingualism*, 5(6), 303-317.
- Bernat, E., & Lloyd, R. (2007). Exploring the gender effect on EFL learners' beliefs about language learning. *Australian Journal of Educational & Developmental Psychology*, 7(1), 79-91. Retrieved from:
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.452.5914&rep=rep1&type=pdf>

- Boroomand, F., & Rostami, A. A. (2013). A gender-based analysis of Iranian EFL learners' types of written errors. *International Journal of Research Studies in Language Learning*, 2(5), 79-92.
- Castillo Jaén, S. J. (2016). *Do corpora benefit the level of vocabulary of 10th year of basic education students? case study in don Bosco Educational Institution in Esmeraldas, 2015-2016*(Doctoral dissertation, Ecuador-PUCESE-Escuela de Lingüística Aplicada mención Inglés).
- Chen, L. L. (2006). The effect of the use of L1 in a multimedia tutorial on grammar learning: An error analysis of Taiwanese beginning EFL learners' English essays. *The Asian EFL Journal Quarterly*, 8(2), 76-110.
- Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and brain sciences*, 22(6), 991-1013.
- Constitution of Ecuador, (2008). "Good life plan". Constituent Assembly. Retrieved from: https://www.constituteproject.org/constitution/Ecuador_2008.pdf
- Corder, S. P. (1967). The significance of learner's errors. *IRAL-International Review of Applied Linguistics in Language Teaching*, 5(1-4), 161-170.
- Dahlmeier, D., Ng, H. T., & Wu, S. M. (2013, June). Building a Large Annotated Corpus of Learner English: The NUS Corpus of Learner English. In *BEA@NAACL-HLT* (pp. 22-31)
- Darus, S., & Subramaniam, K. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. *European Journal of Social Sciences*, 8(3), 483-495.
- Díez-Bedmar, M. B. (2009). Written learner corpora by Spanish students of English: An overview. In *A survey of corpus-based research [Recurso electrónico]* (pp. 920-933).
- Douglas, D., Selinker, L., Tarone, E. E., Gass, S. M., & Cohen, A. D. (1994). Research methodology in context-based second-language research. *Research methodology in second-language acquisition*, 119-131.
- Flowerdew, J. (2009). 19 Corpora in Language Teaching. *The handbook of language teaching*, 327.

- Golomb, J., Kluger, A., De Leon, M. J., Ferris, S. H., Mittelman, M. P., Cohen, J., & George, A. E. (1996). Hippocampal formation size predicts declining memory performance in normal aging. *Neurology*, 47(3), 810-813.
- Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching. *Corpora and language teaching*, 33, 13-32.
Retrieved from: https://books.google.com.ec/books?hl=es&lr=&id=3yfOJSMTRhoC&oi=fnd&pg=PA13&dq=learner+corpora+granger&ots=wsiMXIVZxA&sig=LE6BRPVXR4OQ2m8PFQtENoAqHRI&redir_esc=y#v=onepage&q=learner%20corpora%20granger&f=false
- Halpern, D. F., & La May, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, 12(2), 229-246.
- Ho, C. M. L. (2003). Empowering English teachers to grapple with errors in grammar. *The Internet TESL Journal*, 9(3).
- Hourani, T. M. Y. (2008). An analysis of the common grammatical errors in the English writing made by 3rd secondary male students in the Eastern Coast of the UAE. Retrieved from: <http://bpace.buid.ac.ae/bitstream/1234/225/1/20050055.pdf>
- Hsu, J. Y., & Chiu, C. Y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *Asian EFL Journal*, 10(1), 181-204.
- Hunston, S. (2002). *Corpora in applied linguistics*. Ernst Klett Sprachen.
- James, C. (2013). *Errors in language learning and use: Exploring error analysis*. Routledge.
- Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *Tesol Quarterly*, 37(3), 467-487.
- Kimura, D., & Clarke, P. G. (2002). Women's advantage on verbal memory is not restricted to concrete words. *Psychological reports*, 91(3_suppl), 1137-1142.
- Kindt, D., & Wright, M. (2001). Integrating language learning and teaching with the construction of computer learner corpora. Retrieved May, 22, 2003.

- Kramer, J. H., Delis, D. C., Kaplan, E., O'donnell, L., & Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology*, 11(4), 577.
- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647-672.
Retrieved from:
http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9922.2010.00621.x/epdf?r3_referer=wol&tracking_action=preview_click&show_checkout=1&purchase_referrer=onlinelibrary.wiley.com&purchase_site_license=LICENSE_DENIED
- López Estupiñán, B. (2017). *Common written english grammatical mistakes in the students from the applied linguistics school at PUCE in Esmeraldas* (Doctoral dissertation, Ecuador-PUCESE-Escuela de Lingüística Aplicada).
- Mahidin Abd Kadir, N. (2014). *Grammatical errors across proficiency levels in the spoken English of matriculation students/Nurjasmin binti Mahidin Abd Kadir* (Doctoral dissertation, University of Malaya).
- McEnery, T., & Xiao, R. (2010). *Corpus-based contrastive studies of English and Chinese*. Routledge.
- Meunier, F., & Gouverneur, C. (2009). New types of corpora for new educational challenges. *Corpora and language teaching*, 179-201. Retrieved from: https://books.google.com.ec/books?hl=es&lr=&id=3yfOJSMTRhoC&oi=fnd&pg=PA179&dq=Types+of+Corpora+pdf&ots=wsiMXIU2tG&sig=Yomfq-Y7MIPq0AIFj_boS8D7R1M&redir_esc=y#v=onepage&q=Types%20of%20Corpora%20pdf&f=false
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art—and beyond. *Corpus technology and language pedagogy: New resources, new tools, new methods*, 5-24
- O'keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press
- Oxford, R. L., & Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for

- Language Learning (SILL). *System*, 23(1), 1-23. Retrieved from: <http://www.tesl-ej.org/wordpress/issues/volume5/ej20/ej20r4/7>
- Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in cognitive sciences*, 6(11), 456-463.
- Richards, J. C., & Renandya, W. A. (Eds.). (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge university press.
- Rundell, M. (2002). *Macmillan English dictionary: for advanced learners*. Macmillan.
- Sinclair, J. (1987). *Collins COBUILD English language dictionary*. Harper Collins Publishers.
- Squire, L. R., & Knowlton, B. J. (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. *The new cognitive neurosciences*, 2, 756-776.
- Tribble, C. (2010). What are concordances and how are they used. *The Routledge handbook of corpus linguistics*, 167-183.
- Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of cognitive neuroscience*, 9(2), 266-276.
- Varley, S. (2009). I'll just look that up in the concordancer: integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, 22(2), 133-152.
- Xiao, R., & Tao, H. (2007). A corpus-based sociolinguistic study of amplifiers in British English. *Sociolinguistic studies*, 1(2), 241-273.
- Zoghi, M., Kazemi, S. A., & Kalani, A. (2013). The effect of gender on language learning. *Journal of Novel Applied Sciences*, 1124-1128.

8. APPENDICES

Appendix 1

docs.google.com/forms/d/e/1FAIpQLSe0KWtJ4mKPk0b7u4VB8QmihUG8ph1CduqftG

Writing in Students of the PUCE Esmeraldas

Name:

Tu respuesta

Gender

- Male
- Female

Age

- 17-18
- 19-21
- 21-25
- 25+

Appendix 2

[.google.com/forms/d/e/1FAIpQLSe0KWtJ4mKPk0b7u4VB8QmihUG8ph1Cc](https://www.google.com/forms/d/e/1FAIpQLSe0KWtJ4mKPk0b7u4VB8QmihUG8ph1Cc)

How many years have you studied English?

- 1-2
- 3-4
- 5-6
- 7-8
- 9+
- Otro:

What degree are you following?

Tu respuesta _____

EXAM Question: Write a short text of your habits during the weekend. You can talk about likes and dislikes, favourite places, hobbies, personal relationship. You can also write about the best weekend you have ever had. (100 words aprox)

Tu respuesta

EXAM Question: Do you like to travel? Tell us about the places that you have visited and the places you would like to go. (100 words aprox.).

Tu respuesta

Appendix 3

	B	C	D	E	G
1	: Boris	Male	21-25	I like (Vf) travel, (srun) I went to 3 countries, (srun) I	5-6
2	: Martin Stopper	Male	21-25	i (mec) like (vf) made (vt) differents (WOadv) things	9+
3	: Marlon Andrés V	Male	21-25	I do my homework on weekends, I (trans) also(WOinc)	7-8
4	: Yilio Prado	Male	21-25	Well, every weekend I travel to Limones. I love to p Sunday at noon we(Pform) usually go to (**)church	7-8
5	: Karol Zamora	Female	21-25	I like traveling a lot, one of my biggest dreams is to	9+
6	: Maria Jose More	Female	19-21	I like traveling much. I have visited (ArtOrDet) small the (mec) places visited are: Baños, Puyo, Yasuni, In the yasuni was found diversity of plants and anim I would like to travel the (**) world, after I finish my Travel places like: the seven wonders of the world, I	3-4
7	: Jair Abdel Monta	Male	21-25	I'm studing at Catholic University in Esmeraldas, bu	9+
8	: José Miguel Vera	Male	19-21	In my weekend (srun) I like to play the guitar, (trans) The best weekend (ArtOrDet) I had, was when I was	3-4
9	: Carlos Castro	Male	21-25	I like to travel, I have visited some places of Ecuad	9+
10	: Varinia Guacham	Female	21-25	I like to have fun, to spend (WOinc) with my family	3-4
11	: Johan Garces	Male	21-25	I like to play soccer in my free time, because I am a I also like to spend time with my family because I	3-4
12	: Mayra Angulo	Female	21-25	My best weekend was spent with my family on the	3-4
13	: Nicole Arizala	Female	19-21	My weekend is much more than a time dedicated (V My best weekend was when I went to Baños with m (trans) And in the evenings we went out (WOinc) to	3-4
14	: Soliz Castillo Xic	Female	19-21	I have visited Guayaquil, Quito, Papallacta, Baños,	3-4
15	: Katherine	Female	21-25	My best weekend or my best weekends are (Vform)	3-4
16	: Raquel Suarez	Female	19-21	"If I have traveled I went to Colombia (trans) was of My family lives there, that's why I got there, (srun) I would like to visit Europe because I find it very int "	7-8
17	: Paola chilian toal	Female	21-25	IF I LIKE TRAVELING (**) MY FAVORITE PLACE	5-6
18	: ROGER VERA	Male	21-25	My habits during the weekend are do excercise in th	3-4
19	: Kenia Vásquez	Female	21-25	I like to spend time with my family and go out to ha	7-8
20	: Eduardo Vera An	Male	21-25	During the weekend I like (vt) go to the beach with In the weekend I take the day to do the most duties	9+
21	: María Nazareno	Female	21-25	The Saturday usually I am dedicated to see series	9+
22	: Fernando Franco	Male	21-25	On weekends I like to go to the beach to sunbathe,	9+
23	: Josselyn Mera	Female	19-21	I'm a person who like the quiet places. For this reas During the weekend, I prefer to stay in (**) my hom	3-4
24	: Melina Parrales	Female	19-21	I like to travel and I have known different place (nn) I would like to know Paris as (trans) I like to learn al	9+

fx During the weekend I like (vt) go to the

B A [List Icon] [Eraser Icon] [Clipboard Icon] [Share Icon]

Appendix 4

ERROR TAG	ERROR CATEGORIES	MALE	FEMALE	TOTAL
VERBS				
VT	VERB TENSE	12	8	20
VFORM	VERB FORM	4	1	5
SUBJECT-VERB-AGREEMENT				
SVA	SUBJECT-VERB-AGREEMENT	1	1	2
ARTICLES/DETERMINERS				
ArtOrDet	ARTICLE OR DETERMINER	4	2	6
NOUNS				
Nn	NOUN NUMBER	2	3	5
Npos	NOUN POSSESSIVE	2	2	4
PRONOUNS				
Pform	PRONOUN FORM	6	1	7
Pref	PRONOUN REFERENCE	0	5	5
Wcip	WRONG COLLOCATIONS, IDIOMS AND PREPOSITIONS	2	3	5
SENTENCE STRUCTURE				
Srun	RUNNONS, COMMA SPLICE	31	32	63
WORD ORDER				
Woinc	INCORRECT SENTENCE FORM	9	11	20
Woadv	ADVERB/ADJECTIVE POSITION	6	1	7
TRANSITIONS				
Trans	LINK WORDS/PHRASES	9	20	29
MECHANICS				
Mec	PUNCTUATION, CAPITALIZATION, SPELLING AND TYPOS	30	15	45
REDUNDANCY				
Rloc	LOCAL REDUNDANCY	1	4	5
TOTAL		119	109	228
		11	13	24
Average		10,82	8,38	9,5
%		56%	44%	

Appendix 5

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Concordance Hits 63

Hit	KWIC	File
51	cional and international soccer (srun) also play videogames in computer or play	gaby_18.txt
52	n computer or play station and (Srun) I dislike give me orders and do	gaby_18.txt
53	places. My girlfriend is Josselyn (srun) have (vt) 21 years old (srun) is the	gaby_18.txt
54	yn (srun) have (vt) 21 years old (srun) is the best student of the carrier	gaby_18.txt
55	out to have fun with my friends,(srun) I love playing volleyball every weekend and	gaby_19.txt
56	o to the party always weekend (srun) but my mother dislike her when i (gaby_20.txt
57	e place where they sell seafood (srun).</text> <years>9+</years> Ingenier\xEDa en G	gaby_20.txt
58	ams on the internet. (Mec) Also (srun) clean and ordered my home and washed	gaby_21.txt
59	uties of the university. At night (srun) I go out with my family to	gaby_21.txt
60	med exercises with my cousins (srun) and (trans) finally study my classes and	gaby_21.txt
61	at I really like. Some weekends (srun) I usually travel to visit my grandmother	gaby_22.txt
62	h your family and friends. Also (srun) I like to visit Guayaquil because it'	gaby_23.txt
63	mong the places are Guayaquil (srun) this city is very beautiful, it is (gaby_24.txt

Search Term Words Case Regex Search Window Size 50

Start Stop Sort

Kwic Sort Level 1 1R Level 2 2R Level 3 3R

Total No. 24 Files Processed

Clone Results

Appendix 6

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

File View Hits 2 File gaby_1.txt

<gender>male</gender> <age>21-25</age> <text>I like (Vf) travel, (srun) I went to 3 countries, (srun) Republica (Wform) Dominicana (WOadv), United States an (mec) Mexico, (srun) this (nn) countries is (vform) impresionated (mec), (srun) (Pform) have diferent (mec) places wonderful(WOadv) and (srun) the culture is diferente (mec), (trans) the food is delicious. In United States I went to Miami and Florida, in Republic Dominican (WOadv) I went to Santo Domingo, (srun) an (mec) in Mexico I went (Wcip)a diferent (mec) city but for my(Pform) the most exciting was Acapulco. I want to go this year to Peru and the next year to Cuba, for me Cuba hawev (sva) a (ArtOrDet) culture (WOadv) interesting. I like (vform) travel around the world because (mec) learn diferent (mec) languages,(trans) culture, (srun) (mec) have (ArtOrDet) good experience.</text> <years>5-6</years> Comercio Exterior 6to

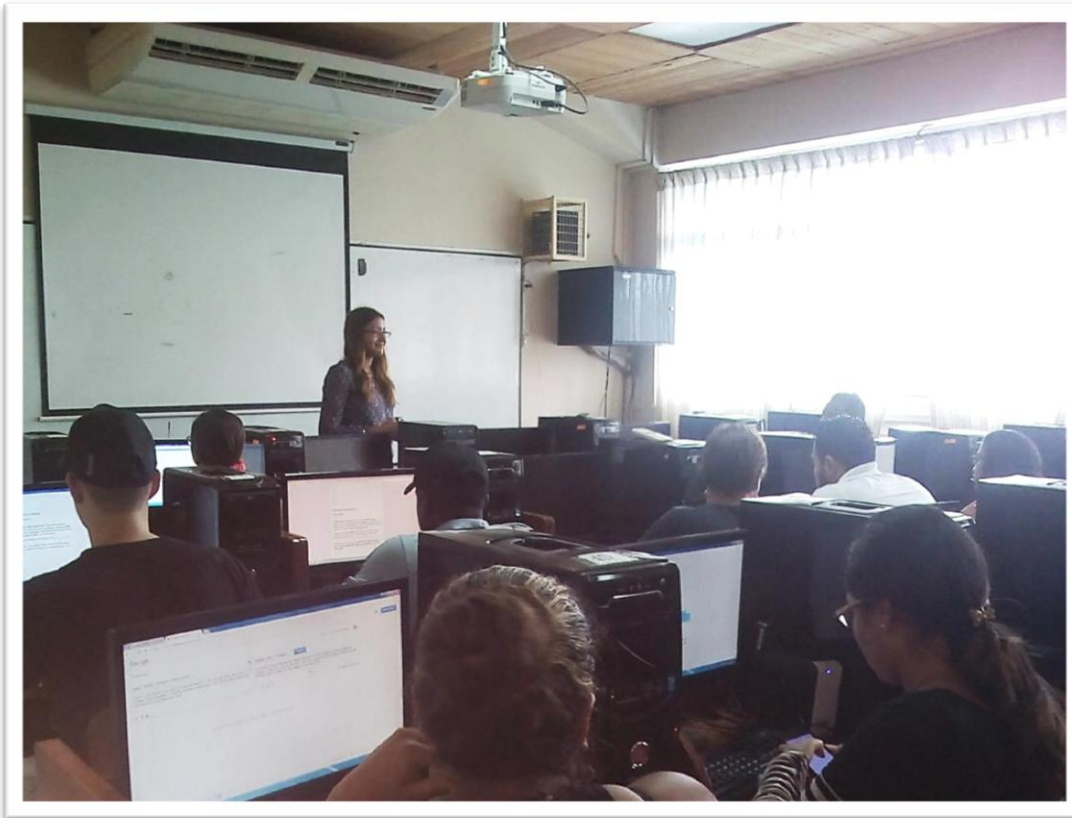
Search Term Words Case Regex Hit Location 2

Start Stop

Total No. 24 Files Processed

Clone Results

Appendix 7



Appendix 8

