The Use of Discourse Markers in Argumentative Compositions by Learners of Spanish as a Foreign Language

by

Sofia Lorena Sanchez Sanchez

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Arts

in
Applied Linguistics

Department of Modern Languages and Cultural Studies
University of Alberta

Abstract

The role of discourse markers in determining the quality of compositions has received increasing attention in the literature on second language learning. The present study investigates the use of discourse markers in compositions by students of Spanish as a foreign language, a topic that, so far, has barely been addressed in the field of ELE (Español como Lengua Extranjera "Spanish as a foreign language"). This study examined the frequency, the variety, and the accuracy in the use of discourse markers in argumentative compositions by learners of Spanish as a foreign language. The study also explored the correlation between the use of discourse markers (ratio, tokens, and variety) and the quality of the compositions. A mixed method approach was used; a total of 64 compositions from 5 levels of proficiency were analyzed to identify connectors, discursive operators, and metatextual connectors, following Calsamiglia & Tusón's (2001) classification of discourse markers. Findings from the study revealed that there is a statistically significant increase in the number and the variety of discourse markers used from beginners to advanced levels. The accuracy in the use of discourse markers decreased as the proficiency level increased. Furthermore, discourse markers were shown to be more important for the quality of the compositions at the advanced levels; a higher number and a broader variety of discourse markers correlate more strongly with quality scores at the higher levels. The results of the study contribute to an understanding of the use of discourse markers in Spanish L2 compositions.

Preface

This thesis is an original work by Sofia Lorena Sanchez Sanchez. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "Discourse Markers and the Quality of Argumentative Writing Among University Learners of Spanish as a Foreign Language," No. Pro00068197, 14/10/2017.

For Gaby

Acknowledgments

This thesis was financially supported by the Mexican National Council of Science and Technology (CONACYT) Scholarship.

I would like to express my gratitude to my supervisor, Dr. Yvonne Lam, for her guidance and endless patience for the completion of this project and during this phase of my life. I also want to extend my appreciation to Dr. Odile Cisneros and Professor Anne Malena for their support and encouragement to pursue my goals.

To Leanne, Bashair, Delma, Sarah , Houssem, Amelia, Raquel, Vanesa, and Manu, I want to give thanks for sharing laughs and contributing to my psychological well-being.

I would also like to thank my sister, Lola, for holding my hand along this very bumpy road; my brother, Pepe, for reminding me that we should always reflect the beauty of the world through our art; and Dan, for seeing through my eyes and opening a new door.

Last but not least, I give thanks to my parents for all their love, encouragement and support to succeed in every single endeavor I have pursued, from learning new languages to becoming a dancer, and lately, to change our geographies.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Every language instructor, or any instructor for this matter, has likely asked herself at some point: “How am I supposed to cover all these topics in one term? How are my students going to learn all of this? And —in the case of language classes— how could my students possibly develop all these skills in 50 hours of class time?” As language instructors, we all face the challenge of fitting the expected outcomes for the unit's reading, writing, speaking, and listening components into our class planning. Fifteen minutes are a treasure —should I include this game, or would it be better to...? We must ponder each activity and each topic; as a result, some aspects of the language receive less, or in many cases, very little attention. The textbook is the most common didactic material used in language classrooms, and the instructors may rely on these materials to plan their lessons; in consequence, if a topic is not part of the unit content, it may be unintentionally neglected. In this respect, Nogueira da Silva (2011, 2012) reported that discourse markers are scarcely included in ELE (Español como Lengua Extranjera / Spanish as a Foreign Language) textbooks. This research investigates the use of discourse markers in ELE compositions and its correlation with their overall quality. It aims to contribute to the current literature to guide the decision-making regarding teaching discourse markers in the ELE classroom.

It has been broadly described in the literature that writing in another language (L2) is a complex process. In this context, Intaraprawat and Steffensen (1995) said that L2 writers face many challenges: using a different morphology, syntax, and lexicon; and learning the new conventions of the L2 community. Alderson (2002) addressed how L2 writers struggle with the cultural practices of the target language, such as type of text conventions, and their current writing practices. Also, Berril (1996) talked about the difficulties related to the learners' first

language (L1) writing skills; for instance, a first-year undergraduate student with poor writing skills in the L1 may not intuitively consider using discourse markers in a L2 composition. In order to offer adequate support to students, instructors need to consider that L2 learners will, generally, have fewer opportunities to practice and improve their writing skills in settings other than the institution where language instruction takes place. This is also true for instruction on discourse markers. The input we provide and the explicit instruction on the use of these lexical expressions is the main contact the students will have with them.

The adequate use of discourse markers is a fundamental part of mastering an L2. Once the learner has mastered the basics of the target language, these elements of discourse are fundamental for a better production and understanding of the different discursive genres that are produced in a L2 (Nogueira da Silva, 2012, p. 79). Discourse markers are an essential part of language interactions; their pragmatic function is crucial for the adequate inference and production of discourse. For example, Mirzaei Jegarlooei and Allami (2018) researched the relevance of discourse markers in the perception and production of politeness. Lahuerta and Pelayo (2003) explored the difficulties L2 Spanish learners experience in interpreting discourse markers in authentic newspaper texts. There is an abundant literature about the role of discourse markers in English-as-second-language (ESL) compositions; in fact, almost any contemporary ESL textbook devotes some space to address their use, either by providing a chart, an appendix, or a box calling for attention to the use of these elements.

In the last few decades, the study of discourse markers has gained relevance. As Meléndez Quero (2010) said, the study of discourse markers is one of the most interesting topics in the field of foreign language teaching; it is, as well, one of the most complicated ones (p.138). In Spanish, there are multiple theoretical and methodological approaches to the study of these

expressions; they have been classified as *enlaces supraoracionales*, *conectivos*, *conectores pragmáticos*, *ordenadores (léxicos) del discurso*, *operadores pragmáticos o discursivos*, *marcadores de interacción*, *marcadores del discurso*, among others (Mancera & Placencia, 2011, p. 147). The multiplicity of names for discourse markers in Spanish reflects the conceptual complexity of these elements, and, as stated by Montolío, it allows us to understand the difficulty experienced by L2 learners to master the connectors and, in general, the discourse markers in a foreign language (1998, p. 114).

One of the most important definitions of discourse markers is from Martín Zorraquino and Portolés:

> Los marcadores del discurso son unidades lingüísticas invariables que no ejercen una función sintáctica en el marco de la predicación oracional y que poseen un cometido coincidente en el discurso, a saber: el de guiar, de acuerdo con sus distintas propiedades morfosintácticas, semánticas y pragmáticas, las inferencias que se realizan en la comunicación (Martín Zorraquino y Portolés, 1998, § 63.1.2)[1].

Calsamiglia and Tusón (2001) explain that discourse markers give cohesion and structure to the text and that they are used to guide the reader or listener to interpret the text. In this respect, De Beaugrande and Dressler (1981) propose that coherence and cohesion are two of the seven principles of textuality. In academic compositions, it is fundamental to produce cohesive and coherent texts; in order to do so, the writer needs to develop a pragmatic-discursive skill which involves successful lexical selection according to the discursive mode, and the adequate use of discourse markers (Andriessen & Coirier, 1999; Calsamiglia & Tusón, 2001).

[1] "Discourse markers are invariable linguistic units that don't play any syntactic function in the predicate of the sentence, and they have a role in the discourse, that is to guide, depending on their distinctive morpho-syntactic, semantic, and pragmatic properties, the inferences that are made in communication." (my translation)

Research in the field of Second Language Writing suggests that discourse markers are part of the verbal devices that can help produce effective and satisfactory compositions (Rahimi, 2011). Consequently, the lack of discourse markers or the inappropriate use of discourse markers in a foreign language composition could impede successful communication or might lead to the lack of comprehension of the text. As research has shown (Nogueira da Silva, 2011; Calsamiglia & Tusón, 2001; Montolío Durán, 2001), certain discourse markers are more common in a given type of text. For instance, Cuenca (1995) identified three main types of discourse markers in Spanish argumentative writing: contrastive, causal and consecutive, and distributive.

There is a considerable amount of research about the correlation between the use of discourse makers and the quality of the composition. This topic has been addressed from numerous perspectives, as discourse markers, cohesive devices, lexical bundles, type of chunks, among others; however, the findings of these studies have been somewhat contradictory. While some studies found a positive correlation between the use of discourse markers and the quality of the composition, other studies didn't find any statistical correlations.

In the context of ESL, several studies identified a positive correlation between the quality of the composition and the use of discourse markers. Liu and Braine (2005) investigated the use of cohesive devices in 50 argumentative compositions written in English by Chinese undergraduate non-English majors. They found that the quality of writing co-varied significantly with the number of lexical devices and the total number of cohesive devices used. Intaraprawat and Steffensen (1995) analyzed the metadiscourse features of 47 persuasive essays written by ESL university students. They evaluated the use of connectors such as adverbs, adverbial phrases, and coordinators, reminders, and topicalizers. They also analyzed code glosses, illocutionary markers, validity markers, attitude markers, and commentaries. They found that

good essays had more metadiscourse markers; the readability of the text increased with the level of use of these features. Conversely, a low number of metadiscourse elements in the essays suggested a poorer quality of writing because, among other reasons, there was not enough redundancy and attention to the audience. Jalilifar (2008) investigated the types of discourse markers in 598 descriptive compositions produced by Iranian ESL students. The results revealed that discourse markers were used with different degrees of occurrence. The most frequently used discourse markers were elaborative, followed by inferential, contrastive, causative, and topic-relating markers. He also identified a direct positive relationship between the quality of the compositions and the number of correctly-used discourse markers. Yang and Sun (2012) compared the correct and incorrect uses of cohesive devices in 60 argumentative compositions produced by second-year and fourth-year undergraduate Chinese English-as-a-foreign-language (EFL) learners. The results revealed a positive correlation between the correct use of cohesive devices and the writing quality. Qin and Uccelli (2016) analyzed a total of 200 argumentative and narrative compositions written by Chinese secondary-school EFL learners. They found that lexico-syntactic complexity and diversity of organizational markers were good predictors of the quality of argumentative compositions. Lahuerta (2004) analyzed 78 expository texts produced by first-year ESL university students. The study revealed a statistically significant correlation between the number of discourse markers used and the quality of the compositions.

On the other hand, the findings of two studies contradict the results obtained above. Rahimi (2011) investigated the frequency and types of discourse markers used in the argumentative and expository writing of 47 ESL students. He identified that the most frequently used discourse markers in both essay types were as follows: elaborative markers (mainly “and”), contrastive, inferential, reason, exemplifier, and conclusive markers. His corpus revealed that

discourse markers were more frequently used in the argumentative essays than in the expository ones; the research also showed that the use of discourse markers did not correlate with writing quality. Meisuo (2000) investigated the use of discourse markers in 107 expository texts produced by Chinese university ESL students. This study also didn't reveal any statistically significant correlation between the numbers of cohesive elements used and the quality of the compositions.

In ELE, very little research has been done about the use of discourse markers in terms of frequency of discourse markers used. Rodriguez (2016) analyzed the frequency of use of discourse operators of concretion and argumentative reinforcement in the Corpus of Learners of Spanish as a Foreign Language (CAES). The corpus consisted of 3,878 compositions ranging from A1 to C1 proficiency level. In the entire corpus, he identified only 19 occurrences of discourse operators of concretion and 152 occurrences of discourse operators of argumentative reinforcement. Lahuerta (2002) analyzed 50 compositions from intermediate-level students enrolled in a Spanish literature class at the University of Oviedo and found a direct relationship between the level of competence in writing and the use of discourse markers.

It is clear from the literature review presented above that it is imperative to do more research about discourse markers in the ELE classroom. Not only there are very few studies about how learners use discourse markers, but also —to my knowledge— there are no studies that investigate the correlation between the use of discourse markers and the quality of compositions in Spanish as a foreign language. This research aims to contribute to this discussion by studying the variety of discourse markers used in compositions and by analyzing the correlation between the use of discourse markers and the quality of the composition in ELE. It should be noticed that the variety of discourse markers has not been examined in other studies

of the same kind as Jalilifar (2008) and Rahimi (2011) for ESL, and Lahuerta (2002) for Spanish. In addition, this study explores the use of discourse markers and the errors in their use along five levels of proficiency with the purpose of gathering empirical data that could be used to make instructional decisions on this topic.

The present study was guided by the following research questions:

1. Does the frequency of discourse markers use change with proficiency level?
2. How varied is the use of discourse markers?
3. How accurate is the use of discourse markers?
4. Is there a correlation between the use of discourse markers and the overall quality of the compositions?

The analyses performed provide information that could potentially be used to enhance the explicit instruction of discourse markers in the university context. Ultimately, the goal is to answer the questions of any instructor: *How relevant is this topic? How much emphasis should I give to teach this?* and *What should I focus on?*

Chapter 2. Methodology

This chapter will present a description of the discourse markers considered for this study, the characteristics of the participants, the procedures for the data collection, a description of the tasks, and the analysis procedures.

The present study was designed as a mixed methods research, as defined by Leech and Onwuegbuzie (2009), with the purpose of investigating two aspects of argumentative compositions in Spanish as a foreign language: the patterns of use of discourse markers and the correlation between the use of discourse markers and the quality of the compositions. The qualitative analysis provides information about the use of discourse markers, such as the most frequently used discourse markers per category, and describes the most common mistakes made by Spanish learners. The quantitative analysis investigates the correlation between the use of discourse markers and the overall quality of the compositions; these analyses combined increase the accuracy of the data obtained (Denscombe, 2008). This study not only outlines the stages in the use of discourse markers but also documents the challenges that learners face while integrating them into their compositions.

This research explores the use of discourse markers at different levels of proficiency, seeking to provide a more complete picture of the phenomenon than previous studies. The data collected includes writing samples from beginner to advanced-level learners. While the number of samples for some levels is low, this study serves as a starting point for future research regarding the use patterns of discourse markers.

2.1 Participants

All the participants were enrolled in Spanish foreign language classes at a postsecondary institution in Western Canada. The Beginners I group corresponds to the first term of language courses, Beginners II group to the second term, Intermediate I to the third term, Intermediate II to the fourth term, and Advanced to the fifth term.

A computer-based background questionnaire (Appendix 5) was administered along with the consent form to obtain relevant information that may be significant for this research, such as other languages that the students know, study abroad experiences, writing practices, the use of dictionaries, and peer work practices. The initial sample was 110 compositions; however, in order to have a sample as homogeneous as possible, only texts produced by participants that declared English as their first language were included. Likewise, participants who reported being heritage speakers were not considered for this study. The final sample was 64 compositions.

Regarding participants' language skills, 34 participants only knew English and Spanish (bilinguals), and 29 participants were multilinguals; the classification was done based on self-report, and only the participants reporting the ability to write in another language (or languages) than English and Spanish were labeled as multilinguals. There were 42 females and 22 males, ranging in age from 17 to 27 years, with a median age of 21 years. Table 2.1 presents the demographics.

Table 2.1 *Descriptive Data of the Participants*

	Total no. of participants	Females	Males	Bilingual	Multilingual
Beginners I	21	16	5	11	10
Beginners II	8	2	6	7	1
Intermediate I	23	13	10	12	10*
Intermediate II	8	7	1	4	4
Advanced	4	3	1	0	4
Total	64	41	23	34	29

*One participant did not respond to the question.

2.2 Data Collection

The data were collected in one session per group, at the end of the term. The participants were enrolled in five different levels of proficiency: two groups for Beginners I, one group for Beginners II, two groups for Intermediate I, one group for Intermediate II, and one group for Advanced Spanish. The participants went to a language laboratory where each participant had access to a personal computer. Participants first completed the background questionnaire and gave consent for the use of the materials gathered from the session. Then, the screen-recording tool of *Microsoft* was manually turned on in order to video-record the process of writing the composition. Although the video recordings were not analyzed for this study, it is possible that the awareness that they were being recorded influenced the students' behaviour.

The computers were connected to the Internet and the participants were allowed to use the online resources they considered pertinent, like dictionaries or textbook online resources; however, they were strongly encouraged not to use translation software such as *Google Translate*. As the task was included in the regular compositions done during the language course, the researcher and the instructor were watching for the overuse of the provided resources.

Participants wrote the composition in *Microsoft Word*, with the language settings set in Spanish and the spell checker enabled. Editing and writing tools were provided to emulate the students' current writing practices. In addition, from a capacity theory perspective (McCutchen, 1996), the use of these tools can reduce the time spent revising spelling and typography features, thereby freeing cognitive resources that can be used for more demanding aspects of writing (Barkaoui, 2016). All groups were given 50 minutes to write the text.

With the objective of reducing the Observers' Paradox (Labov, 1972), the topic of the research was not disclosed to the participants. Participants were told that the research was about writing, without giving further details. However, they had a general idea of what the task required, as it was described in the course outline (Appendix 1).

2.3 Tasks

As mentioned above, the tasks were part of the course assignments and were adapted from the regular class tasks to elicit argumentative texts. I chose this type of text because argumentation is present in a lot of discursive acts in private and social life; as Calsamiglia and Tusón (2001) mentioned, we argue in everyday conversation, in a job interview, in publicity, in editorial articles, opinion articles, school papers, etc. (p. 294). In addition, this type of text is known to be the most difficult in comparison with other text types (Gleason, 1999; Richards & Schmidt, 1992), and its form is language specific (Golder & Coirier, 1996). The difficulty of the prompt took into consideration the proficiency level of each group; the topics were linked to the thematic content studied in the fifth unit of each level. The exact instructions given to the

students on the day of the task are presented in Appendix 2 and each is also briefly described below.

Beginners I. The prompts for Beginners I and II require the participants to provide reasons, which are the supporting elements of an argumentative text (Adam, 2011, p. 130). The task for this level consisted of four questions regarding lifestyle. The first was about the activities students do during their free time, the frequency with which they do them, and with whom they practice them. The second question was about the frequency and time they spend with their families and how it benefits their emotional health. The third question required them to mention personal non-healthy habits and ways to improve their general health. The last question asked them to make recommendations for good physical and mental health, and to explain why such advice would be beneficial. The minimum word count was 120 words.

Beginners II. Three ads offering different types of accommodation were presented to the students. The task was to choose a room for a hypothetical study abroad in Madrid, to explain why they choose that option, and to provide information about the type of home (house, apartment…), the location, etc. The minimum word count was 150 words.

Intermediate I. Participants had to select an invention or discovery and form an opinion about why it is the most important discovery or invention in human history. Then, they had to explain the characteristics, usefulness, and advantages of their choice. They were asked to explain why it is the most important invention or discovery and to make reference to how the invention or discovery changed life before and after it appeared. The minimum word count was 200 words.

Intermediate II. Participants had to write a formal petition following the Spanish criteria for an *instancia* (a request letter) in which they had to write to the municipal administration requesting something they considered important and to provide arguments for the request. The minimum word count was 250 words.

Advanced I. Participants were asked to argue if they agree with the idea of free will or not and to support their position. The requested word count was 250-300 words.

2.4 Types of Discourse Markers

Spanish discourse markers have been studied by several scholars in last decades (Portolés, 1998; Martín Zorraquino & Portolés, 1998; Montolío, 1993; Pons, 1998; Cuenca, 1995). I chose to follow the classification presented by Calsamiglia and Tusón in *Las cosas del decir* because it is clearly structured, the labels of the categories are comparable with those for other languages, and it is accessible to both instructors and students, as well as to the academic community. In addition, studies like Nogueira da Silva (2011) follow this classification, promoting a compatible dialog between studies.

Calsamiglia and Tusón (2001) identified three types of discourse markers: connectors, discursive operators, and metatextual connectors. There was no pre-determined list of discourse markers to be identified in the data; the final list of discourse markers was elaborated based on the discourse markers that were used by the participants. The three categories are described in detail below.

Connectors: Calsamiglia and Tusón (2001) defined connectors as discourse markers that establish logical-semantic relationships among textual segments, where the segment can be formed by a statement or a set of statements. These discourse markers provide cohesion to the text and are closely related with the microstructure of the text (p. 247). The authors classified the connectors according to five subcategories: additive connectors, contrastive connectors, causal base connectors, temporal connectors, and spatial connectors. The causal base category, that is discourse markers that have a cause foundation, is subdivided into Causatives, Consecutives, Conditionals, and Final goal. I considered the types of connectors listed below; as there were no occurrences in the data, the subcategory of spatial connectors is not included:

a) Additives: *y* "and," *también* "also," *además* "in addition"...
b) Contrastive: *pero* "but," *sin embargo* "however," *a pesar de "*although"…
c) Causatives: *porque* "because," *puesto que* "given that"...
d) Consecutives: *entonces* "then," *por eso "*therefore," *así que* "so that"…
e) Conditionals: *si* "if," *siempre que* "as long as"...
f) Temporal: *mientras* "while," *entonces* "then," *luego* "then"…

Discursive Operators: Calsamiglia and Tusón describe them as discourse markers that introduce specific discursive operations by indicating the writer's position towards the utterance or by directing towards a specific treatment of the information. These discourse markers give coherence to the text and are closely related with the macrostructure of the text (p. 247). Some examples of this type of discourse markers are:

a) Point of view: *creo que* "I believe that," *en mi opinion* "in my opinion"...
b) Expression of certainty: *es evidente que* "it is evident that," *de hecho* "in fact"...

c) Presentation of a theme: *respecto* "regarding"…

d) Exemplifiers: *por ejemplo* "for example," *pongamos por caso* "as an example"…

e) Reformulation: *es decir* "in other words," *esto es "*that is"…

Metatextual Connectors: These discourse markers indicate the logical development of the text. Metatextual connectors are used in the introduction, body, and conclusion of the text. One of their characteristics is that they are displayed in time and space, contributing to the global organization of the text, and therefore forming the macrostructure of the text (p. 247). Some examples of this type of discourse markers are:

a) Starters: *antes que nada* "first of all," *primero de todo* "first of all" …

b) Distributives: *por un lado* "on the one hand," *por otro* "on the other hand"…

c) Orderers: *primero* "first," *en segundo lugar "*secondly"…

d) Continuatives: *entonces* "therefore," *en este sentido* "in this respect"…

e) Additives: *además* "furthermore"…

f) Conclusives: *en conclusión* "in conclusión," *en suma* "to sum up"…

2.5 Data Analysis

The analysis followed five main stages: the selection and coding of the writing samples and the assessment of their quality, the analysis of the frequencies of discourse markers used, the analysis of the variety of discourse markers used, the error analysis, and the analysis of correlation between the use of discourse markers and the quality of the composition. Below I present a general description of these stages.

The first stage was the selection and coding of the data. For this purpose, all the identifiers were first removed from the samples and the demographic dataset, and each participant was assigned an ID indicating the number of the participant and the level. Second, the final set of compositions was selected using the demographic questionnaire; data from participants who reported having a L1 other than English or who identified themselves as heritage speakers were removed in order to have a sample as homogeneous as possible. The final corpus was 64 samples. Third, the compositions were scored using a rubric (Appendix 3) based on two grading guides similar to the one proposed by Jacobs (1981), but with different weights for the assessment criteria: content (15 pts.), organization (15 pts.), grammar (35 pts.), vocabulary (25 pts.), and mechanics (10 pts.). In order to ensure the reliability of the scores, 30% of the samples per level were randomly selected using *random.org*; then, the samples were scored by two native speakers of Spanish who were experienced second language (ESL) teachers as well as by the researcher. The interrater reliability of the scores was estimated through Cronbach's Alpha formula; the reliability index was 0.94, which indicates that the scoring was highly consistent among the three raters. Finally, the occurrences of all discourse markers, correct and incorrect, were coded and classified by category and subcategory.

The second stage concerned the analysis of the frequencies of the discourse markers used. The density of use of the discourse markers was calculated and compared using an omnibus one-way ANOVA followed by a Tukey's post-hoc test.

The third stage addressed the variety of the discourse markers used. First, a descriptive analysis was performed, followed by an ANOVA and a Tukey's post-hoc test. Second, the categories and subcategories of the discourse markers used were analyzed per group and as a dataset.

The fourth stage was the error analysis. For this purpose, an analysis of variance of the error percentages was performed using a Kruskal-Wallis test, followed by a post-hoc test using Dunn tests. Then, for the cases where the accuracy rate was different from 100%, Pearson correlations on the bootstrapped data (10,000 replications) were performed between the accuracy scores and the total numbers of tokens produced, and the variety of discourse markers used. Finally, the errors were classified into three types: word choice, element left out, and word form.

The fifth stage was the correlation analysis. Bootstrapped (10,000 replications) Pearson correlations were calculated between the quality scores and the ratio of discourse markers used, the variety of discourse markers used, and the total number of discourse markers used.

Chapter 3. Results

3.1 Does the frequency of discourse markers use change with proficiency level?

In order to examine the frequency of use of discourse markers across the levels, I considered the density of use, that is, the number of discourse markers used (both correct and incorrect tokens) as a proportion of the total number of words produced by the group. Figures 3.1 presents a boxplot of the density of discourse markers used across the five levels. The descriptive statistics for the groups are presented in table 3.1.

Figure 3.1 *Density of Discourse Markers*

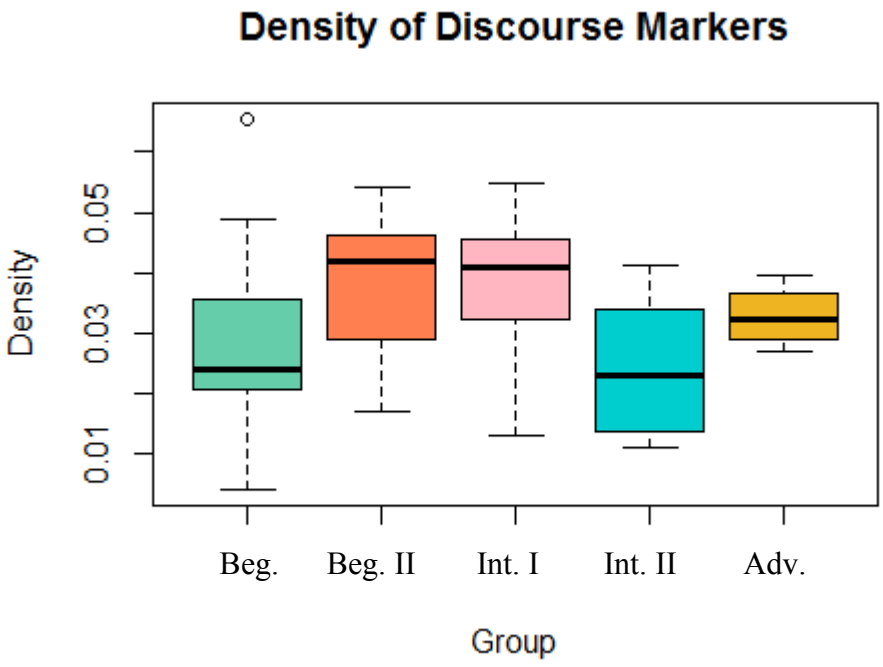


Table 3.1 *Density of Discourse Markers Used*

Group	Number of participants	Total number of words	Total number of discourse markers used	Mean # of discourse markers correct + incorrect tokens (SD)	Density of discourse markers
Beg. I	21	5,286	151	7.19 (4.11)	2.82%
Beg. II	8	1,289	49	6.12 (1.95)	3.82%
Int. I	23	5,554	209	9.08 (2.84)	3.85%
Int. II	8	2,115	49	6.12 (2.47)	2.42%
Adv.	4	1,248	43	10.75 (6.94)	3.27%
Total	64	15,492	501	7.64 (3.57)	3.29 (1.33)

It was hypothesized that the higher levels of proficiency would use more discourse markers independently of the accuracy of use. In order to determine if there were statistically significant differences in the density of use of discourse markers, an omnibus one-way ANOVA was carried out on the data with an alpha value set at 0.05. There was a significant effect of level on the density of discourse markers [$F(4,59)=3.06$, $p=0.023$]. A Tukey's post-hoc test indicated that there were significant differences in the density of discourse markers between the following pairs: a) Beginners I and Intermediate I, and b) Intermediate I and Intermediate II.

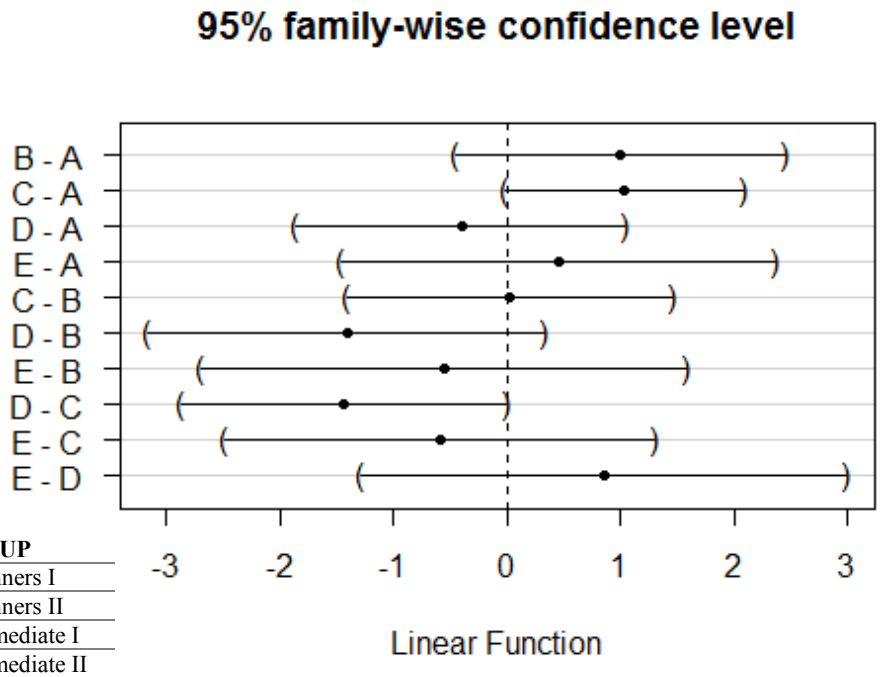
Following Cumming (2012), Larson-Hall and Plonsky (2015), Plonsky (2015) and others, Confidence Intervals (CI) and effect sizes are used rather than traditional *p*-values because *p*-values vary as a function of sample size (Plonsky, 2015, p. 25). Using CIs— "a range of

plausible values for the corresponding parameter"[2] (Kline, 2004, p. 27) that "shows us how much confidence we can have that a point estimate, such as a mean score, is a good estimate of the effect" (Larson-Hall, 2015, p. 130)—provides the researcher with a standardized measure that describes the size of the effect that is being studied; furthermore, the size of the effect can be compared across studies on the grounds that effect sizes do not change no matter how many participants there are (Maxwell & Delaney, 2004, p. 100). When an effect size is presented along with the CI, the researcher and the reader can better interpret how important the result is in practical terms.

The CIs for the post-hoc test are visually presented in Figure 3.2. The arms represent a 95% CI and make predictions about the population mean (μ) that is unknown in real life (Cumming, 2012, p. 79); note that it does not represent the width of the interval in the data analyzed (Larson-Hall, 2015, p. 131). The dotted line in the middle (zero) is used to show whether a comparison is statistically significant or not, similar to an alpha level in a post-hoc test; the lines that go through zero indicate non-statistically significant comparisons. The mean difference for each pairwise comparison is indicated by the dot. For instance, the arm that represents the comparison between groups C and D (Intermediate I and Intermediate II) does not go through zero, indicating that this comparison is statistically significant. The mean in this comparison could be as low as -2.26 or as high as 0.49 in the population, with an estimated mean of -1.43; thus, the CI is noted as [-2.26, 0.49].

[2] For example the true mean, correlation coefficient, difference in scores, etc.

Figure 3.2 *CIs for the Tukey's Post-hoc Test for Density of Discourse Markers Used*



IDENTIFIER	GROUP
A	Beginners I
B	Beginners II
C	Intermediate I
D	Intermediate II
E	Advanced

Given the fact that the number of samples per group is very small, an ANOVA and a post-hoc test would lack power; consequently, further methods of data analysis were followed. Plonsky (2015) recommends using bootstrapping —a type of robust statistic— where values are randomly resampled from the original sample to form new samples (Larson-Hall, 2015, p. 474). This resampling procedure is useful to overcome the lack of statistical power and to avoid Type II errors[3] resulting from analyses based on small samples (Plonsky, 2015, p. 47). Since CIs are based on standard errors, which are really only useful if the data are distributed normally or when the sample size is greater than 30 (Gries, 2013, p. 130), for cases where the samples are not normally distributed or are too small, such as here, Gries (2013) recommends the use of bootstrapping to obtain CIs.

[3] The error of not rejecting H0 when it is false (Howell, 2014, p. 169)

Using bootstrapping (10,000 replications), pairwise comparisons for the five groups were conducted. Table 3.2 reports the mean difference, the 95% bias-corrected and accelerated (BCa) CI [4], Cohen's *d*[5] effect size, and BCa CI for Hedges' *g*[6].

Table 3.2 *Group Comparison Using Bootstrapping (10,000 replications)*

Group comparison	Mean difference	95% BCa CI	Cohen's d	BCa CI for Hedges' g
Beginners I-Beginners II	1.00	[-0.14, 1.90]	- 0.75	[-0.08, 1.53]
Beginners I-Intermediate I	1.03	**[0.18, 1.71]**	-0.79	[0.13, 1.49]
Beginners I-Intermediate II	-0.39	[-1.40, 0.53]	0.30	[-0.93, 0.55]
Beginners I- Advanced	0.45	[-0.34, 1.15]	-0.42	[-0.28, 0.99]
Beginners II-Intermediate I	0.03	[-0.84, 0.98]	-0.02	[-0.79, 0.87]
Beginners II-Intermediate II	-1.40	**[-2.38, -0.18]**	1.17	[-2.37, 0.06]
Beginners II-Advanced	-0.54	[-1.33, 0.50]	0.58	[-1.83, 0.51]
Intermediate I-Intermediate II	-1.43	**[-2.26, -0.49]**	1.24	[-2.11, -0.29]
Intermediate I-Advanced	-0.57	[-1.15, 0.10]	0.65	[-1.16, 0.14]
Intermediate II-Advanced	0.85	[-0.05, 1.68]	-0.94	[-0.23, 2.05]

Note: Statistically significant CIs are boldfaced.

The statistically significant comparisons were Beginners I and Intermediate I, Beginners II and Intermediate II, and Intermediate I and Intermediate II. All the CIs for the mean differences between groups in the bootstrapped samples are large, which indicates that we cannot place too much confidence in them.

[4] The bias-corrected and accelerated (BCa) bootstrap adjusts for the median not being equal to the mean (Larson-Hall, 2016, p. 474).

[5] Cohen's *d* was calculated with the formula Cohen's $d = M_1 - M_2 / \sigma_{pooled}$ where $\sigma_{pooled} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$, using the online calculator provided by the University of Colorado Spring available at: https://www.uccs.edu/lbecker/

[6] Hedges' *g* is a measure of effect size similar to Cohen's *d* but differs from it in the denominator. Hedges' $g = M_1 - M_2 / SD_{pooled}$

Following Plonsky and Oswald's (2014) guidelines, where *d*=0.4 is a small effect size, *d*=0.7 is a medium effect, and *d*=1.0 is a large effect, there was a medium effect size in the mean difference between the Beginners I and Intermediate I group, with Cohen's *d* = -0.79. Based on the 95% CI obtained from the bootstrapped (10,000) analysis (CI [0.18, 1.71]), we are confident that there was a real increase in the density of discourse markers used that ranged from 18% to 170%; even though the range is wide, we are confident that the increase was of at least 18%.

The two strongest effects were observed in the comparisons with the Intermediate II group. The comparison between the Beginners II and Intermediate II groups showed a Cohen's *d*=1.17, indicating that the difference in means decreased from Beginners II to Intermediate II. There was also a strong effect size for the Intermediate I and Intermediate II groups (Cohen's *d*=1.24). The analyses showed that the Intermediate II group had a significantly lower density of discourse markers than the Beginners I and Intermediate II groups. A possible explanation for this decrease may be the task itself, in the form of a letter that requires personal information from the petitioner and has specific sections that must be followed; thus, the difference in density in the use of discourse markers may be due in part to a methodological flaw, although the average length of the compositions did increase from 161.12 words (Beginners II) to 264.37 words (Intermediate II), providing enough opportunities to use discourse markers.

It was expected to find significant differences between the groups as the levels progressed; however, the bootstrapped comparison didn't find any significant difference at the .05 or .1 levels between the Advanced group and any other group. There was a medium (but non-significant) effect size between the Intermediate II and Advanced group (Cohen's *d*=-0.94); however, considering the unexpectedly low density of discourse markers in the Intermediate II group, the real effect is likely smaller. The comparisons also showed a small effect size between

the Beginners I and Advanced group (Cohen's d=-0.42). Even though the comparison was not statistically significant, the effect cannot be neglected.

In sum, the comparisons of the density of discourse markers used by the five groups revealed the following findings. There was a statistically significant increase in the density of discourse markers used between the Beginners I and Intermediate I groups. The Intermediate II group used a statistically significant lower density of discourse markers than the Beginners II and Intermediate I groups. There was a small effect size in the comparison between Beginners I and Advanced group and a medium effect size between the Intermediate II and Advanced groups; however, neither was significant. Based on the data gathered, the hypothesis that higher levels would use more discourse markers than lower levels is supported for certain levels but not others, although a general (non-significant) increase is observed for the other comparisons.

3.2 How Varied is the Use of Discourse Markers?

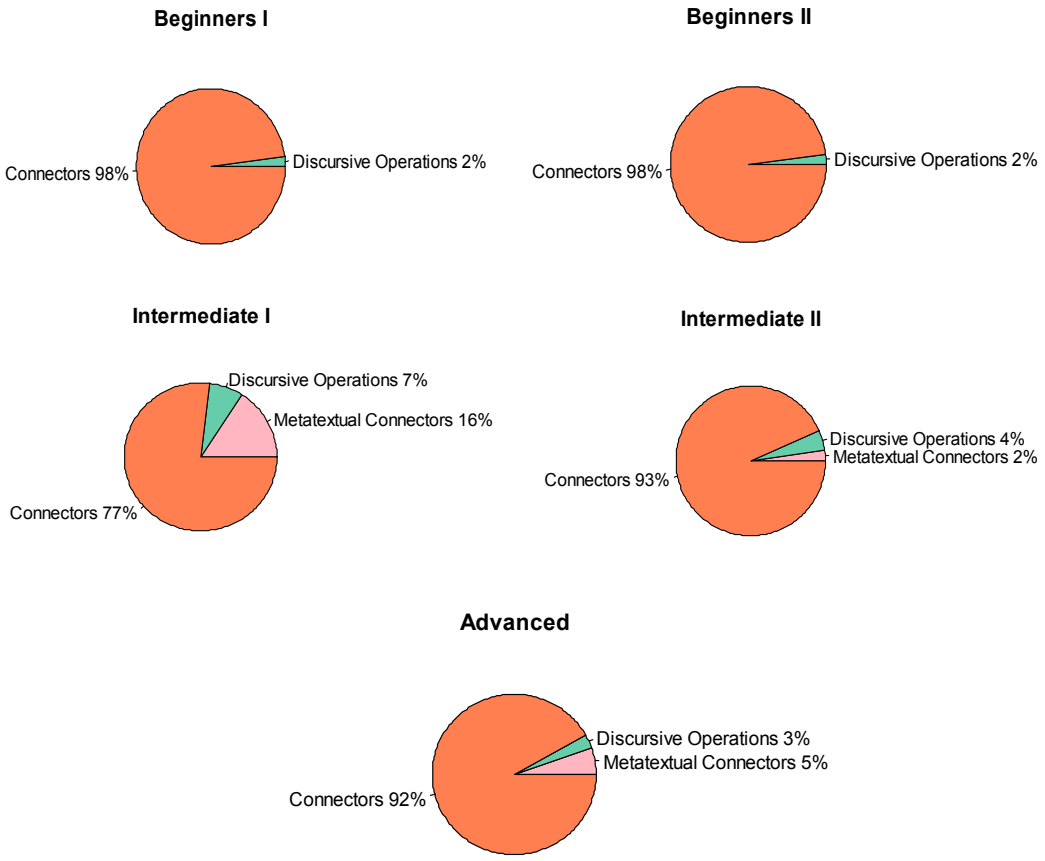
In this section, I discuss the variety of discourse markers used by the participants. I only consider accurate uses because it was difficult to assign a function to the incorrect forms. First, the percentage of use of each category of discourse markers is provided to show the distribution of use of the different types of discourse markers along the five levels. Second, the frequency of use of each individual discourse marker is presented for each level to investigate any possible use patterns. Third, the percentage of use of each discourse marker within its respective category is provided to show the pattern of discourse marker use at the category level. Lastly, the results of the statistical analysis of the variety of discourse markers used are presented to examine differences in the use of discourse markers.

3.2.1 Frequency of discourse markers by category

This section provides a general picture of the use pattern of discourse markers. When the variety of discourse markers was considered in terms of the different categories, it was observed that not all the levels attempted to use the three types of discourse markers. Both Beginner levels predominantly used connectors (98% of all occurrences); the discursive operators represented the remaining 2%. The Intermediate and Advanced groups used all three categories of discourse markers. It is noticeable that the Intermediate I group used more metatextual connectors (16%) compared with the Intermediate II (2%) and Advanced groups (5%). Figure 3.3 present the pie charts with the corresponding percentages of discourse markers per category.

It was expected to find a predominance of connectors over discursive operators and metatextual connectors because the category connectors encompasses coordinating conjunctions that join independent clauses (Intaraprawat & Steffensen, 1995); therefore, they will have a higher frequency of occurrence. The other two categories have a clause-level or higher scope (Intaraprawat & Steffensen, 1995); as a result, the opportunity for use is lower for discursive operators and metatextual connectors.

Figure 3.3 *Percentages of Discourse Markers per Category*



3.2.2 Frequency of individual discourse markers

The four most frequently used discourse markers were connectors: *y* "and" (142 tokens), *porque* "because" (130 tokens), *también* "also" (59 tokens), and *pero* "but" (41 tokens). Together they account for the 76.54% of the total occurrences. The metatextual connector with the highest number of occurrences (18 tokens) was *antes* "before", which represents 3.70% of the total. The discursive operators with the highest number of instances (11 tokens each) were *si* "if," *en mi opinión* "in my opinion," and *por ejemplo* "for example", at 2.26% each. The rest of the discourse markers had very few occurrences and the percentages ranged from 0.21% (1 token) to 1.65% (8 tokens). Table 3.3 shows the occurrences of each discourse marker and as a percentage of the total count of discourse markers.

Table 3.3 *Overall Frequency of Individual Discourse Markers*

Type	Discourse marker	Tokens	%
C	y and	142	29.22
C	porque because	130	26.75
C	también also	59	12.14
C	pero but	41	8.44
MT	antes before	18	3.70
DO	por ejemplo for example	11	2.26
DO	si if	11	2.26
DO	en mi opinión in my opinion	10	2.06
MT	ahora now	8	1.65
C	además in addition	8	1.65
C	sin embargo however	8	1.65
MT	después after	7	1.44
C	entonces then	7	1.44
C	por eso therefore	5	1.03
C	así que so that	5	1.03
C	entonces back then, at that time	3	0.62
C	a pesar de although	2	0.41
C	por lo tanto therefore	2	0.41
MT	luego later	1	0.21
MT	por otro lado on the other hand	1	0.21
MT	en fin finally	1	0.21
DO	de hecho in fact	1	0.21
C	por el contario on the contrary	1	0.21

Table 3.3 (continued) *Overall Frequency of Individual Discourse Markers*

Type	Discourse marker	Tokens	%
C	no obstante nevertheless	1	0.21
C	pues then	1	0.21
C	por consiguiente in consequence	1	0.21
C	luego then	1	0.21
Total # of discourse markers		486	

C=Connectors; MT=Metatextual Connectors; DO=Discursive Operators

By looking at the frequency of discourse markers by level, it was observed that Beginners I used predominantly *y* "and" *porque,* "because," and *también* "also," which together represent 88.73% of all the tokens used by the participants in this level. Table 3.4 presents the frequency of discourse markers observed in Beginners I and as a percentage of the total count of discourse markers for that level.

Table 3.4 *Frequency of Discourse Markers: Beginners I*

Type	Discourse marker	Tokens	Percentage of total # of tokens
C	y and	52	34.43%
C	porque because	45	29.80%
C	también also	37	24.50%
C	pero but	10	6.62%
DO	por ejemplo for example	3	1.98%
C	además in addition	2	1.32%
C	así que so that	1	0.66%
C	si if	1	0.66%
	Total # of tokens	151	

C=Connectors; DO= Discursive Operators

Beginners II used mostly *porque* "because" followed by *pero* "but" and *también* "also." The occurrences of these three discourse markers represent 83.34% of all the tokens used by the participants at this level (Table 3.5).

Table 3.5 *Frequency of Discourse Markers: Beginners II*

Type	Discourse marker	Tokens	Percentage of total # of tokens
C	porque because	27	56.25 %
C	pero but	8	16.67 %
C	tambien also	5	10.42 %
C	y and	4	8.33 %
C	sin embargo however	2	4.17 %
C	entonces then	1	2.08 %
DO	por ejemplo for example	1	2.08 %
	Total # of tokens	48	

C=Connectors; MT=Metatextual Connectors; DO= Discursive Operators

The Intermediate I level used predominantly the discourse markers *y* "and" and *porque* "because" which represent 54.89% of the tokens (Table 3.6). This group began to use metatextual connectors; the third most frequently used discourse marker *antes* "before" is part of this category.

Table 3.6 *Frequency of Discourse Markers: Intermediate I*

Type	Discourse Marker	Tokens	Percentage of total # of tokens
C	y and	72	35.29 %
C	porque for example	40	19.60 %
MT	antes before	18	8.82 %
C	también also	13	6.37 %
C	pero but	12	5.88 %
DO	en mi opinión in my opinion	9	4.41 %
MT	ahora now	8	3.92 %
MT	después after	6	2.94 %
DO	por ejemplo for example	5	2.45 %

Table 3.6 (Continues) *Frequency of Discourse Markers: Intermediate I*

Type	Discourse Marker	Tokens	Percentage of total # of tokens
C	sin embargo however	5	2.45 %
C	además in addition	4	1.96 %
C	entonces then	4	1.96 %
C	por eso therefore	2	0.98 %
C	entonces back then	2	0.98 %
MT	luego later	1	0.49 %
DO	de hecho in fact	1	0.49 %
C	no obstante nevertheless	1	0.49 %
C	por lo tanto therefore	1	0.49 %
	Total # of tokens	204	

C=Connectors; MT=Metatextual Connectors; DO= Discursive Operators

The Intermediate II group used mostly *y* "and" and *porque* "because" followed by *pero* "but"; these three discourse markers represent 58.67% of all the tokens used by this level (Table 3.7).

Table 3.7 *Frequency of Discourse Markers: Intermediate II*

Type	Discourse marker	Tokens	Percentage of total # of tokens
C	y and	12	26.08 %
C	porque because	10	21.73 %
C	pero but	5	10.86 %
C	así que so that	4	8.69 %
C	también also	3	6.52 %
DO	por ejemplo for example	2	4.34 %
C	además in addition	2	4.34 %
C	por eso therefore	2	4.34 %
C	si if	2	4.34 %
MT	después after	1	2.17 %
C	entonces then	1	2.17 %
C	por lo tanto therefore	1	2.17 %
C	luego later	1	2.17 %
	Total # of tokens	46	

C=Connectors; MT=Metatextual Connectors; DO= Discursive Operators

The Advanced group used predominantly the discourse markers *porque* "because," *si* "if," and *pero* "but." These three discourse markers represent 59.45% of all the tokens used by this level (Table 3.8).

Table 3.8 *Frequency of Discourse Markers: Advanced*

Type	Discourse marker	Tokens	Percentage of total # of tokens
C	porque because	8	21.62 %
C	si if	8	21.62 %
C	pero but	6	16.21 %
C	y and	2	5.40 %
C	a pesar de although	2	5.40 %
MT	por otro lado therefore	1	2.70 %
MT	en fin finally	1	2.70 %
DO	en mi opinión in my opinion	1	2.70 %
C	también also	1	2.70 %
C	sin embargo however	1	2.70 %
C	por el contario on the contrary	1	2.70 %
C	entonces then	1	2.70 %
C	por eso therefore	1	2.70 %
C	pues then	1	2.70 %
C	por consiguiente in consequence	1	2.70 %
C	entonces back then	1	2.70 %
	Total # of tokens	37	

C=Connectors; MT=Metatextual Connectors; DO= Discursive Operators

3.2.3 Frequency of use of connectors

As mentioned in the methodology section, connectors establish logical-semantic relationships among textual segments and provide cohesion to the text. Connectors was the category of discourse markers with the highest frequency of use, at 88.07% of the total number of tokens. As can be observed in table 3.9, the connectors *y* "and" and *porque* "because" were the most frequently used—each had more than twice as many uses as the connector in third

place, *pero* “but"—and they account for more than two-thirds (77.34%) of all the connectors used. Table 3.9 presents the list of connectors with the number of total tokens in the data and as the percentage of the category.

Table 3.9 *Frequency of Use of Connectors*

Discourse marker	Tokens	Percentage of the category
y and	142	33.18%
porque because	130	30.37%
también also	59	13.79%
pero but	41	9.58%
si if	11	2.57%
además in addition	8	1.87%
sin embargo however	8	1.87%
entonces then	7	1.64%
por eso therefore	5	1.17%
así que so that	5	1.17%
entonces back then	3	0.70%
a pesar de although	2	0.47%
por lo tanto therefore	2	0.47%
por el contario on the contrary	1	0.23%
no obstante nevertheless	1	0.23%
pues then	1	0.23%
por consiguiente in consequence	1	0.23%
luego then	1	0.23%
Total # of tokens	428	

In the data analyzed, the participants used connectors that belong to six subcategories: Additives, Causatives, Contrastives, Consecutives, Temporal, and Conditionals. In the following sections, I examine the connectors following these subcategories. It was expected to find a broader variety of discourse markers as the proficiency of the groups increased. For several of the subcategories, as described below, this was indeed the case; however, this wasn’t consistent among all the subcategories.

Additive connectors

This subcategory includes two of the most frequently used discourse markers in the whole corpus: *y* and *también*. The connector *y* "and" (142 tokens) has the highest frequency, representing 29.22% of the total discourse markers used by the participants. The connector *también* "also" (59 tokens) is the third most frequently used discourse marker in all the categories, with a frequency of 12.14%. The frequency of the connector *además* "in addition" was low (8 tokens), 1.65% of all the tokens in the data. Table 3.10 presents the raw count for each additive connector and the percentage of that additive connector out of all the additive connectors used by that level.

Table 3.10 *Frequency of Use of Additive Connectors*

Group	y and		también also		además in addition	
	Tokens	%	Tokens	%	Tokens	%
Beg. I (n=21)	52	57.14 %	37	40.66 %	2	2.20 %
Beg. II (n=8)	4	44.44 %	5	55.55%	-	-
Int. I (n=23)	72	80.90 %	13	14.61 %	4	4.49 %
Int. II (n=8)	12	70.59 %	3	17.65 %	2	11.76 %
Adv. (n=4)	2	66.67 %	1	33.33 %	-	-
Total	142	67.94 %	59	28.22 %	8	3.82 %

Within the subcategory of additives, the connector *y* "and" was used by 78.12% of all the participants. The Beginners I group used a lower percentage of this connector (57.14%) compared with the other four groups; this can be explained by the fact that the Beginners I group used all three connectors in this subcategory. The Beginners II group's use of the connector *y*

"and" represents 44.44% of the additive connectors for the level. For the Intermediate I, this connector represented 80.90 % of the total of additive connectors produced by the level.

The additive connector *también* "also" was used by 48.43% of the participants and by all the groups except Beginners II. The Beginners I and the Advanced groups showed a higher percentage of use of this discourse marker.

The connector *además* "in addition" was taught explicitly in the second unit of the Beginners II class. Consequently, we would expect Beginners II and all subsequent levels to use it; however, the Beginners II and the Advanced group didn't use *además* at all. In fact, only 6 out of 64 participants of the study (9.3% of all the participants) used this connector, from which we can infer that there is a low awareness of its existence.

Causative connectors

The connector *porque* "because" was the second most used discourse marker (130 tokens) at 26.75% of the total tokens in the study and at 30.37% of all the connectors. The number of tokens per level was as follows: Beginners I (45 tokens), Beginners II (27 tokens), Intermediate I (40 tokens), Intermediate II (10 tokens), Advanced (8 tokens).

Given the argumentative nature of the tasks, it was expected that participants would use this connector frequently. All the groups used this discourse marker, and 78.12% of all the participants used it at least once in their compositions. It is important to notice that it was the only causative connector used.

Contrastive connectors

From the data collected, there were five different discourse markers that belong to the contrastive connectors subcategory: *pero* “but,” s*in embargo* “however,” *por el contario*” on the contrary,” *a pesar de* “although,” *no obstante* “nevertheless.” *Pero* “but” was the contrastive connector that was used with the highest frequency (41 tokens); this connector was the fourth most used discourse marker in the data, and it represents 9.58% of all the connectors, used by 40.62% (26) of the participants. Table 3.11 presents the frequencies of use for the contrastive connectors per level.

The diversity of contrastive connectors increased at the Intermediate I level with one occurrence of *no obstante* “nevertheless”; therefore, it was important to observe if the token was produced by a participant who was also using *pero* “but” and *sin embargo* “however” to rule out that the increase of variety was being reported due to a single participant who had a high awareness of contrastive connectors. The token was produced by a participant who didn’t use *pero*, which suggests a diversification in the awareness of discourse markers.

The connector *sin embargo* "however" was used by the Beginners II, Intermediate I and Advanced groups, producing a total of 8 tokens that represent 15.09% of all the contrastive connectors. The connectors *a pesar de* "although" and *por el contrario* "on the contrary" were only used by the Advanced group, with one occurrence each. The connector *no obstante* "nevertheless" had only one token in the data, and it was used by the Intermediate I group.

The variety of contrastive discourse markers increased parallel to the proficiency of the groups: Beginners I group only used one connector of this subtype, Beginners II used two different connectors, Intermediate I used three different connectors, while the Advanced group

produced the greatest variety with four different discourse markers. It must be noted that the Intermediate II group only used *pero* "but", reflecting the unexpectedly the lower level of performance seen earlier.

Table 3.11 *Frequency of Use of Contrastive Connectors*

Group	pero but		sin embargo however		por el contario on the contrary		a pesar de although		no obstante nevertheless	
	Tokens	%	Tokens	%	Tokens	%	Tokens	%	Tokens	%
Beg. I (n=21)	10	100.00	-	-	-	-	-	-	-	-
Beg. II (n=8)	8	80.00	2	20.00	-	-	-	-	-	-
Int. I (n=23)	12	66.67	5	27.78	-	-	-	-	1	5.56
Int. II (n=8)	5	100.00	-		-	-	-	-	-	-
Adv. (n=4)	6	60.00	1	10.00	1	10.00	2	20.00	-	-
Total	41	77.35	8	15.09	1	1.88	2	3.77	1	1.88

Consecutive connectors

Among the connectors, consecutive connectors was the subcategory with the greatest variety (six types). Of all the participants, 20.31% (13) used at least one consecutive connector. The Beginners I group used only one consecutive connector, *así que* "so that," with one occurrence. Beginners II used only the connector *entonces* "then," with one token as well. Intermediate I group used three types: *entonces* "then," *por eso* "therefore," and *por lo tanto* "therefore." Intermediate II used four types: *entonces* "then," *por eso* "therefore," *así que* "so that," and *por lo tanto* "therefore." The Advanced group used *entonces* "then," *por eso* "therefore," *pues* "then," and *por consiguiente* "in consequence." The variety of consecutive connectors used increased along with proficiency (see Table 3.12).

Table 3.12 *Frequency of Use of Consecutive Connectors*

Group	entonces then		por eso therefore		así que so that		por lo tanto therefore		por consiguiente in consequence		pues then	
	Tk	%	Tk	%	Tk	%	Tk	%	Tk	%	Tk	%
Beg. I (n=21)	-	-	-	-	1	100	-	-	-	-	-	-
Beg. II (n=8)	1	100	-	-	-	-	-	-	-	-	-	-
Int. I (n=23)	4	57.14	2	28.57	-	-	1	14.28	-	-	-	-
Int. II (n=8)	1	12.50	2	25.00	4	50.00	1	12.50	-	-	-	-
Adv. (n=4)	1	25.00	1	25.00	-	-	-	-	1	25.00	1	25.00
Total	7	33.33	5	23.80	5	23.80	2	9.52	1	4.76	1	4.76

*Tk=Tokens

Temporal connectors

Table 3.13 presents the occurrences of each connector and as a percentage of the total count of temporal connectors per level. The connector *entonces* "back then" was only used by the Intermediate I and the Advanced groups. The connector *luego* "later" was used once by a participant in the Intermediate II group. Each token was used by a different participant, which means that only 4 out of 64 participants (6.25%) used temporal connectors. No tokens in this subcategory were identified in the samples collected from Beginners I, Beginners II, and Intermediate I groups.

Table 3.13 *Frequency of Use of Temporal Connectors*

Group	*n*	entonces back then		luego then	
		Tokens	%	Tokens	%
Beg. I	21	-	-	-	-
Beg. II	8	-	-	-	-
Int. I	23	2	100	-	-
Int. II	8	-		1	100
Adv.	4	1	100	-	-
Total		3	75	1	25

Conditional connectors

Si "if" was the only conditional connector identified in the data collected. Overall, it was used by 7 participants (10.9%). The Advanced group showed the highest frequency of use with a total of 8 occurrences; each of the participants in the Advanced group used it at least once. The Intermediate II group registered only two tokens, with 2 out of 8 participants using this connector. In the case of the Beginners I group, there was only one occurrence among the 21 participants. Beginners II and Intermediate I groups didn't use it at all.

In this section, I have presented the frequency of use of connectors. The data showed that the most frequently used connectors were *y* "and," *porque* "because," *también* "also," and *pero* "but." It also revealed that the variety of certain subcategories of connectors used increased with the level of proficiency, except for the group Intermediate II that behaved in an unusual way, producing a lower use of discourse markers than the other four groups. Among all the types of connectors found in the data, consecutive connectors was the subcategory with the greatest variety (six different types).

3.2.4 Frequency of use of discursive operators

In the data collected, three discursive operators were identified. Each of them belongs to a different subcategory. The discourse marker *por ejemplo* "for example" presents an example. The discourse marker *en mi opinión* "in my opinion" introduces point of view. The discourse marker *de hecho* "in fact" belongs to the category of discursive operators that is used to expresses certainty (Calsamiglia & Tusón, 2001, p. 247). Table 3.14 presents the raw count of discursive operators and as a percentage of the category.

Table 3.14 *Frequency of Use of Discursive Operators*

Discourse marker	Tokens	Percentage of the category
por ejemplo for example	11	50.00 %
en mi opinión in my opinion	10	45.45 %
de hecho in fact	1	4.55 %
Total # of tokens	22	

Discursive operators represent 4.5% of all the discourse markers in the data. In total, 17 participants (26.5%) used at least one discourse marker from this category. When analyzing the numbers of participants that used one, two or all three discursive operators, it was observed that only two participants used both *en mi opinión* "in my opinion" and *por ejemplo* "for example," and only one used all three discourse markers; therefore, the tokens are not being produced by the same participants. On the other hand, these three discursive operators fulfill very different functions, so if the participants didn't use them, it could simply be because they didn't need it or because there were other forms to express the idea without a discourse marker. Table 3.15 presents the raw count of discursive operators per group. No percentages are given because each discursive operator belonged to a different subcategory.

The discourse marker *en mi opinión* "in my opinion" was used by 15.62% of the participants, nine in the Intermediate I and one in the Advanced group. Given the argumentative nature of the task, this was a lower-than-expected number of tokens; however, in Spanish there are several constructions to express opinion, such as *yo pienso que "*I think that," *yo creo que* "I believe that," *me parece que* "it seems to me that," *yo opino que* "in my opinion" and so on, that are not considered discourse markers.

The discourse marker *por ejemplo* "for example" was used by 10 participants as well (15.62%). The Advanced group was the only group where no occurrences of this exemplifier were observed. Note that the low number of tokens doesn't mean the participants didn't know the discourse marker; the absence of the form could be because it wasn't needed in the text or because they used an alternative form.

Table 3.15 *Frequency of Use of Discursive Operators*

	Point of view	To present an example	To express certainty
	en mi opinión in my opinion	por ejemplo for example	de hecho in fact
Group	Tokens	Tokens	Tokens
Beg. I (n=21)	-	3	-
Beg. II (n=8)	-	1	-
Int. I (n=23)	9	5	1
Int. II (n=8)	-	2	-
Adv. (n=4)	1	-	-
Total	10	11	1

3.2.5 Frequency of use of metatextual connectors

There were six discourse markers used that belonged to this category and which are grouped into three subcategories. Table 3.16 presents the raw count of metatextual connectors and as a percentage of all the metatextual connectors used at that level.

The Intermediate I group used the most metatextual connectors, although all from the same category; the two Beginners groups didn't use any metatextual connector. The Advanced group was the only one that used organizers and conclusion connectors.

Table 3.16 *Frequency of Use of Metatextual Connectors*

Group	Temporal								Organizers		Conclusion	
	antes before		ahora now		después after		luego later		por otro lado on the other hand		en fin finally	
	Tk	%	Tk	%	Tk	%	Tk	%	Tk	%	Tk	%
Beg. I (n=21)	-	-	-	-	-	-	-	-	-	-	-	-
Beg. II (n=8)	-	-	-	-	-	-	-	-	-	-	-	-
Int. I (n=23)	18	54.55	8	24.24	6	18.18	1	3.03	-	-	-	-
Int. II (n=8)	-	-	-	-	1	100	-	-	-	-	-	-
Adv. (n=4)	-	-	-	-	-	-	-	-	1	50.00	1	50.00
Total	18	50.00	8	22.22	7	19.44	1	2.78	1	2.78	1	2.78

Tk=Tokens

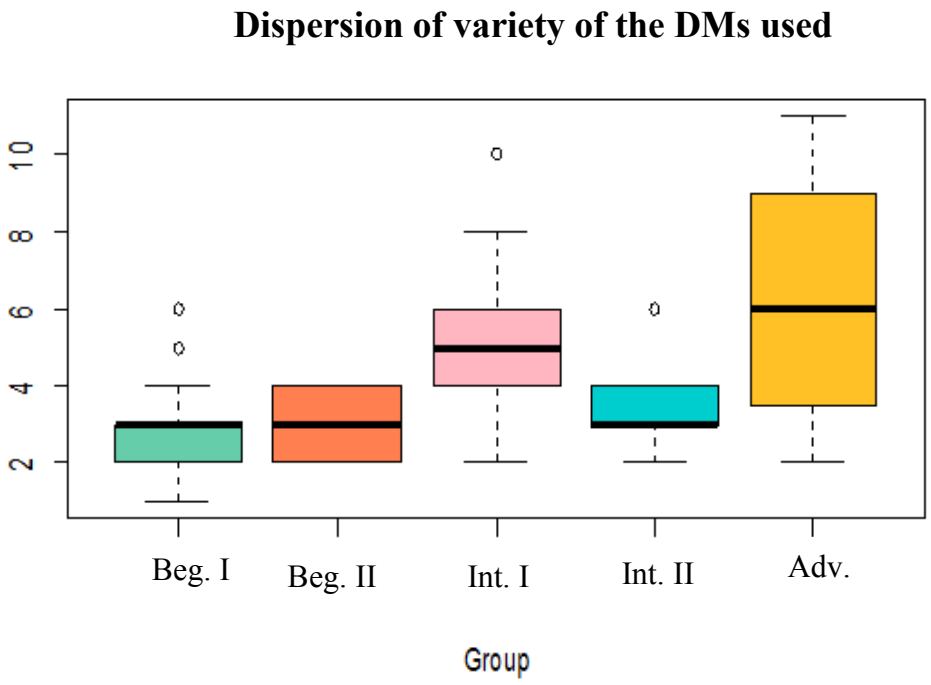
3.2.6 Statistical comparison of variety of discourse markers

It was hypothesized that the variety of discourse markers would increase parallel to the proficiency of the groups. Variety was calculated considering the types of discourse markers used regardless of the category they belong to. Figure 3.4 shows the dispersion of the variety of discourse markers used by each group. The descriptive statistics for the groups are presented in Table 3.17.

Table 3.17 *Descriptive Statistics for Variety of Discourse Markers Used*

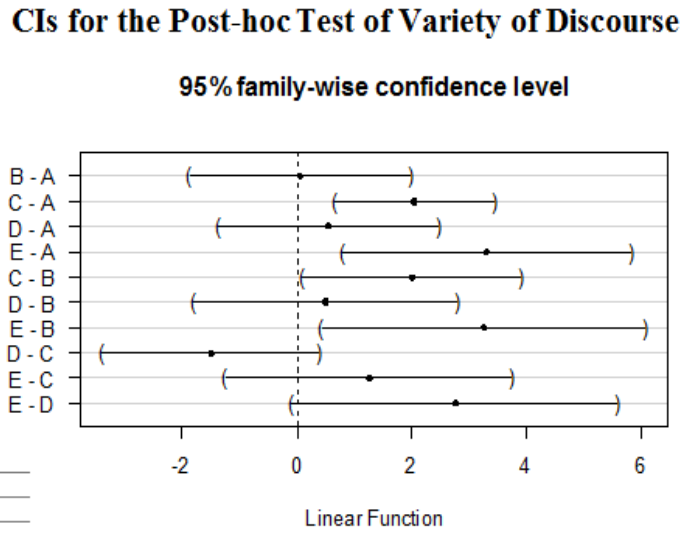
Group	*n*	Mean # of types	SD
Beginners I	21	2.9	1.16
Beginners II	8	3.0	0.92
Intermediate I	23	5.0	1.85
Intermediate II	8	3.5	1.19
Advanced	4	6.2	3.77
Total	64	4.0	1.95

Figure 3.4 *Dispersion of the Variety of Discourse Markers Used by Group.*



In order to determine if there were statistically significant differences in the variety of discourse markers used, an omnibus one-way ANOVA was performed. A Tukey's post-hoc test indicated that there were significant differences in the comparisons between the following three pairs: a) Beginners I with Intermediate I and with Advanced, b) Beginners II with Intermediate I and with Advanced, and c) Intermediate II with Advanced. The CIs for the post-hoc test are visually presented in Figure 3.5.

Figure 3.5 *CIs for the Post-hoc Test for Variety of Discourse Markers*



IDENTIFIER	GROUP
A	Beginners I
B	Beginners II
C	Intermediate I
D	Intermediate II
E	Advanced

Using bootstrapping (10,000 replications), pairwise comparisons for the five groups were conducted. Table 3.18 reports the mean difference, the 95% bias-corrected and accelerated (BCa) CI, Cohen's *d* effect size, and BCa CI for Hedges' *g*.

There were two strong effect sizes observed in the comparison of the Beginners I group with Intermediate I (*d*=1.32) and Advanced (*d*=1.18). The comparisons between the Beginners II group and Intermediate I and between Beginners II and Advanced also showed strong effect sizes (*d*=1.36 and *d*=1.18, respectively). There was a medium effect size for the comparison between Intermediate I and Intermediate II, with a Cohen's *d*=-0.96 due to the unexpectedly low use of discourse markers in the Intermediate II group.

Table 3.18 *Comparison of Variety of Discourse Markers Using Bootstrapping (10,000 replications)*

Group comparison	Mean difference	95% BCa CI	Cohen's d	BCa CI for Hedges' g
Beginners I-Beginners II	0.05	[-0.79, 0.77]	0.04	[-0.69, 0.79]
Beginners I-Intermediate I	2.05	**[1.17, 2.95]**	1.32	[0.62, 1.87]
Beginners I-Intermediate II	0.55	[-0.31, 1.50]	0.47	[-0.39, 1.25]
Beginners I- Advanced	3.30	**[0.25, 6.73]**	1.18	[-0.27, 4.18]
Beginners II-Intermediate I	2.00	**[1.09, 2.99]**	1.36	[0.51, 1.71]
Beginners II-Intermediate II	0.50	[-0.37, 1.62]	0.47	[-0.72, 1.29]
Beginners II-Advanced	3.25	**[0.25, 6.87]**	1.18	[-1.15, 3.51]
Intermediate I-Intermediate II	-1.50	**[-2.52, -0.43]**	-0.96	[-1.47, -0.06]
Intermediate I-Advanced	1.25	[-1.89, 4.65]	0.42	[-1.20, 2.36]
Intermediate II-Advanced	2.75	[-0.37, 6.25]	0.98	[-0.70, 3.10]

These results show that there is a statistically significant difference in the variety of discourse markers that are used between the two Beginners groups and the other levels (except Intermediate II). Based on the 95% CI obtained from the bootstrapped (10,000 replications) analysis (CI [0.25, 6.73]), we are confident that there was a real increase in the variety of discourse markers used between Beginner I and Advanced that ranged from 25% to 673%. Although the range of the increase is wide, we are confident that the increase in variety was of at least 25%.

In this section, I have presented the distribution of use of the three categories of discourse markers along the five levels. In addition, I have discussed the frequency of use of each

discourse marker, and showed how the variety of discourse markers used increased with the level of proficiency, except for the Intermediate II group that behaved in an unusual pattern.

3.3 How Accurate is the Use of Discourse Markers?

The purpose of the third research question was to explore the appropriateness and accuracy of the discourse markers used by the participants. The appropriateness of each token was decided based on the acceptability of the sentence. In some cases it was difficult to determine the word the participant originally intended to use; nevertheless, it was possible to deduce the intended meaning based on context and the discourse markers that students are likely to know at a given level.

It was hypothesized that higher levels of proficiency would use discourse markers more accurately. In order to investigate how accurate the use of discourse markers was, the number of correct tokens over the total tokens produced by the participant was converted into a percentage to perform an analysis of variance. As a cursory glance at Table 3.19 reveals, the results go against the original hypothesis; therefore, the data were first analyzed to test for normal distribution and homogeneity of variance. After visual exploration (Figure 3.6), it was determined that the accuracy rate of the five groups was not normally distributed; a Levene test of variance produced a *p*-value of 0.095, indicating that there is no homogeneity of variance (Larson-Hall, 2015, p. 122). Consequently, a Kruskal-Wallis test (non-parametric analysis of variance) was performed. The Kruskal-Wallis test revealed a significant effect of Accuracy Rate on Group ($X2(4) = 16.66$, $p < 0.002$); it was followed by a post-hoc test using Dunn tests with Bonferroni correction. The pairwise comparisons showed two statistically significant differences:

a) between the Beginners I and the Advanced groups (p=0.001); and b) between the Intermediate I and the Advanced groups (p=0.028).

Figure 3.6 *Boxplot of Correct Use of Discourse Markers*

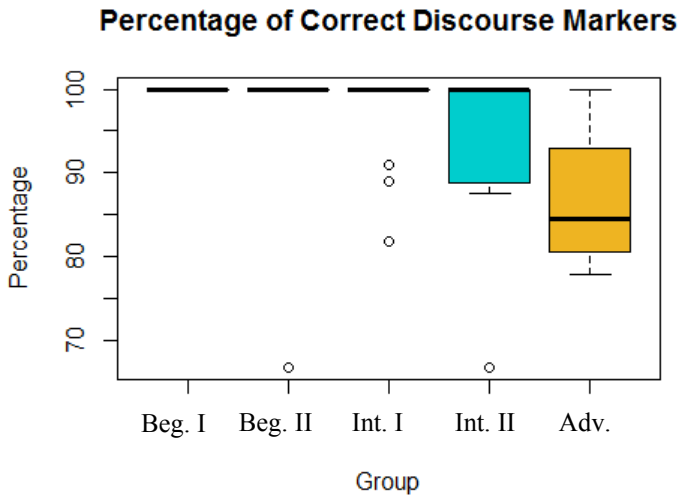


Table 3.19 *Mean Number of Tokens, Correct and Incorrect Tokens for all Discourse Markers*

Group	Total of tokens	Incorrect tokens	Tokens produced Mean	Tokens produced SD	Correct tokens Mean	Correct tokens SD	Accuracy rate
Beg. I (n=21)	151	0	7.19	4.11	7.19	4.11	100.00%
Beg. II (n=8)	49	1	6.12	1.96	6.00	2.20	97.95%
Int. I (n=23)	209	5	9.08	2.84	8.60	2.80	97.60%
Int. II (n=8)	49	3	6.12	2.47	5.75	2.37	93.87%
Adv. (n=4)	43	6	10.75	6.94	9.25	5.90	86.04%
Total	501	15	7.87	3.72	7.64	3.57	97.00%

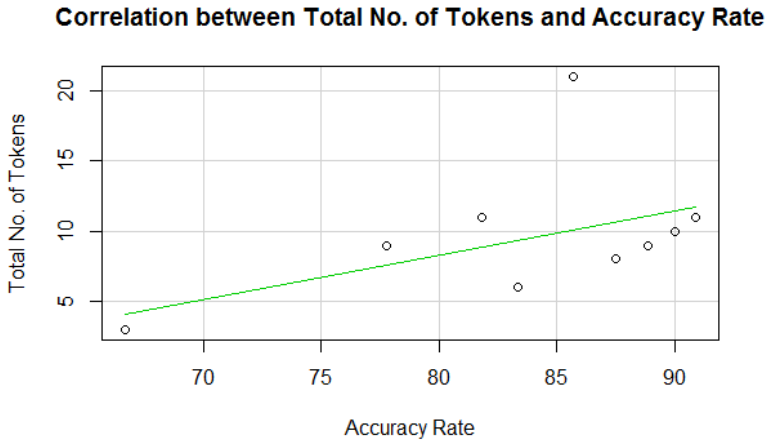
These results yield two new questions:

- Is the decrease in accuracy correlated with the number of tokens used, that is, the more discourse markers used, the more chances of error?
- Is the decrease in accuracy correlated with an increase in the variety of discourse markers used?

3.3.1 Is the decrease in accuracy correlated with the number of tokens used?

In order to determine if there is a correlation between the number of tokens used and the accuracy rates, a visual exploration of the dataset was done through a scatterplot. A 100% accuracy score nullifies any possible underlying correlation of these two factors. A second analysis considering only the cases where the accuracy rate was less than 100% was performed. The scatterplot revealed a possible correlation (Figure 3.7).

Figure 3.7 *Correlation Between the Number of Tokens Used and the Accuracy Rates*

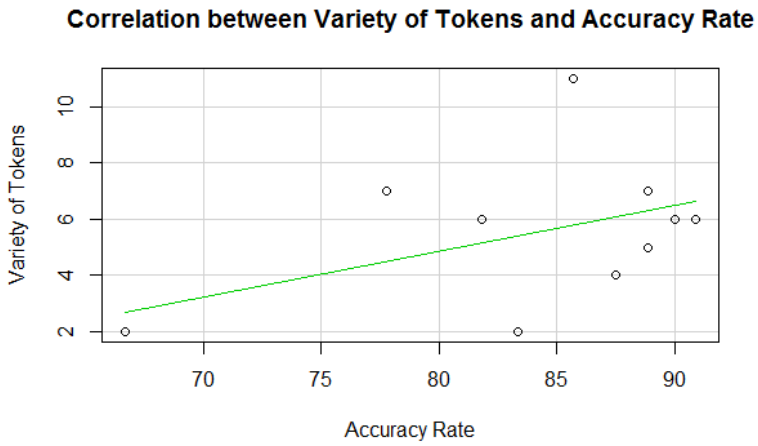


A Pearson correlation on the bootstrapped data of the cases where the accuracy rate was less than 100% (10,000 replications) found no correlation, with a medium effect size and a fairly wide CI (95% CI[-0.19, 0.98], N=11, r= 0.51, r^2=0.26). The fact that there was no correlation between the total number of tokens used and the accuracy score rejects the hypothesis that the decrease in accuracy is due to higher chances of making mistakes as a result of more discourse markers being used, that is, more opportunity for error.

3.3.2 Is the decrease in accuracy correlated with an increase in the variety of discourse markers used?

In order to answer the second question, a visual exploration was performed for the cases where the accuracy rate was lower than 100%. A linear regression of these data showed a possible positive correlation between the decrease in accuracy and the increase in variety of discourse markers used (Figure 3.8).

Figure 3. 8 *Correlation between Variety of Tokens and Accuracy Rate*



A Pearson correlation on the bootstrapped data of the cases where the accuracy rate was different from 100% (10,000 replications) found no correlation, with a medium effect size and a wide CI (95% CI [-0.31, 0.91], N=11, r= 0.40, r^2=0.16). These results show that a greater variety of discourse markers used is not linearly related to a decrease in accuracy.

The original hypothesis was that higher levels of proficiency would use discourse markers more accurately. The analysis showed it wrong; in fact, more proficient groups had a lower level of accuracy. After this unexpected result, two possible explanations were considered. The first was that a higher use of discourse markers would translate into more chances of error. The second was that a higher number of types of discourse markers (variety) would mean more chances of error. None of these hypotheses was confirmed; therefore, in order to gain a better understanding of the types of errors that affect the accurate use of discourse markers, a qualitative observation is required.

3.3.3 Error Analysis

The qualitative analysis of the inaccurate uses was done following Spillner's (1991) steps for error analysis. The definition of error was adopted from Ferris (2011): "Errors are morphological, syntactic and lexical forms that deviate from the rules of the target language, violating the expectations of a literate adult native speaker" (p. 3). It was difficult to differentiate between incorrect lexical choices and errors in form, as the types of errors were based on ESL conventions for error correction (Ferris, 2006); this issue was resolved by trying different error categories and exploring the tagging consistency between three instructors. The first attempts

showed low reliability scores which indicated the coding categories were not clear. The final coding version is described below.

There were three types of errors identified in the data I collected: a) Word Choice, which includes cases where the word that was used is inappropriate and interferes with the understanding of the sentence; b) Element Left Out, where some discourse markers are used in pairs and the absence of one of the elements affects the understandability of the clause, so sentences with missing or unnecessary words were included in this category; c) Word Form, where the intentionality of the clause is clear even though the form or the discourse marker is not accurate. Errors where one of a pair of discourse marker was omitted were excluded from this category because they fit better in the Element Left Out category.

The data were coded by the researcher, and then all the excerpts were randomly organized using *random.org*. These samples were then coded by a native-speaker university Spanish instructor with more than 10 years of teaching experience. The agreement percentage was originally 80%. The three tokens where the classification differed were then reviewed by the researcher and the supervisor and a coding decision was reached. Table 3.20 shows the total count of errors by type.

Table 3.20 *Total Count of Errors by Type*

		Error type		
Group	*n*	Word Choice	Element Left Out	Word Form
Beg. II	8	1	-	-
Int. I	23	2	3	-
Int. II	8	2	1	-
Adv.	4	3	-	3
Total	**64**	8	4	3

Examples of each error type are given below. The excerpts are reproduced verbatim, despite multiple errors in the texts which are not being addressed in the research. The translations are free versions that attempt to depict, whenever possible, the word selection and the sentence structure used by the participants.

Word Choice

This category registered the highest number of errors, eight, or 53.30% of all the errors coded, and this type of error was present in the Beginners II, Intermediate I, Intermediate II, and Advanced groups. It was the only error coded for the Beginners II group, shown in Excerpt 1.

Excerpt 1 [Beg. II/Part 6B]:

> [...] La grande habitación es la mejor, especialmente porque es no compartido, pero allí es una compartida sala por socialización. ***Tan***, tú no son viviendo solo en una nueva pías [...].
>
> "[…] The large room is the best, especially because it is not shared, but there is a shared room for socializing. So, you won't be living alone in a new country […]."

The comma after the first word indicates that the participant was attempting to connect the sentence with the previous one; in this context, an additive connector would be the best fit. If one searches Google Translate for the translation of "so," *tan* is the second option, although it refers to the "so" used in comparisons *(una habitación tan grande* "a room so big").

There were two errors in the Word Choice category in the Intermediate I group. Excerpt 2 exemplifies a use where it wasn't possible to recover with certainty the intended meaning of *no obstante* "however" or "nevertheless." In Spanish, *no obstante* has a counter-argumentative

(contrastive) nature; therefore, its use here is semantically inaccurate, violating the expectations of the reader.

Excerpt 2 [Int. I/Part. 7C]:

No tiene que cortar el pan en rodajas y las rodajas están siempre del mismo tamaño- normalmente más delgadas. ***No obstante***, hoy muchos lugares del mundo venden pan de varios tamaños.

"You don't have to cut the bread in slices and the slices are always the same size —usually thinner. **However,** nowadays several places in the world sell bread of different sizes."

In excerpt 3, the participant used the contrastive connector *sin embargo* "however" instead of *por otra parte* "on the other hand," a distributive metatextual connector, which would be one of few options to keep semantic coherence in the paragraph.

Excerpt 3[Int. I/Part. 13C]:

Pero él fue un idiota porque el agarraba un pedazo de metal en el aire mientras lo estaba en la tormenta de relámpago. **Sin embargo**, Yo deseo que hay este más un molemente hacia el crear de electricidad de recursos de renovables.

"But he was an idiot because he was holding a piece of metal in the air while he was in a lightning storm. **However**, I wish that there is a (unintelligible) to create electricity from renewable resources."

There were two errors of Word Choice coded for the Intermediate II group. In Excerpt 4, the participant used *así que* "so that" instead of the consecutive connector *de modo que* "in such a way that." As can be seen, the English form is the same for both conjunctive phrases (*conjunciones locutivas*); however, they are not interchangeable in Spanish. The fact that the form is used in the same sentence with no evidence of commas supports the interpretation of a literal translation of "so that." Another interpretation for this error would be choosing the adverb

así instead of the preposition *para,* the sentence could also perfectly accept *para que* "in order to."

Excerpt 4 [Int. II/Part 2D]:

Pido que prepares la guardaría antes de utilizarla ***así que*** el lugar es preparado cuando nos clientes vienen.

"I request that you get the daycare ready before using it **in order** to have the place ready when the clients come."

Another example of this type of direct translation is excerpt 5, where the participant, as above, made the same error using *así que* instead of *para que*.

Excerpt 5 [Int. II/Part.8D]:

Además, necesita que los trenes y los autobuses vengan más frecuentemente ***así que*** la gente no necesita esperar tanto.

"Also, it is necessary that the trains and the buses come more frequently **in order** for people not to wait so long."

There were three errors of Word Choice in the Advanced group. Excerpt 6 illustrates an attempt to use a conclusive metatextual connector. In this example, the participant used the form *en final*. There are a couple of options to interpret this entrance. One is an error in form using **en final* instead of *en fin* that can be translated as "anyway," a finalizing metatextual connector, which it is not coherent with the text; therefore, I disregarded it. The second option I can deduce is that the participant used *en final* having in mind the adverb "finally" *finalmente*; however, the adverb is frequently used in English to sequence elements in an argument and requires a previous position or statement of element that is not present in the composition. In Spanish, the adverb

finalmente at the beginning of a sentence works as a temporal discourse marker[7]. The third possibility is that the intended meaning was *en conclusión* "in conclusion," *para terminar* "to conclude," or *por último* "lastly"; as the last option is a more likely explanation, this error was classified as Word Choice.

Excerpt 6 [Adv./Part.2E]:

Por consiguiente, recibimos nuestras consecuencias de acuerdo con nuestras acciones entonces tenemos que tomar decisiones que no traigan remordimiento.
En final, hay personas de diferentes creencias sobre el destino […]

"Therefore, we get the consequences according to our actions then we have to make decisions that won't bring remorse.

Lastly, there are people with different beliefs about destiny [....]"

Excerpt 7 is a literal translation of the English form "on the other hand"; the correct form is *por otro lado* "on the other hand."

Excerpt 7 [Adv./Part. 1E]:

Esas personas no las dan mucha consideración. ***En el otro mano***, también otras decisiones más difíciles.

"Those people don't give them too much consideration. On the other hand, also more difficult decisions."

In excerpt 8, the participant used the discourse marker *mientras* which has several meanings. In the sense of conditional connector, this discourse marker would be translated to English as "as long as," and in the sense of temporal connector as "meanwhile." Considering the context of this sentence, none of those values is acceptable. The apparent intention of the sentence is to transition to a new argument that explains the writer's opinion in opposition to

[7] REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual.* <http://www.rae.es> [28/02/2018] Example: "rendimiento y la generosidad" entre sus miembros. Finalmente, Inés Alberdi resaltó como rasgo de madure**" [At the end, Inés Alberdi highlighted as a signs of ?]

somebody else's previously stated opinion. The literal translation and first entry in Google Translate of "while" is *mientras*. In order to achieve that transition, the most natural discourse markers would be the metatextual connectors: *por otra parte / por otro lado* "on the other hand," *por mi parte* "for my part," or *en cambio* "on the other hand" or "in contrast."

Excerpt 8 [Adv./Part. 1E]:

> Otras personas creen que gente se controla sus vidas totalmente y que las decisiones son un resultado del libre albedrío. ***Mientras*** yo creo que nosotros controlamos la dirección de nuestras vidas, también yo creo que el libre albedrío no es independiente de su destino.
>
> "Other people believe that people completely control their own lives and that decisions are the result of free will. **On the other hand,** I believe that we control the direction of our life, I also believe that free will is not independent of one's destiny."

Element Left Out

There were four occurrences coded for the category of Element Left Out, three of them in the Intermediate I group and one in the Advanced group. In excerpt 9, the writer used the consecutive connector *entonces* "then" instead of *en ese entonces* "back then." This interpretation is also supported by the use of *actualmente* "nowadays" in the following sentence, indicating the user was establishing a contrast in time. The missing words *en ese* interfere with the understanding of the sentence because the form *entonces* stands alone and has a different meaning. As a result, the coherence of the text suffers and the reader has to go back to the previous sentence to infer the meaning.

Excerpt 9 [Int. I/Part. 10C]:

> También, cuando las personas iban de vacaciones era difícil encontrar un teléfono. ***Entonces***, era difícil si querían llamar a sus parientes en un otro país y no podían usar el internet si se perdían.

Actualmente, es muy fácil cuando tienes que contactar a alguien. El móvil es como un ordenador pequeño.

"Also, when people used to go on vacation it was difficult to find a telephone. **Back then**, it was difficult if they wanted to call their relatives in another country and they couldn't use the Internet if they were lost. Nowadays, it is very easy when you have to reach someone. The cellphone is like a portable computer."

Another form of Element Left Out was when a set of interdependent discourse markers was truncated and one of the required elements was omitted in the composition. An example of this type of pattern in English would be "either...or." Excerpt 10 illustrates how the participant uses *no sólo* "not only" and misses the second element *sino* "but also" before the conjunction *que* "that."

Excerpt 10 [Int. I/Part. 6C]:

No sólo la bombilla crear una solución fácil y asequible para tener luz en la noche, [sino] que tenía beneficios económicos.

"**Not only** the light bulb created an easy and accessible solution to have light at night, [but also] it had economic benefits."

In excerpt 11, the participant used *más que nunca* "more than ever" at the beginning of the sentence, with an English-like funcion. In Spanish *más que nunca* is usually delimited by commas and preceded by a temporal adverb as *hoy* "today" or *ahora* "now," as in *hoy/ahora, más que nunca,[...]*.

Excerpt 11 [Int. I/ Part. 7C]:

La invención de pan rebanado cambió la industria panadera. ***Más que nunca*** las personas están comiendo pan y los panderos siempre están ocupados.

"The invention of the slice bread changed the bread industry. **More than ever** people are eating bread and bakers are always busy."

There was only one occurrence of Element Left Out identified in the Intermediate II corpus. Excerpt 12 illustrates how the participant used *el primero* "the first" for *primero* "first." This metatextual connector orders or sequences elements in the argument; therefore, it requires at least one more element in the listing, i.e. *segundo* "second." The participant listed a second policy the government should adopt; however, there is no explicit reference to the second element in the text. One would expect to find forms such as *segundo* "second" or *el segundo* "the second." Another possible interpretation of this error could be an agreement error where the intended meaning was *La primera [cosa]* "the first [thing]"; however, the use of comma separating *el primero* from the verb *es* "is" suggests that the participant was organizing the elements, making the first interpretation more likely.

Excerpt 12 [Int. II/ Part.4D]:

Hay muchas cosas que el gobierno necesite hacer a resolver la cuestión de 'calentamiento global'. ***El primero***, es a hágale una prioridad que la mayoría, si no toda la población sabe lo que pueden hacer a hacer una diferencia para el planeta.

"There are many things that the government must do to solve the problem of global warming. **First,** it is to make it a priority that the majority of people, if not all the population, know what they can do to make a difference for the planet."

Word Form

Only three errors were coded for the Word Form category, all of them in the Advanced group. In excerpt 13, the participant used *del* "prep. *de* + art. *el*" instead of the preposition *por* "for." The correct form would be *por otro lado* "on the other hand." This use is not accurate in form, but it was coded as an accurate use.

Excerpt 13 [Adv./Part. 4E]:

Si somos destinos por algo, es nuestras decisiones que van nos guiar a nuestras ultimas destinaciones. ***Del otro lado***, si no hay destino, solo tenemos el libre albedrio como guía.

"If we are destined for something, it is our decisions that will guide use to our final destinations. **On the other hand,** if there is no destiny, we only have free will as a guide."

In excerpt 14, the participant used the preposition *en* "at" instead of the form *por el contrario* "on the other hand."

Excerpt 14 [Adv./Part. 2E]:

Anna era una persona con la mentalidad de que quedamos donde estamos porque eso es nuestro destino, nada puede cambiarlo. ***Pero en contrario*** mi madre y sus otras hermanas creían que el destino es el resultado de nuestras acciones y decisiones.

"Anna was a person with the mentality that we stay where we are because that is our destiny, nothing can change it. But **On the other hand**, my mother and her other sisters believed that destiny is the result of our actions and decisions."

Similarly, the correct form in excerpt 15 is *por el contrario* "on the other hand," but the participant used the wrong preposition *en* "in/on/at" instead of *por* "for"

Excerpt 15 [Adv./ Part. 2E]:

Si ella hubiera tenido la mentalidad positiva como sus hermanas, Marina habría tomado decisiones mejores para recibir las oportunidades que ella quería y su destino, o sea las consecuencias de sus acciones y decisiones, no la dirigía. ***En contrario***, mi madre y sus otras hermanas no tuvieron el pensamiento que el destino nos define.

"If she had had a positive mentality like her sisters, Mariana would have made better decisions to get the opportunities she wanted and her destiny, that is, the consequences of her actions and decision were not driving her. On the other hand, my mother and her other sisters didn't hold the thought that destiny defines us."

To sum up, this research question investigated the accuracy of the discourse markers used. The original hypothesis was that higher levels of proficiency would use discourse markers more accurately; this was found to be false, as more proficient groups had a lower level of accuracy. As a result, two new questions arose: a) Does the decrease in the accuracy of discourse markers correlate with the number of tokens used? and b) Does the decrease in the accuracy of discourse markers correlate with the variety of discourse markers used? Two bootstrapped correlation analysis were performed to answer the new questions. The first statistical analysis showed that the accuracy rates were not correlated to a higher use of discourse markers, that is, more discourse markers did not necessarily translate into more chances of error. The second statistical analysis showed that a greater variety of discourse markers used didn't correlate with more chances of error, either. As none of these hypotheses was confirmed, a qualitative observation of the errors was deemed necessary.

The qualitative error analyses showed that the errors registered in the data collected could be classified into three groups: errors in word choice, omission of an element of discourse markers that come in pairs, and errors in the form of the discourse marker. The most frequent error type was Word Choice with 8 cases, the second more frequent type of error was Element Left Out with 4 occurrences, and only three cases of Word Form errors were coded, the latter all in the Advanced group.

3.4. Is There a Correlation Between the Use of Discourse Markers and the Overall Quality of the Compositions?

In this section, I first explain the procedures to determine the quality of the compositions, the reliability scores, and the distribution of the quality scores. I made three hypotheses based on previous research that identified a positive correlation between the use of discourse markers and the quality of the compositions: Yang and Sun (2012), Jalilifar (2008), Liu and Braine (2005), and Lahuerta (2004) reported a positive correlation for the number of discourse markers used, and Intaraprawat and Steffensen (1995) and Wenjuan (2016) reported positive correlations for the variety of discourse markers used. As the number of incorrect uses identified were very low (15 cases in the entire sample), I discarded analyzing the data for a correlation between accuracy in discourse markers used and quality of the compositions. The hypothesis are examined in the following order:

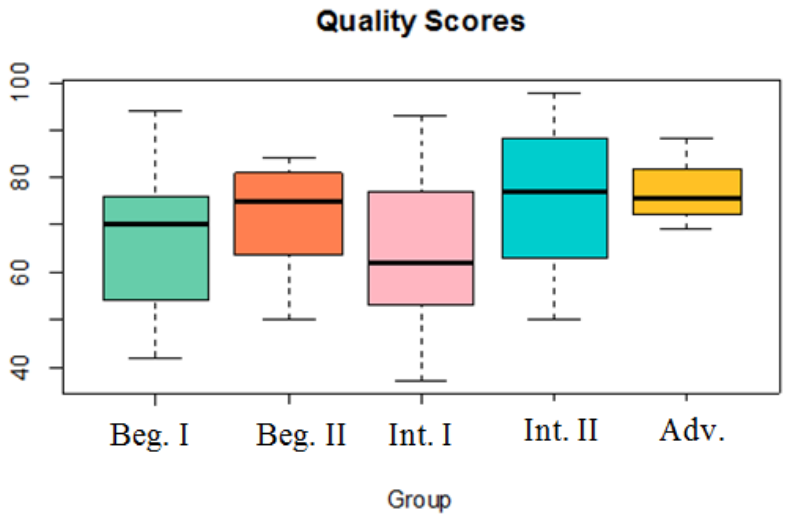
- H1. There is a correlation between the ratio of discourse markers used and the overall quality score of the compositions.
- H2. There is a correlation between the number of discourse markers used and the overall quality score of the compositions.
- H3. There is a correlation between the variety of discourse markers used and the overall quality score of the compositions.

3.4.1 Quality assessment

The samples were assessed using the rubric presented in Appendix 3. The raw quality scores for each composition can be found in the Appendix 4. In order to determine the reliability of the quality assessment, 30% of the compositions per level were randomly selected (Dörnyei, 2007) through *random.org* and rated by two interraters. The total score for reliability was calculated using Cronbach's Alpha formula, yielding a score of 0.947, which indicated a high consistency in the scores (Dörnyei, 2007; Roever, 2017).

Figure 3.9 presents the dispersion of the quality scores using boxplots. The mean score obtained in the five groups (with standard deviations in brackets) are as follows: Beginners I = 66.14 (14.54), Beginners II = 71.62 (12.12), Intermediate I = 63.52 (15.31), Intermediate II = 75.50 (16.37), Advanced = 77.00 (7.95). An ANOVA revealed that there were no significant differences in the quality scores between groups at the $p<.05$ level [$F(4,59)=1.66$, $p=0.17$].

Figure 3.9 *Boxplot of Quality Scores*



3.4.2 Hypothesis 1. There is a correlation between the ratio of discourse markers used and the overall quality score of the compositions.

Ratios were calculated in order to determine the proportion of discourse markers used in a composition based on the length of the text. The raw number of discourse markers (tokens), the total number of words, and the corresponding ratios are presented in Table 3.21.

The mean ratios obtained in the five groups (with standard deviations in brackets) were: Beginners I = 0.028 (0.014), Beginners II = 0.037 (0.013), Intermediate I = 0.037 (0.011), Intermediate II = 0.022 (0.011), Advanced = 0.028 (0.004). Figure 3.10 illustrates the distribution of the ratio of discourse markers used.

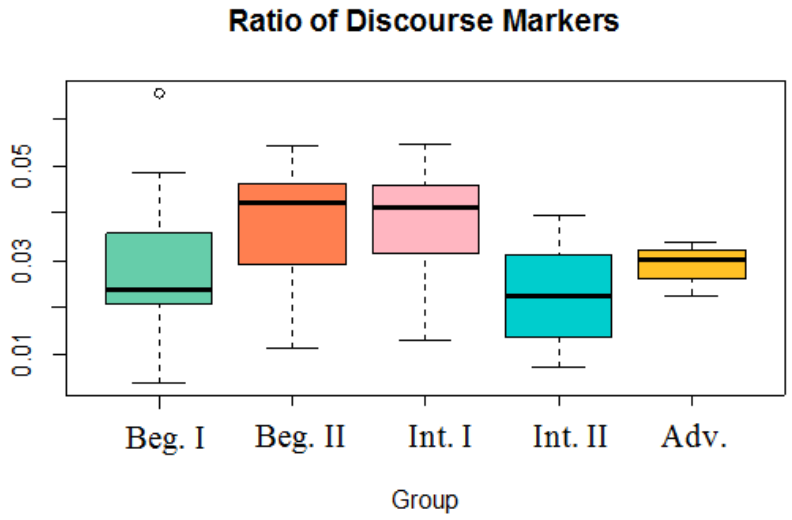
Table 3.21 *Raw Number of Discourse Markers (Tokens) Used*

Sample	Tokens	Words	Ratio	Sample	Tokens	Words	Ratio
2A	14	287	0.048	4C	15	344	0.043
3A	5	209	0.023	5C	12	227	0.052
4A	10	267	0.037	6C	8	238	0.033
5A	16	244	0.065	7C	9	282	0.031
6A	8	224	0.035	8C	10	237	0.042
7A	15	427	0.035	9C	8	216	0.037
8A	7	234	0.029	10C	11	271	0.040
9A	2	170	0.011	11C	10	398	0.025
13A	10	218	0.045	12C	7	226	0.030
14A	4	236	0.016	13C	8	191	0.041
16A	5	210	0.023	14C	11	201	0.054
18A	1	246	0.004	15C	10	209	0.047
21A	10	232	0.043	16C	9	218	0.041
22A	8	319	0.025	17C	10	203	0.049
23A	4	269	0.014	18C	12	253	0.047
26A	3	282	0.010	19C	4	218	0.018
28A	5	240	0.020	20C	7	272	0.025
30A	6	262	0.022	21C	7	209	0.033
31A	8	243	0.032	22C	3	229	0.013
32A	5	243	0.020	23C	12	271	0.044

Table 3.21 (Continued) *Raw Number of Discourse Markers (Tokens) Used*

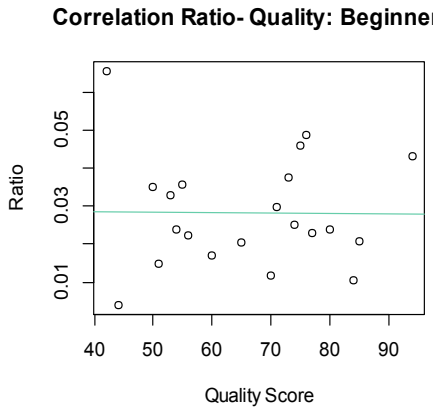
Sample	Tokens	Words	Ratio	Sample	Tokens	Words	Ratio
40A	5	224	0.022	1D	4	275	0.014
1B	7	170	0.041	2D	7	279	0.025
2B	9	166	0.054	3D	8	315	0.025
3B	7	152	0.046	4D	2	269	0.007
4B	7	164	0.042	5D	5	246	0.020
5B	7	150	0.046	6D	4	311	0.012
6B	2	177	0.011	7D	7	178	0.039
7B	4	156	0.025	8D	9	242	0.037
8B	5	154	0.032	1E	8	267	0.029
1C	9	220	0.040	2E	18	531	0.033
2C	4	201	0.019	3E	7	229	0.030
3C	11	220	0.050	4E	5	221	0.022

Figure 3.10 *Boxplot of the Ratio of Discourse Markers Used.*



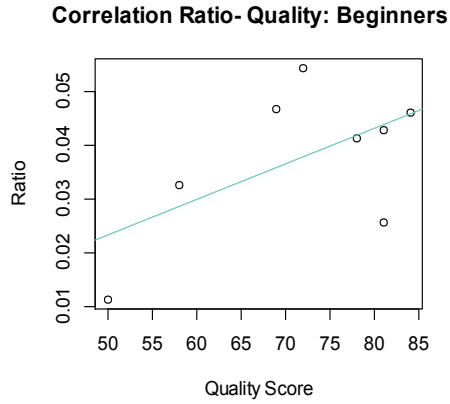
A bootstrapped (10,000 replications) Pearson correlation between the ratio of discourse markers and the quality scores of the compositions for Beginners I found the effect size of the correlation was tiny, and the CI extremely wide (95% CI [-0.44, 0.53], r=0.04, N=21, R^2=.001), meaning the correlation coefficient is not highly reliable. Figure 3.11 presents a graph of a linear correlation test for illustrative purposes. We can see that the correlation is very weak.

Figure 3.11 *Correlation Between Ratio of Discourse Markers Used and Quality of the Compositions: Beginners I.*



A bootstrapped (10,000 replications) Pearson correlation between the ratio of discourse markers and the quality scores of the compositions of Beginners II also found that the effect size of the correlation was small, and the CI fairly wide (95% CI [-0.71 ,0.99], r=0.22, N=8, R^2=0.4). Figure 3.12 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.12 *Correlation Between Ratio of Discourse Markers Used and Quality of the Compositions: Beginners II.*



For Intermediate I, the bootstrapped (10,000 replications) Pearson correlation again found that the effect size of the correlation was small, and the CI wide (95% CI [-0.20, 0.63], r=0.26,

N=23, R^2=0.06). Figure 3.13 presents the graph of a linear correlation test for illustrative purposes.

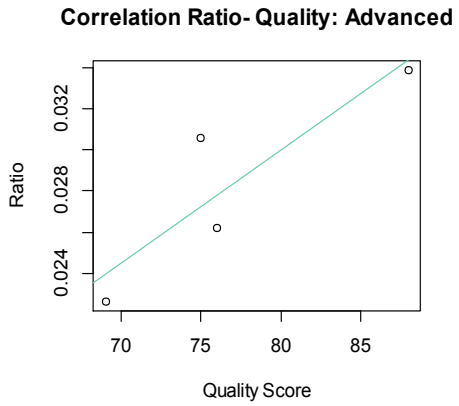
Figure 3.13 *Correlation Between Ratio of Discourse Markers Used and Quality of the Compositions: Intermediate I.*

For Intermediate II, the bootstrapped (10,000 replications) Pearson correlation showed no correlation between the ratio of discourse markers and the quality scores of the compositions of Intermediate II, with a negligible effect size, and the CI spanning (i.e. as a line) (95% CI [-0.74 , 0.74], *r*=0.01, N=8, R^2=0.001). Figure 3.14 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.14 *Correlation Between Ratio of Discourse Markers Used and Quality of the Compositions: Intermediate II.*

Finally, for the Advanced group, the bootstrapped (10,000 replications) Pearson correlation found the effect size of the correlation was large, and the upper bound of the CI non-determined due to the very small sample size (95% CI [0, n/a], r=0.98, N=4, R^2=0.96). Figure 3.15 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.15 *Correlation Between Ratio of Discourse Markers Used and Quality of the Compositions: Advanced.*



The histogram of the five groups shows several outliers. An O-type bootstrapped (10,000 replications) correlation was performed between the ratio of discourse markers and the quality scores of all five groups. The outliers were detected and removed by mathematical methods, and the Pearson correlation was calculated with the remaining data. The bootstrapped (10,000 replications) Pearson correlation found the effect size of the correlation was negligible, and the CI narrow (95% CI [-0.24, 0.34], r=0.04, N=64, R^2=0.0016). Figure 3.16 presents the graph of a linear correlation test for illustrative purposes.

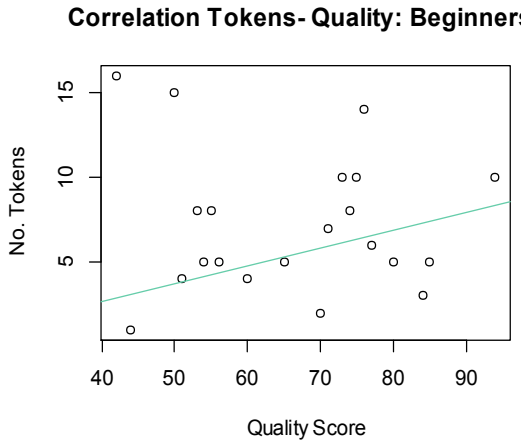
Figure 3.16 *Correlation Between Ratio of Discourse Markers Used and Quality of the Compositions: All Groups.*

The analyses showed that there was a small effect size for the ratio of discourse markers and the quality of the composition for the Beginners II and Intermediate I groups. The effect size was large for the Advanced group.

3.4.3. Hypothesis 2. There is a correlation between the number of discourse markers used and the overall quality score of the compositions.

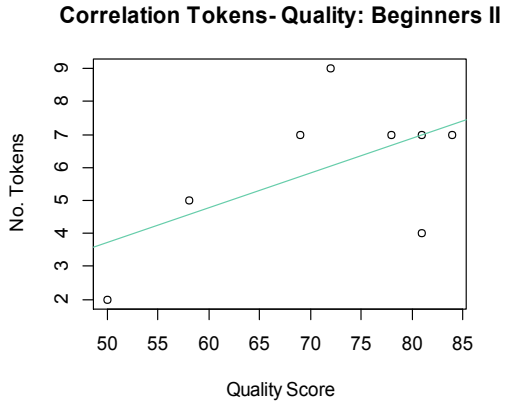
A bootstrapped (10,000 replications) Pearson correlation between the number of discourse markers used and the quality scores of the compositions for Beginners I found the effect size was very small, and the CI extremely wide (95% CI [-0.47, 0.52], r=0.03, N=21, R^2=0.0009), meaning the correlation coefficient is not highly reliable. Figure 3.17 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.17 *Correlation Between Number of Tokens Used and Quality of the Compositions: Beginners I.*



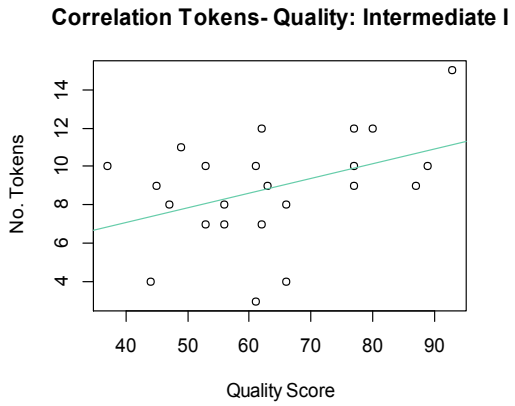
For Beginners II, the bootstrapped (10,000 replications) Pearson correlation found the effect size to be small, and the CI fairly wide (95% CI [-0.62, NA], r=0.27, N=8, R^2=0.07). Figure 3.18 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.18 *Correlation Between Number of Tokens Used and Quality of the Compositions: Beginners II.*



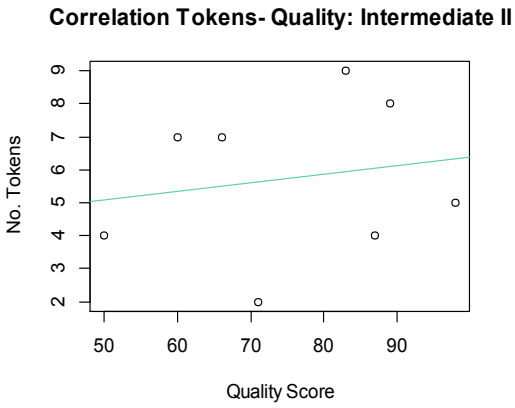
For Intermediate I, the bootstrapped (10,000 replications) Pearson correlation found the effect size to be medium, and the CI was wide (95% CI [0.04, 0.75], r=0.48, N=23, R^2=0.23). Figure 3.19 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.19 *Correlation Between Number of Tokens Used and Quality of the Compositions: Intermediate I.*



For Intermediate II, the bootstrapped (10,000 replications) Pearson correlation found that the effect size of the correlation was very small, and the CI was wide (95% CI [-0.61, 0.95], r=0.09, N=8, R^2=0.008). Figure 3.20 presents the graph of a linear correlation test for illustrative purposes.

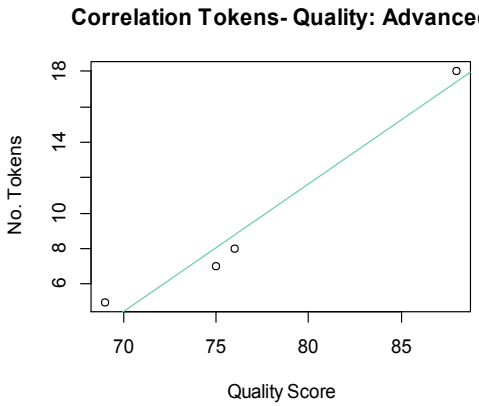
Figure 3.20 *Correlation Between Number of Tokens Used and Quality of the Compositions: Intermediate II.*



Finally, for the Advanced group, the bootstrapped (10,000 replications) Pearson correlation found the effect size of the correlation to be large, and the CI narrow (95% CI [0.99,

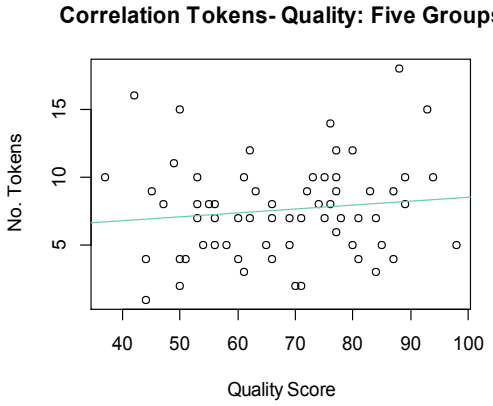
NA], *r*=0.99, N=4, R^2=0.98). Figure 3.21 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.21 *Correlation Between Number of Tokens Used and Quality of the Compositions: Advanced.*



A bootstrapped (10,000 replications) Pearson correlation for all five levels found the effect size of the correlation to be very small, and the CI [missing word here] (95% CI [-0.14, 0.37], *r*=0.11, N=64, R^2=0.01). Figure 3.22 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.22 *Correlation Between Number of Tokens Used and Quality of the Compositions: All Groups.*

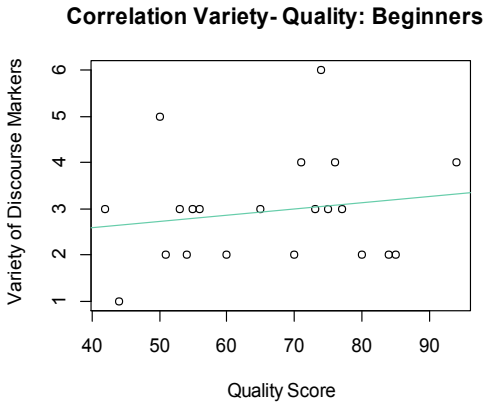


The analyses showed that there is a correlation between the number of discourse markers used and the overall quality of the compositions. The effect was negligible for the Beginners I and Intermediate II groups. There was a small effect for the Beginners II group. The effect was medium for the Intermediate I group, and large for the Advanced group.

3.4.4 Hypothesis 3. There is a correlation between the variety of discourse markers used and the overall quality score of the compositions.

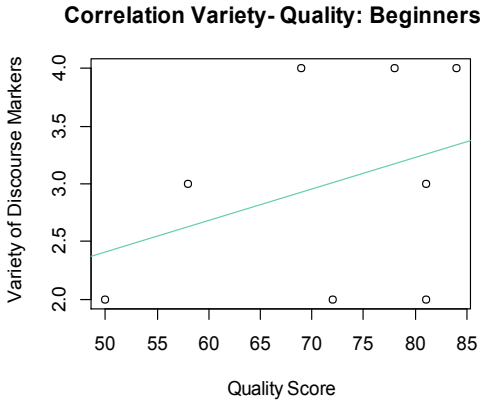
A bootstrapped (10,000 replications) Pearson correlation between the variety of discourse markers used and the quality scores of the compositions for Beginners I found the effect size of the correlation was small, and the CI narrow (95% CI [-0.33, 0.57], r=0.11, N=21, R^2=0.01). Figure 3.23 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.23 *Correlation Between Variety of Tokens Used and Quality of the Compositions: Beginners I.*



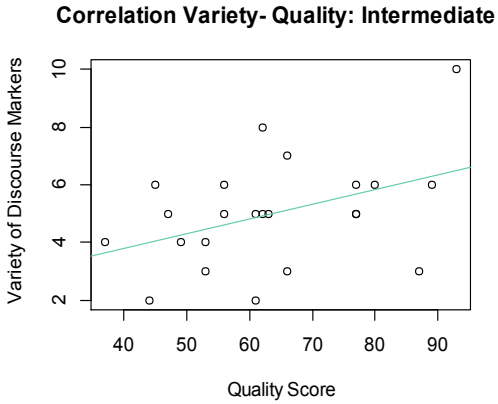
For Beginners II, the bootstrapped (10,000 replications) Pearson correlation correlation found the effect size to be small, and the CI wide (95% CI [-0.51, NA], r=0.25, N=8, R^2=0.06). Figure 3.24 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.24 *Correlation Between Variety of Tokens Used and Quality of the Compositions: Beginners II.*



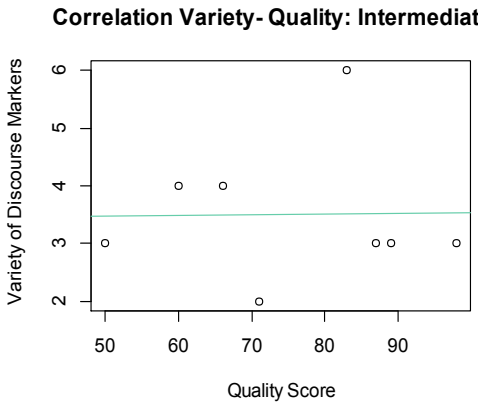
For Intermediate I, the bootstrapped (10,000 replications) Pearson correlation found the effect size to be medium, and the CI narrow (95% CI [0.04, 0.72], r=0.42, N=23, R^2=0.17). Figure 3.25 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.25 *Correlation Between Variety of Tokens Used and Quality of the Compositions: Intermediate I.*



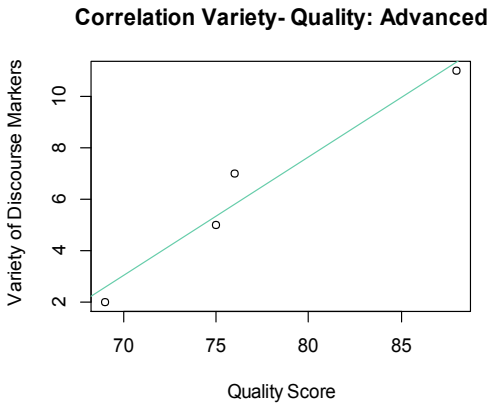
For Intermediate II, the bootstrapped (10,000 replications) Pearson correlation found the size of the correlation to be small, and the CI spanning (95% CI [-0.89, NA], r=-0.20, N=8, R^2=0.04). Figure 3.26 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.26 *Correlation Between Variety of Tokens Used and Quality of the Compositions: Intermediate II.*



For the Advanced group, the bootstrapped (10,000 replications) Pearson correlation found the effect size to be large, and the CI showed a linear correlation (95% CI [0.98, NA], r=0.98, N=4, R^2=0.96). Figure 3.27 presents the graph of a linear correlation test for illustrative purposes.

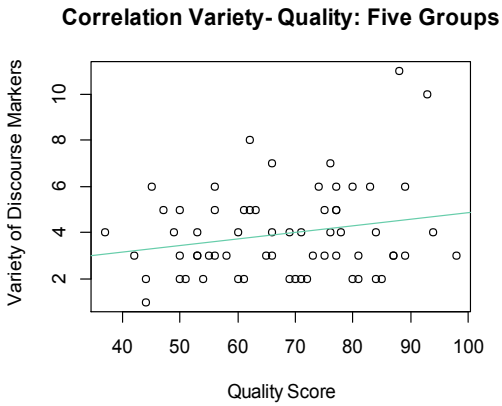
Figure 3.27 *Correlation Between Variety of Tokens Used and Quality of the Compositions: Advanced.*



Finally, the bootstrapped (10,000 replications) Pearson correlation between the variety of discourse markers used and the quality scores of the compositions of the five groups found the

effect size was small, and the CI was narrow (95% CI [-0.13,0.37], r=0.09, N=64, R^2=0.008). Figure 3.28 presents the graph of a linear correlation test for illustrative purposes.

Figure 3.28 *Correlation Between Variety of Tokens Used and Quality of the Compositions: All Groups.*



In this section I examined the correlation between the quality of the compositions and the ratio, the number, and the variety of discourse markers used. The analyses showed a positive correlation between the number of discourse markers used and the quality of the compositions for the Beginners II, Intermediate I, and Advanced groups, and between the variety of discourse markers used and the quality of the composition for the Intermediate I and the Advanced groups. No correlation was found between the ratio of discourse markers and the quality of the compositions.

Chapter 4. Discussion

In this chapter, the findings that were presented in Chapter 4 are discussed in relation to previous research. I begin with a brief summary of the findings for each research question. Following the summary, the results are discussed in turn. In the final part of the chapter, I discuss the implications of this research for teaching Spanish as a Foreign Language. The limitations of the study are addressed subsequently.

4.1 Does the frequency of discourse markers use change with proficiency level?

The goal of this research question was to determine if the density of discourse markers used by the participants would increase as the proficiency level increased. It was hypothesized that more advanced students would use more discourse markers. In this regard, I found that there was a significant effect of group on the density of discourse markers for certain levels but not others. The bootstrapped comparisons indicated that the difference was statistically significant between Beginners I and Intermediate I, with an increase of at least 18% and possibly as high as 170%. There was a medium (but non-significant) effect size between the Intermediate II and Advanced group (Cohen's *d*=-0.94), and a small (but non-significant) effect size between the Beginners I and Advanced group (Cohen's *d*=-0.42). These increases may provide support in favor of the hypothesis, perhaps due to more exposure via language instruction. I would like to emphasize that even though the statistical analysis didn't reveal significant differences between the Beginners I and Advanced groups or between the Intermediate I and Advanced groups, the increase in the density of use of discourse markers is not negligible because the width of the SD curves of both groups overlap considerably, making it difficult to find statistically significant

values. The small sample size of the Advanced group could have made it difficult to reach statistical significance as well.

I would like to point out that the Intermediate II group performed in an unexpected way, in that they produced a lower density of discourse markers than the other four groups. The bootstrapped comparisons showed strong effect sizes in the comparisons between Beginners II and Intermediate II (Cohen's *d*=1.17), and between Intermediate I and Intermediate II (Cohen's *d*=1.24). This decrease could be related to the task itself; as already mentioned, the format of the letter required students to write information such as name, recipient, address, etc., that would be unlikely to require the use of discourse markers. Nonetheless, the average word count of the texts produced by the Intermediate II group increased by 40% when compared with Beginners II and by 10% when compared with Intermediate I, providing enough opportunity for the use of discourse markers. It could well be that for this particular task, discourse markers were not particularly useful; however, the statistically significant lower density and variety of discourse markers used requires further study.

There are few studies to date that provide the overall frequency of the devices analyzed —discourse markers, connectors, cohesive devices, discursive operators, etc. — while comparing different levels of proficiency or length of language instruction. In a multidimensional analysis of Spanish L2 compositions, Asención Delaney and Collentine (2011) found that second- and third-year university students used few cohesive devices in general; the frequencies were not reported, so there is no available data to be compared with the findings of this study. In their literature review, Hu and Li (2015) quote Lu (2012) who found that second-year students of English use discourse connectives with a higher frequency that first-year students. Due to language barriers, I couldn't retrieve Lu's (2012) thesis to compare with my own

data. Jalilifar (2008) found that more proficient English learners, the Graduate group, used a higher frequency of cohesive devices than the Senior and Junior groups. The reported ratio of discourse markers was 33.7 per 1,000 words for the most experienced group and 20.83 per 1,000 words for the Junior group. The latter group is most comparable with the Advanced group in my research, who used 32.7 discourse markers per 1,000 words, considerably higher than Jalilifar's findings and more comparable to the Graduate group in Jalilifar's study, although the Advanced group in my study is likely to be of a lower proficiency level. Unfortunately, there is not enough available information to compare findings across studies.

In sum, the literature on the topic and the present research show that the frequency of discourse markers used generally increases with the proficiency level; however, this increase hasn't always been statistically significant. On the other hand, it has been noted that the overuse of discourse markers is a common practice among L2 learners (Meisuo, 2000; Jalilifar, 2008; Rahimi, 2011; Hu & Li, 2015). For these reasons, it is necessary to also examine the variety of discourse markers used and the relationship between the use of discourse markers and the overall quality of the compositions.

4.2 How Varied is the Use of Discourse Markers?

The second research question concerned the variety of discourse markers used by the participants. It was hypothesized that the variety of discourse markers would increase parallel to the proficiency of the groups. The results showed that the groups Beginners I and Beginners II used only two types of discourse markers, connectors and discursive operators, whereas the Intermediate I, Intermediate II, and Advanced groups used all three types, connectors, discursive

operators, and metatextual connectors. The connectors that were most frequently used were *y* "and," *porque* "because," *también* "also," and *pero* "but." The metatextual connectors with the highest number of occurrences were *antes* "before," *ahora* "now," and *luego* "later." The discursive operators with the highest number of instances were *si* "if," *en mi opinión* "in my opinion," and *por ejemplo* "for example." The cross-analysis of frequency of use and level showed that the variety of connectors increased for the Contrastive and Consecutive subcategories. For the discursive operators, the three used were each from a different subcategory, so it was not possible to observe any progressive pattern in their use. As for the metatextual connectors, there were too few tokens overall to hypothesize a pattern. Regarding the statistical analysis, the results showed statistically significant differences between: a) Beginners I with the Intermediate I and Advanced groups; b) Beginners II with the Intermediate I and Advanced groups; and c) Intermediate I and Intermediate II. This last comparison is due to the unusual and unexpected behavior of the Intermediate II group mentioned earlier. If we disregard it, the hypothesis that the variety of discourse markers increases alongside proficiency seems to hold.

Within the connector category, additive connectors showed the largest percentage of use (42.38%), followed by causative connectors (26.74%), contrastive connectors (10.90%), consecutive connectors (4.32%), conditional connectors (2.26%) and temporal connectors (0.82%). The predominant use of additive connectors is consistent with the findings of Green et al. (2000), Meisuo (2000), Lahuerta (2002), Lahuerta (2004), Jalilifar (2008), Rahimi (2011), and Hu and Li (2015). The additive connectors *y* "and," *también* "also," and *además* "in addition" link segments at the same level, accounting for their extensive use, even in different languages. It is worth noting that the use of additive connectors decreased dramatically in the

Advanced group. Although it wasn't the aim of this study to analyze syntactic complexity in the compositions, it was evident that the syntactic complexity did increase (see Asención and Collentine, 2011). The question then arises as to whether there is a relationship between the increase of syntactic complexity and the use of fewer additive connectors. On the other hand, not all the studies have found prominence of additive connectors. Field and Yip (1992) found that contrastive connectors were more prominently used by three of the four groups they studied, although this wasn´t observed in my data.

Regarding metatextual connectors, Kobayashi and Rinnet (2008) found higher frequencies of metatextual connectors over inter-sentential markers (connectors in my analysis). These results contradict what I found; in my data, the six metatextual connectors used constitute only 7.40% of the tokens. Four of these discourse markers are temporal: *antes* "before," *ahora* "now," *después* "after," and *luego* "later." Only one of the items in this category is a discourse organizer, *por otro lado* "on the other hand," and one item is used to conclude, *en fin* "finally." The majority of these temporal metatextual connectors (33/36) were produced by the Intermediate I class, possibly as a result of these connectors being taught in the unit or of the nature of the task for this group. Furthermore, the scarce use of metatextual connectors —only 3 occurrences in the other four groups— suggest that there is low awareness of this type of connectors among the students across the levels. Considering the results of Hernández (2011) in which the combination of explicit instruction and input flood had a positive effect on students' use of discourse markers to narrate a past event in comparison to a group that didn't receive any type of instruction, and de la Fuente (2009) in which she advises the use of consciousness-raising tasks to foster explicit learning and metalinguistic awareness as necessary cognitive steps to

learn L2 discourse markers (p. 2018), it may be beneficial to target the instruction of metatextual connectors in the classroom.

For the category of discursive operators, only three discourse markers were identified in the data; as a group they represent 4.5% of all the discourse markers tallied. The discourse marker *por ejemplo* "for example" was the most frequently used one. It was produced by the Beginners and the Intermediate groups, although surprisingly, there were no occurrences in the advanced group. The overall low frequency of use of *por ejemplo* "for example" aligns with the results of Rodriguez (2016), where *por ejemplo* "for example" had a low frequency and wasn't used, either, by the most proficient writers. The discourse marker *en mi opinión* "in my opinion" was mainly used by the Intermediate I group (9/10 occurrences). This uneven distribution was unexpected considering the argumentative nature of the tasks; however, further review of the data revealed that other participants chose alternative forms to express their opinion. There was only one occurrence of the discourse marker *de hecho* "in fact" in the Intermediate I group. These results are also in line with Rodriguez (2016), who found a low frequency of this discourse marker, with most tokens produced by the students at the C1 (Proficient User) level, as well as occurrences at the A1, B1, and B2 levels.

In sum, with respect to the variety of discourse markers, a statistically significant increase in the variety of discourse markers used was observed from Beginners I and II to Intermediate I, and from Beginners I and II to Advanced. In addition, the increase in variety from Intermediate I to Advanced had a small, though not significant, effect. Additive connectors were overwhelmingly the most frequently used discourse markers, except in the Advanced group, perhaps due to an increase in syntactic complexity, and there was an increase in variety for the contrastive and consecutive subcategories as proficiency increased. Metatextual connectors,

mostly from the temporal subcategory, were the least used discourse markers and were not produced at all by the two Beginner groups. Lastly, it seemed that learners, especially at the beginner and intermediate levels, associated one form with one function for the majority of the discourse markers, with the exception of the additive connectors *y* "and" and *también* "also" that were used simultaneously by participants in all levels.

4.3 How Accurate is the Use of Discourse Markers?

At the beginning of the research, it was hypothesized that higher levels of proficiency would use discourse markers more accurately. The Kruskal-Wallis analysis of variance revealed the opposite, with a statistically significant decrease between the Beginners I and Advanced groups, and between the Intermediate I and Advanced groups. The Beginners I group had an accuracy rate of 100%, but they overwhelmingly used connectors (98% of tokens). A possible explanation for the high accuracy rate is that the similarity between the Spanish connectors and the English equivalents reduced the chances of confusion. The decrease in accuracy in the Advanced group aligns with Yang and Sun's (2012) results; they found that the senior students had a higher frequency of error in cohesion conjunctions [M=0.279 (SD=0.107)] than the sophomore students [M=0.219 (SD=0.040)]. These results could perhaps be explained by the length of the compositions, a higher number of discourse markers used, or by the variety of tokens used. Following this logic, more tokens or more variety of discourse markers would translate into more chances of error. However, the statistical analyses found no correlation between these factors.

In light of these indeterminate results, a qualitative error analysis was performed. None of the previous studies consulted for this research analyzed the type of errors in the use of discourse markers in L2 Spanish compositions. There are several studies in EFL that report the misuse of contrastive connectors as a prominent type of error (Field & Yip, 1992; Meisuo, 2000; Lu, 2012), and I also identified the misuse of contrastive discourse markers in 3 out of the 15 errors (20%), under the error category of Word Choice. There were only four errors coded under the category of Element Left Out, and they were all produced by the Intermediate I and Intermediate II groups, who seemed unaware that some discourse markers appear in two parts. Otherwise, I couldn't identify any further pattern in these errors. The third type of error was Word Form; only three cases were identified and all of them were observed in the advanced group. These errors had a milder effect on the intelligibility of the text. It seems that part of the challenge for the accurate use of discourse markers is not only the type of discourse marker or the semantic function of the marker, but also the number of words that form it; for instance, discourse markers with three words represented 6 out of 15 errors, 5 of the errors were discourse markers formed by two words, and only 2 of the errors were one word discourse markers. In the data collected, there were 5 three-word discourse markers, 8 two-word discourse markers, and 14 one-word discourse markers. It seems that errors were more likely to occur when the discourse marker consists of more than one word.

4.4 Is There a Correlation Between the Use of Discourse Markers and the Overall Quality of the Compositions?

The fourth research question investigated the relationship between the quality of the compositions and the use of discourse markers The hypothesis was that there is a correlation between the quality of the compositions and the ratio of discourse markers used, the raw number of discourse markers used, and the variety of discourse markers used. The results revealed a positive correlation between the ratio of discourse markers and the quality of the compositions for the Beginners II, Intermediate I, and Advanced groups, with a small effect size for the first two groups and a large effect size for the Advanced group. There was a statistically significant correlation between the number of discourse markers and the quality of the compositions for the Beginners II, Intermediate I, and Advanced groups, ranging from a small correlation in the Beginners II group to a medium effect for the Intermediate I group to a strong correlation in the Advanced group. There was also a positive correlation between the variety of discourse markers used and the quality of the composition for all but the Intermediate II group, with a small effect size for Beginners I and II, a medium effect fo Intermediate I, and a large effect for the Advanced group.

The analysis of the correlation coefficients shows that the ratio of discourse markers (discourse markers divided by number of words) is not a good predictor of the quality of the compositions. Even though the analysis showed a strong correlation for the Advanced group, there was no consistency in the results. A small correlation was identified in the Beginners II and Intermediate II groups; however, the CIs were fairly wide, indicating that these correlations are not trustworthy. In addition, calculating ratios according to the number of words may not be entirely informative. Discourse markers join different levels of the text, that is, metatextual

connectors connect paragraphs and multi-sentential chunks of discourse, while connectors join sentences; therefore, other ways to calculate ratios in future research could be T-units, lexical bundles, etc. I opted to use the number of words as the denominator because the samples I gathered reflect a broad range of proficiency levels, making it difficult to establish clear divisions between the sentences. Therefore, the difference in syntactic complexity between the groups would interfere in the count of units.

Regarding the correlation between the number of discourse markers used and the quality of the compositions, a positive but not significant correlation was found for Beginner II, and a positive and significant correlation for the Intermediate I and Advanced groups, but not for Beginners I and Intermediate II. It is not surprising that there is no clear correlation for the Beginners I group; as it has been described before, this group overused connecting devices, which would detract from the quality, while the performance of the Intermediate II group was irregular, making it difficult to determine if there is a correlation. The results of the present study seem to contradict those of Meisuo (2000) in which no correlation was found between the number of cohesive ties (conjunction ties, reference ties, and lexical ties) and the quality of the compositions. My results also differ from those of Yang and Sun (2012) who found a medium correlation (r=.405, p< .05) between the use of conjunctions and the quality scores for the Senior group but not for the Sophomore group, whereas I found positive correlation both for the Intermediate I group (similar to their Sophomore group) and the Advanced group (similar to their Senior group). Jalilifar (2008) and Liu and Braine (2005) also found a strong positive correlation between the number of discourse markers and the score of the compositions in both the Junior and the Advanced groups, while Lahuerta (2004) found a statistically significant relationship between the scores of the compositions written by first-year English students and the number of

discourse markers used (r=0.64). Considering the results from the studies outlined above where a positive correlation between the number of tokens and the quality of the compositions was found across different levels of proficiency, I conclude that the number of discourse markers in a composition is a valid predictor of the quality of compositions in Spanish as a foreign language.

With respect to the variety of discourse markers, the correlation scores ranged from very small for the Beginners I (r=0.11), small for Beginners II (r=0.25), medium for the Intermediate I group (r=0.42), small for the Intermediate II group (r=0.20), and strong for the Advanced group (r= 0.98); however, the CI were very wide, so the correlation may not be a very good estimate. These results are in line with Intaraprawat and Steffensen (1995) where the good essays showed a greater variety of metadiscourse features within each category than the poor essays. Similarly, Wenjuan (2016) found that a diversity of organizational markers and lexico-syntactic complexity were predictors of quality for EFL argumentative essays.

From the data analyzed, we can see that the variety of discourse markers used increased as the proficiency increased, except for the Intermediate II group; similarly, the correlation between the variety of discourse markers used and the quality of the composition increased as the proficiency increased. Thus, the variety of discourse markers may also be a good predictor of the quality of the compositions. In particular, the compositions of advanced learners who use a broad variety of discourse markers may receive higher quality scores. The correlation between the variety of discourse markers used and the quality scores was smaller at the lower levels; this result could partially be explained by the fact that other aspects were more notable at the beginner levels, such as the repetition of discourse markers, which was also documented by Field and Yip (1992) and Meisuo (2002). In general, the analyses of the correlation for ratio, number of tokens, and variety of discourse markers and the quality of the compositions revealed a

stronger correlation as the proficiency level increased. This suggests that the use of discourse markers is perceived to be more important at higher levels than at lower levels.

4.5 Limitations of this Study

There are five main limitations to the present study. First, the size of the groups were uneven, and some groups had a low number of participants, making it difficult to arrive at statistical conclusions. The original data sample was 110 compositions; however, 46 of those were excluded because the participants had an L1 different than English. Including the data of participants from different linguistic backgrounds would increase the sample size; however, in order to preserve the comparability of the compositions, I decided to use only data produced by participants who identified themselves as native English speakers, and robust statistical methods were used to compensate for the reduced sample size. Second, as the participants were allowed to use dictionaries, it is not possible to differentiate between the use of discourse markers and searching skills. A longitudinal study would be necessary to have a clearer picture of the acquisition pattern of discourse markers. Third, while the classification criteria for the discourse markers was selected considering its accessibility for language instructors, it is not a widespread nomenclature. In addition, the low number of studies addressing this topic in Spanish and the high variability in classification criteria in the English studies make it difficult to compare findings across languages. Fourth, the tasks were not comparable across groups as the researcher took into account the curriculum content in each course. Indeed, this lack of comparability may account for why Intermediate II performed unexpectedly, as the topic of their writing assignment was quite different from the other four groups. Future studies may wish to use a single topic and

composition format for all proficiency levels. The current study also did not test the proficiency level of the groups; it may be that the Intermediate II group was not significantly more proficient than the Intermediate I group, as course placement is not always a good reflection of actual proficiency. It is also possible that there were other differences among the groups that may have affected their performance on the task, such as variation in instruction, differences in the type of feedback they have received from current and previous instructors, as well as individual motivation and anxiety levels when faced with writing a composition in class.

Chapter 5. Conclusion

This study investigated the use of discourse markers in argumentative compositions written by learners of Spanish as foreign language and its correlation with the quality of the texts. The main findings are as follows.

The results revealed that there was a statistically significant increase in the use of discourse markers for certain levels, but not others. Therefore, it is necessary to examine other aspects of use. It was observed that there is a significant increase in the variety of discourse markers used from Beginners I and II to Intermediate I, and from Beginners I and II to Advanced. The Beginner levels used only two types of discourse markers, connectors and discursive operators, whereas the Intermediate and Advanced levels used the three types, connectors, discursive operators, and metatextual connectors. This suggests that the variety of discourse markers used increases with proficiency level. In contrast, the accuracy in the use of discourse markers decreased as the proficiency level increased. However, the gravity of the error seems to be less severe at the higher levels, involving Element Left Out or Word Form errors rather than Word Choice. This is an encouraging sign, as the use of discourse markers seems to have increasing relevance for the quality of the composition as the proficiency level increases, that is, more tokens and a broader variety of discourse markers correlates more strongly with quality scores at the higher levels.

The multiple analyses performed in this study have provided valuable information to answer the original pedagogical questions that motivated this research. The results suggest that the use of discourse markers may be indeed relevant for the quality of the compositions, and its importance increases with the level of proficiency; therefore, students would benefit from

explicit instruction of these lexical expressions. The emphasis that we should give to teach them would vary across levels, as lower-level learners seem to struggle more with the function of the discourse markers and produce a smaller variety of tokens, while higher-level learners show greater variety, though they do not always have the correct form. Nogueira da Silva (2011, 2012) recommends teaching discourse markers in relation to text sequences (see Adam, 2011). Finally, regardless of the method of instruction that we choose to follow in our classes, it is important to raise awareness among students about three aspects of discourse markers: the non-equivalence of discourse markers between English and Spanish, the specific form of some of discourse markers, such as paired use, and the counterproductive effect of overusing them. Even though, overuse was not particularly addressed on this research, it was noted that lower levels overused the same discourse marker. Therefore, it is important to address the situation in two dimensions: promoting the use of a broader spectrum of discourse markers, particularly metatextual connectors, and simultaneously raising awareness of the risks of overusing them (Meisuo, 2000; Jalilifar, 2008; Rahimi, 2011; Hu & Li, 2015). Given the importance of discourse markers for producing cohesive and coherent compositions, it is imperative that we find more ways to incorporate them into L2 writing instruction.

References

Adam, J. (2011). *Les textes types et prototypes: séquences descriptives, narratives, argumentatives, explicatives, dialogales et genres de l'injonction-instruction* (3rd ed.) Nouv. éd. entièrement rev. et augm. Paris: A. Colin.

Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35 (2), 79-113. https://doi.org/10.1017/S0261444802001751

Andriessen, J., & Coirier, P. (1999). *Foundations of argumentative text processing.* Amsterdam: Amsterdam University Press.

Asención Delaney, Y., & Collentine, J. (2011). A multidimensional analysis of a written L2 Spanish corpus. *Applied Linguistics*, (32)3, 299-322. Retrieved from https://doi-org.login.ezproxy.library.ualberta.ca/10.1093/applin/amq053

Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *Modern Language Journal,* 100(1), 320-340. Retrieved from https://doi.org/10.1111/modl.12316

Calsamiglia, H., &Tusón, A. (2001). *Las cosas del decir. Manual de análisis del discurso*. Barcelona: Ariel.

Cuenca, M. J. (1995). Mecanismos lingüísticos y discursivos de la argumentación. *Comunicación, Lenguaje y Educación*, *7*(2), 23-40. Retrieved from https://doi.org/10.1174/021470395321340411

Cumming, G. (2012). *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis.* New York: Routledge.

De Beaugrande, R., & Dressler, W. U. (1981). *Introduction to text linguistics.* London: Longman.

De la Fuente, M. (2009). The role of pedagogical tasks and focus on form in acquisition of discourse markers by advanced learners. In Leow, R.P, Campos H., & Lardiere, D. (Eds.), *Little words: their history, phonology, syntax, semantics, pragmatics, and acquisition* (pp. 211-21). Washington, DC: Georgetown University Press.

Dörnyei, Z. (2007). *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies.* Oxford: Oxford University Press.

Field, Y., & Yip, L. (1992). A comparison of internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal*, *23*(1), 15-28. Retrieved from https://doi.org/10.1177/003368829202300102

Gleason, M. M. (1999). The role of evidence in argumentative writing. *Reading & Writing Quarterly*, 14, 81-106.

Golder, C., & Coirier, P. (1996). The production and recognition of typological argumentative text markers. *Argumentation*, 10, 271-282.

Green, C. F., Christopher, E. R., & Jaquelin, L. K. M. (2000). The incidence and effects on coherence of marked themes in interlanguage texts: A corpus-based enquiry. *English for Specific Purpose*, 2, 99-113. Retrieved from https://doi.org/10.1016/S0889-4906(98)00014-3

Gries, S. Th. (2013). *Statistics for linguistics with R: a practical introduction* (2nd rev. ed.). Berlin: De Gruyter Mouton.

Hernández, T. A. (2011). Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research*, 15(2), 159-182. Retrieved from https://doi.org/10.1177/1362168810388694

Howell, D. C. (2014). *Fundamental statistics for the behavioral sciences* (8th ed.). Belmont, Calif.: Wadsworth, Cengage Learning.

Hu, C., & Li, Y. (2015). Discourse connectives in L1 and L2 argumentative writing. *Higher Education Studies*, *5*(4), 30. Retrieved from DOI:10.5539/hes.v5n4p30

Intaraprawat, P., & Steffensen, M. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing*, 4(3), 253-272. Retrieved from https://doi.org/10.1016/1060-3743(95)90012-8

Jacobs, H. L. (1981). *Testing ESL composition: a practical approach.* Rowley, MA.: Newbury House.

Jalilifar, A. (2008). Discourse markers in composition writings: The case of Iranian learners of English as a foreign language. *English Language Teaching*, 1(2), 114-122. Retrieved from DOI:10.5539/elt.v1n2p114

Kline, R. B. (2004). *Beyond significance testing: reforming data analysis methods in behavioral research.* Washington, DC.

Kobayashi, H., & Rinnert, C. (2008). Task response and text construction across L1 and L2 writing. *Journal of Second Language Writing*, 17(1), 7-29. Retrieved from https://doi.org/10.1016/j.jslw.2007.08.004

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Lahuerta, A. C. (2002). El uso de los marcadores del discurso por aprendices de español como lengua extranjera. *Aula Abierta*, (80), 153-176. Retrieved from https://dialnet-unirioja-es.login.ezproxy.library.ualberta.ca/servlet/articulo?codigo=307654

Lahuerta, A. C. (2004). Discourse markers in the expository writings of Spanish university students. *Ibérica*, *8*, 63-80.

Lahuerta, A. C.,& Pelayo, M. F. (2003). Usos marginales de los marcadores del discurso: Su efecto en la comprensión lectora en español como lengua extranjera. *Ibérica*, 5, 49-68.

Larson-Hall, J. (2016). A *guide to doing statistics in second language research using SPSS and R.* New York, NY: Routledge.

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65(1), 127-159.

Leech, N. L. & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & Quantity: International Journal of Methodology*, (2), 265. Retrieved from http://dx.doi.org/10.1007/s11135-007-9105-3

Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, 33,623-636. Retrieved from https://doi.org/10.1016/j.system.2005.02.002

Lu, Y. J. (2012). A study of discourse markers in English argumentative writing by English majors: A perspective of relevance theory (Master's thesis). Ganan Normal University, Ganzhou, China.

Martín Zorraquino, M. A., & Montolío Durán, E. (Eds.) (1998). *Los marcadores del discurso. Teoría y análisis*. Madrid: Arco Libros.

Martín Zorraquino, M. A., & Portolés, J. (1998). Los marcadores del discurso. In Bosque, I., & Demonte, V. (Eds.), *Nueva gramática descriptiva de la lengua española* (pp. 4051-4213). Madrid: Espasa-Calpe.

Martín-Peris, E. & Sans-Baulenas, N. (2013). *Gente Hoy 1. Libro del alumno + CD.* Barcelona: Difusión.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: a model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, 8, 299-325. Retrieved from http://www.jstor.org.login.ezproxy.library.ualberta.ca/stable/23359419

Meisuo, Z. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal*, 31(1), 61-95. Retrieved from https://doi.org/10.1177/003368820003100104

Meléndez Quero, C. (2010). Comment expliquer la signification des particules discursives d'une langue étrangère? *Cahiers de l'APLIUT*, 29(1), 137-151. Retrieved from http://journals.openedition.org/apliut/994

Montolío Durán, E. (1998). La Teoría de la Relevancia y el estudio de los marcadores discursivos, In Martín Zorraquino, M. A., & Montolío Durán, E. (Eds.), *Los marcadores del discurso. Teoría y análisis* (pp. 93-119). Madrid: Arco Libros.

Montolío Durán, E. (1993). Si me lo permiten… gramática y pragmática; sobre algunas estructuras condicionales regulativas en español. In Haverake, H. et al (Eds), *Aproximaciones pragmáticas al español Diálogos Hispánicos* (pp. 119-147), 12, Amsterdam: Rodopi.

Muñoz, F., & Del Mar Ruiz Domínguez, M. (2016). Los operadores discursivos de concreción o especificación y de refuerzo argumentativo en el Corpus de aprendices de español como lengua extranjera. *RAEL: Revista Electrónica de Lingüística Aplicada,* 15(1). Retrieved from http://www.aesla.org.es/ojs/index.php/RAEL/article/view/274

Nogueira da Silva, A. M. (2011). La enseñanza de los marcadores del discurso del español en relación con los géneros y secuencias textuales. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 9, 1-15.

Nogueira da Silva, A. M. (2012). Los marcadores del discurso y su introducción en los manuales de E/LE. *Philologica Urcitana*, 7, 75-95.

Plonsky, L. (2015). Advancing Quantitative Methods in Second Language Research. New York, NY: Routledge.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878-912.

Pons, S. (1998). Reformulación y reformuladores. A propósito del libro 'Les opérations de reformulation,' *Oralia*, 1, 183-198.

Portolés, J. (1998). *Marcadores del discurso*. Barcelona: Editorial Ariel, 1998.

Qin, W., & Uccelli, P. (2016). Same language, different functions: A cross-genre analysis of Chinese EFL learners' writing performance. *Journal of Second Language Writing*, 33, 3-17. Retrieved from https://doi.org/10.1016/j.jslw.2016.06.001

Rahimi, M. (2011). Discourse markers in argumentative and expository writing of Iranian EFL learners. *World Journal of English Language*, 1(2), 68. Retrieved from https://doi.org/10.5430/wjel.v1n2p68

Richards, J. C., & Schmidt, R. (1992). *Longman dictionary of language teaching & applied linguistics.* London: Longman.

Roever, C., & Phakiti, A. (2017). *Quantitative methods for second language research: a problem-solving approach.* Routledge.

Yang, W. X., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education*, 23(1), 31-48. Retrieved from https://doi.org/10.1016/j.linged.2011.09.004

Appendix 1. Task Descriptions in the Course Syllabus

Beginners I

Salud y hábitos (unidad 5)

You will have to fill out a survey that the blog TU SALUD has prepared to find out about your habits. Write your answers in complete sentences, and in paragraph format for questions B-E. Make sure you follow the instructions given for each question, and that you meet the minimum word length required where appropriate.

This task will be completed in class individually. It is worth 4% and it will be assessed based on content, grammar, vocabulary and mechanics (spelling and punctuation).

Beginners II

Habitación en alquiler (unidad 11)

Imagine that you are planning to rent one of the rooms in your home and you have to prepare an ad. First, you have to design a floorplan of the house or apartment. You can use http://www.floorplanner.com/ (or http://es.floorplanner.com/ if you'd like to try it in Spanish). Alternatively, you can draw the floorplan by hand (as the ones on p. 122).

Then, in class, you will write a description of the house/apartment and the room you are renting out. In your description, you have to include information about the type of home (house, apartment…), the location, the neighbourhood, furniture and appliances, etc. Be sure to use vocabulary pertaining to the unit (furniture, prepositions to describe where things are placed). Print out your floor plan and turn it in together with the description.

The minimum word count of the text is 150 words. This written description is to be done individually in class, and it is worth 4% of the grade of the course. It will be assessed in terms of content, grammar, vocabulary, and mechanics.

Intermediate II

Una instancia (ver p. 87 del libro)

Una instancia o solicitud es un documento que emplean las empresas o particulares para solicitar algo de un organismo público (ayuntamientos, gobierno…). Consta de tres partes:

1) Introducción: contiene los datos personales del solicitante (nombre, apellidos, dirección); pueden aparecer otros datos como la fecha y el lugar de nacimiento, o el teléfono.

2) Comunicación: se redacta en tercera persona y consta de "exposición" y "petición." La exposición se presenta encabezada por la palabra EXPONE en mayúsculas, en un lugar destacado y seguida de dos puntos. A continuación, se explican de forma ordenada y completa los motivos, datos y argumentos en que se basa la solicitud. Cada uno de ellos debe ir en un párrafo iniciado por la palabra QUE. La petición aparece encabezada por la palabra SOLICITA, también escrita en mayúsculas, situada en un lugar destacado y seguida de dos puntos. En este apartado es donde se concreta de forma clara y concisa lo que se solicita.

3) El cierre (aunque en la canción de Serrat no se contempla) está formado por los siguientes elementos: lugar y fecha de la solicitud, firma del solicitante y pie, que se sitúa en la parte inferior, ocupando todo el ancho de la hoja, y en él se escribe con mayúsculas el nombre del organismo o dependencia al cual se dirige la solicitud, o bien el tratamiento y cargo del destinatario. Además, suele añadirse la localidad.

Individualmente y en clase, van a escribir una instancia solicitando algo que les parezca importante. El texto debe ser de 250 palabras como mínimo y tiene que incluir las tres partes. Esta tarea vale un 4% de la nota y se evaluará el contenido, la organización, la gramática, el vocabulario, y la puntuación y ortografía.

Intermediate II

El invento más importante de la historia

Individualmente, tienes que escoger un invento o descubrimiento que, en tu opinión, es el más importante de la historia de la humanidad. Vas a escribir un texto en el que tienes que describir el invento (características, utilidad, ventajas…) y también explicar por qué para ti es el invento más importante de la historia, haciendo referencia a cómo era la vida antes de su aparición y qué problemas solucionó.

El texto se va a escribir en clase y debe tener 200 palabras como mínimo. Esta tarea vale un 4% de la nota y se evaluará el contenido, la gramática, el vocabulario, y la ortografía y puntuación.

Advanced

The composition was not listed in the course outline, as it was counted as part of homework rather than as a separate component.

Appendix 2. Task Instructions Given to Learners

Beginners I

A. ¿Qué actividades o deportes haces en tu tiempo libre? ¿Con qué frecuencia haces estas actividades o deportes? ¿Cómo ayudan a tu salud y/o equilibrio anímico? Si no practicas ningún deporte explica por qué. (Name at least three activities and two sports; explain when, where and with whom you practice them; 50 words)

B. ¿Con qué frecuencia ves a tu familia y a tus amigos? ¿Qué actividades hacen juntos? ¿Cómo ayudan a tu salud y/o equilibrio anímico? (30 words)

C. ¿Tienes hábitos malos para la salud? ¿Cuáles? ¿Qué aspectos de tu estilo de vida puedes cambiar? (30 words)

D. ¿Tienes recomendaciones para una vida saludable? Por favor escribe 2 recomendaciones para estar en forma físicamente, y 2 para la salud mental y explica por qué. (70 words)

Beginners II

Habitación en alquiler

Imagine that you are planning to rent a room for next term in Madrid. Based on the ads presented below, you have to choose a room; then, you have to explain why you chose that option. You have to include information about the type of home (house, apartment…), the location, the neighborhood, furniture and appliances, etc.

The minimum word count of the text is 150 words.

Image removed after thesis defence	**VIVIR CON UNA FAMILIA ESPAÑOLA** **¿Quieres conocer de cerca la cultura española?** • Habitaciones en casas de familias • Barrio familiar • Pensión completa, lavado de ropa semanal • Las habitaciones tienen cama individual, un armario, un escritorio, una silla y un microondas • La sala de estar se comparte con la familia. • Tamaño mínimo de las habitaciones 4.5m x 3.65m • Precio por semana: 210 €

Image removed after thesis defence	**VIVIR EN UNA RESIDENCIA DE ESTUDIANTES** **¿Te gustaría conocer a estudiantes de otros países?** • Desayuno y cena incluidos, wifi, servicio diario de limpieza • Barrio universitario: hay cafés, bares y discotecas en las cercanías • Línea de metro directa a la escuela (20 minutos) • Cocina compartida con los otros residentes • Sala de estar con televisión con cable • Habitaciones de 4.5m x 2.5m con cama individual, escritorio y cafetera • Servicio de lavadora y secadora por piso • Mínimo un mes. Precio: 950 €

Image removed after thesis defence	**VIVIR EN UN DEPARTAMENTO COMPARTIDO** **¿Quieres vivir con toda la independencia pero, no solo?** • Pisos céntricos, bien comunicados • Amplias habitaciones: 4.5m x 3.65m aprox. • Cerca hay supermercado, farmacia, cafeterías, peluquería, parques etc... • Zona tranquila y muy agradable. Barrio para gente joven y dinámica. • Cada piso dispone de tres o cuatro dormitorios individuales, cocina y dos baños • Conexión a internet y sala de estar común • Todas nuestras habitaciones tienen cama matrimonial, dos mesitas, un amplio escritorio, silla ergonómica, estantería y un mini bar. • Hay cuarto de lavado en el edificio. • Contrato de cuatro meses • Precio: 475 €/mes

Intermediate I

El invento más importante de la historia

Individualmente, tienes que escoger un invento o descubrimiento que, en tu opinión, es el más importante de la historia de la humanidad. Vas a escribir un texto en el que tienes que describir el invento (características, utilidad, ventajas…) y también explicar por qué para ti es el invento más importante de la historia, haciendo referencia a cómo era la vida antes de su aparición y qué problemas solucionó. El texto debe tener 200 palabras como mínimo.

Intermediate II

Una instancia

Una instancia o solicitud es un documento que emplean las empresas o particulares para solicitar algo de un organismo público (ayuntamientos, gobierno…). Consta de tres partes:

1) Introducción: contiene los datos personales del solicitante (nombre, apellidos, dirección); pueden aparecer otros datos como la fecha y el lugar de nacimiento, o el teléfono.

2) Comunicación: se redacta en tercera persona y consta de "exposición" y "petición." La exposición se presenta encabezada por la palabra EXPONE en mayúsculas, en un lugar destacado y seguida de dos puntos. A continuación, se explican de forma ordenada y completa los motivos, datos y argumentos en que se basa la solicitud. Cada uno de ellos debe ir en un párrafo iniciado por la palabra QUE. La petición aparece encabezada por la palabra SOLICITA, también escrita en mayúsculas, situada en un lugar destacado y seguida de dos puntos. En este apartado es donde se concreta de forma clara y concisa lo que se solicita.

3) El cierre (aunque en la canción de Serrat no se contempla) está formado por los siguientes elementos: lugar y fecha de la solicitud, firma del solicitante y pie, que se sitúa en la parte inferior, ocupando todo el ancho de la hoja, y en él se escribe con mayúsculas el nombre del organismo o dependencia al cual se dirige la solicitud, o bien el tratamiento y cargo del destinatario. Además, suele añadirse la localidad.

Individualmente en computadora durante la hora de clase en el laboratorio XXXX van a escribir una instancia solicitando algo que les parezca importante. El texto debe ser de 250 palabras como mínimo y tiene que incluir las tres partes. Esta tarea vale un 4% de la nota y se evaluará el contenido, la organización, la gramática, el vocabulario, y la puntuación y ortografía.

Advanced

Se ha dicho que los seres humanos estamos sujetos a nuestro destino, que el libre albedrío no existe. ¿Qué tan de acuerdo estas con esta idea? ¿Por qué? (Extensión mínima 250-300 palabras)

SAMPLE: XXX	Excellent to very good command of the language 12-15	Good to average command of the language 8-11	Comprehensible expression 4-7	Limited command of the language 0-3
Content **____/15**	• addressed the topic; • main ideas are comprehensible; • substantive, thorough development of thesis; • length is appropriate	• generally addressed the topic; • main ideas are comprehensible • good thematic development; • length is appropriate	• partially addressed the topic • some ideas do not make sense/ are not comprehensible • poor thematic development • length is adequate, but brief	• did not address the topic; • a number of ideas don't make sense; • insufficient response to the writing prompt; • length is inadequate=too short! or not relevant
	Excellent to very good command of the language 12-15	**Good to average command of the language 8-11**	**Comprehensible expression 4-7**	**Limited command of the language 0-3**
Organization **____/15**	• fluent expression; • clear statement of ideas; • clear organization: logical and cohesive sequencing; • appropriate to the writing prompt (also in format, e.g. a letter requires greeting and conclusion)	• adequate fluency; • main ideas clear but loosely organized; • sequencing logical but incomplete; • appropriate to the writing prompt (also in format, e.g. a letter requires greeting and conclusion)	• low fluency; • ideas not well connected; • logical sequencing and development lacking • only partially appropriate to the writing prompt (also in format, e.g. a letter without greeting or conclusion)	• ideas not communicated; • organization lacking, or not enough to rate
	Excellent to very good command of the language 27-35	**Good to average command of the language 18-26**	**Comprehensible expression 9-17**	**Limited command of the language 0-8**
Grammar **____/35**	• few errors in agreement, number, tense, word order, articles, pronouns, prepositions; • Some errors which do not interfere with comprehension; • accurate use of relatively complex structures; • sentence structure shows variety if possible on this level of writing (e.g. sub- and coordinating sentences, not only S-V-O structure; use of transitions)	• some errors in agreement, number, tense, word order, articles, pronouns, prepositions simple; • errors mostly do not interfere with comprehension; • constructions used effectively; • some problems in use of complex constructions; • some variety of sentence structures if possible on this level of writing (e.g. sub- and coordinating sentences, not only S-V-O structure; use of transitions)	• frequent errors in agreement, number, tense, word order, articles, pronouns, prepositions • few errors which interfere with comprehension; • significant problems in use of complex constructions; • Limited variety of sentence structures if possible on this level of writing (e.g. sub- and coordinating sentences, not only S-V-O structure; use of transitions)	• text dominated by errors in agreement, number, tense, word order, articles, pronouns, prepositions; • errors interfere strongly with comprehension; • no mastery of simple sentence construction; • does not communicate, or not enough to rate
	Excellent to very good command of the language 20-25	**Good to average command of the language 14-19**	**Comprehensible expression 7-13**	**Limited command of the language 0-6**
Vocabulary **____/25**	• word choice is appropriate and varied; • mostly accurate word/idiom choice; • effective transmission of meaning	• adequate variety of vocabulary; • some errors of word/idiom choice; • effective transmission of meaning	• Limited variety of vocabulary; • inappropriate use of English (e.g., literal translations, invented words); • frequent word/idiom errors, inappropriate choice, usage; • meaning partially effectively communicated	• little knowledge of target language vocabulary, or not enough to rate; • translation-based errors; • meaning not effectively communicated
	Excellent to very good command of the language 9-10	**Good to average command of the language 6-8**	**Comprehensible expression 3-5**	**Limited command of the language 0-2**
Mechanics **____/10**	• masters conventions of spelling, punctuation, capitalization, paragraph indentation, etc.	• Occasional errors in spelling, punctuation, capitalization, paragraph indentation, etc.	• noticeable spelling and punctuation errors; • frequent capitalization, paragraphing errors; • few times meaning is disrupted by formal problems	• no mastery of conventions due to frequency of mechanical errors, or not enough to rate; • meaning disrupted by formal problems
Total Grade **____/100**	**Comments:**			

Appendix 4. Raw Quality Scores

ID	QUALITY SCORE	ID	QUALITY SCORE	ID	QUALITY SCORE	ID	QUALITY SCORE	ID	QUALITY SCORE
Beginners I		*Beginners II*		*Intermediate I*		*Intermediate II*		*Advanced*	
2A	76	1B	78	1C	45	1D	50	1E	76
3A	80	2B	72	2C	66	2D	60	2E	88
4A	73	3B	84	3C	87	3D	89	3E	75
5A	42	4B	81	4C	93	4D	71	4E	69
6A	55	5B	69	5C	77	5D	98		
7A	50	6B	50	6C	66	6D	87		
8A	71	7B	81	7C	77	7D	66		
9A	70	8B	58	8C	61	8D	83		
13A	75			9C	56				
14A	60			10C	89				
16A	54			11C	53				
18A	44			12C	53				
21A	94			13C	47				
22A	74			14C	49				
23A	51			15C	77				
26A	84			16C	63				
28A	85			17C	37				
30A	77			18C	62				
31A	53			19C	44				
32A	65			20C	56				
40A	56			21C	62				
				22C	61				
				23C	80				

A=Beginners, I B=Beginners II , C= Intermediate I, D=Intermediate II, E=Advanced

Appendix 5. Background Questionnaire

Demographics

1. Age
2. Gender
3. What is your mother tongue?
4. What is your mother's first language?
5. What is your father's first language?
6. What language (s) is/are spoken in your home?
7. Are there other languages that you understand?
8. Are there other languages that you can speak?
9. Are there other languages that you can write in?
10. Are there other languages that you studied before?
11. Have you studied or lived abroad?
    a. For how long?
    b. How old were you?
    c. Where did you go?

Practices

Please choose the option that best represents what you usually do.

Always	Frequently	Sometimes	Seldom	Never
○	○	○	○	○

1. I _____ use dictionaries when I write in Spanish.
2. I _____ use English - Spanish dictionaries when I write in Spanish.
3. I _____ use Spanish - Spanish dictionaries when I write in Spanish.
4. I _____ use translators when I write in Spanish.
5. I _____ use a Spanish grammar when I write in Spanish.
6. I _____ use a spell checker when I write in Spanish.
7. I _____ use webpages that help with grammar issues when I write in Spanish.
8. I _____ think in English when I write in Spanish.
9. I _____ think in Spanish when I write in Spanish.
10. I _____ write in a personal glossary the new words I learn in Spanish.
11. I _____ read in Spanish if it is not assigned as a task or homework.
12. I _____ ask a friend that knows more Spanish for comments on my writing