



Máster Universitario en Profesorado de Educación Secundaria  
por la Universidad Pública de Navarra

**THE EFFECTS OF USING COLLABORATIVE WRITING  
VS. PEER REVIEW TREATMENTS ON SUBSEQUENT  
INDIVIDUAL WRITINGS**

Autora: Seila Rodero Albaiceta

Tutora: M<sup>a</sup> Camino Bueno Alastuey

TRABAJO FINAL DE MÁSTER CURSO 2017/2018

ESPECIALIDAD INGLÉS

Pamplona, junio 2018

## Table of contents

Abstract.....	2
1. Introduction .....	2
2. Literature Review .....	3
2.1 Collaborative writing .....	3
2.2 Peer review.....	6
3. Research questions and expected results .....	10
4. Methodology.....	11
4.1 Participants.....	11
4.2 Instruments and materials .....	12
4.3 Procedure .....	12
4.4 Data analysis .....	14
Analysis of the written texts: Complexity, Accuracy and Fluency (CAF) .....	14
Holistic analysis .....	17
5. Results .....	17
5.1 Improvement of EFL students' writing skills .....	18
5.1.1 Global quantitative data analysis.....	18
5.1.2 Holistic data analysis.....	19
5.2 Differences in subsequent individual writings.....	20
5.2.1 Quantitative data analysis.....	20
6. Discussion.....	26
7. Conclusions .....	30
8. Acknowledgments .....	31
9. References .....	32
Appendix I.....	38
Appendix II.....	39
Appendix III .....	41
Appendix IV .....	42
Appendix V.....	43

## Abstract

Previous researches have investigated separately the effects of collaborative writing and peer review. However, to the best of the author's knowledge, there has not been any research comparing both approaches. This study is aimed to analyse the effects of two treatments, collaborative writing and peer review on individual writing, on a subsequent individually written production in terms of fluency, lexical variety, accuracy and complexity. A total of 29 students of lower-intermediate level took part in this study, 16 in the collaborative writing group and 13 in the peer review group. The students wrote a total of 3 compositions using a news article format from which the pre-test and the post-test writings were analysed following a quantitative and a holistic analysis. The results suggest that students who had received collaborative writing as a treatment produced longer writings and used more complex language, while those who carried out peer review in the treatment session improved their final individual writing in terms of lexical variety and accuracy. The implications of these results are discussed.

**Key words:** collaborative writing, peer review, fluency, lexical variety, accuracy, complexity, EFL.

## 1. Introduction

To be proficient in a language, learners have to be competent in the four skills: reading, speaking, listening and writing, which are interrelated since the use of any language normally requires using more than one skill simultaneously. According to constructivist theories, foreign language learners build their own learning through experience, interaction and reflection as they constantly interact with new educational situations (Vygotsky, 1978, 1979).

Writing is one of the main language skills, specifically, a skill that requires production which is an effort to create language. The same as when speaking, writing is used to share ideas, feelings, to convince others, etc. It is considered one of the most arduous tasks due to the mental exertion it requires. When second language learners are writing, they are forced to notice certain grammar and vocabulary structures and select which ones to use from the collection of structures and lexis available. Writing requires

appropriate language use, text construction, lay out, style and effectiveness (Harmer, 2007).

In 1998, Davis stated that “writing is essentially a creative process and good writers must learn to communicate their ideas clearly to an unseen audience” (p.25). The benefits of peer-review are also well-known since this approach enhances learners’ thinking skills when assessing their classmates.

Researchers applying sociocultural theory to the study of L2 learning maintain that learners can have a positive impact on each other’s development because they can act as both novices and experts (Ohta, 2000, 2001; Storch, 2002; Swain & Lapkin, 1998). Because no two learners have the same strengths and weaknesses, when working together, they can provide scaffolded assistance to each other and, by pooling their different resources, achieve a level of performance that is beyond their individual level of competence (Ohta, 2001). There have been studies analysing different types of scaffolded assistance, and the most common types used have been collaborative writing and peer review. Nevertheless, to the best of my knowledge, there are no studies comparing the effect of implementing those two approaches to analyse their effect on subsequent writings.

## 2. Literature Review

### 2.1 Collaborative writing

As described by Swain (2001), collaborative tasks are communicative tasks in the sense that they involve “learners in comprehending, manipulating, producing, or interacting in the target language while their attention is principally focused on meaning rather than form” (Nunan, 1989:10). But collaborative tasks also require that learners work together, sharing ideas and pooling their knowledge, to achieve one common goal (Fernández Dobao, 2012).

As it has been mentioned before, collaborative tasks demand communication, and when communicating, there are moments when two students discuss about the language they are going to use or they have just used. Those moments when learners collaborate to solve grammatical and lexical misunderstandings, or correct each other are called Language-related episodes (LRE) (Swain & Lapkin, 1998).

Many studies about collaborative writing have been carried out in the last decades investigating the benefits of collaborative writing and comparing it to individual tasks. Most of these researches have been conducted from a cognitive perspective, focusing on the effect of task on L2 production—measured in terms of accuracy, fluency, and syntactic complexity (see, for instance, Bygate, Skehan, & Swain, 2001; Ellis, 2003; García Mayo, 2007).

For example, Storch (1999) analysed the impact of collaboration in different kinds of written exercises. In her study, those who were working in pairs took longer to finish and their production was shorter, but more accurate than those who worked individually. In another study from the same author (2005), other aspects such as the effects of collaborative oral interaction were analysed. The results showed that pair work allowed students an opportunity to collaborate on the writing process, to share and to improve their production by providing each other with immediate feedback on the language being used. In subsequent larger-scale studies (Storch & Wigglesworth, 2007), similar results were presented. No differences were found in terms of fluency and complexity, but the texts written in pairs were significantly more accurate than those written individually. The authors concluded that due to the LREs that had taken place, these collaborative writings tended to have more accurate language.

A 16-week quasi-experimental study conducted by Shehadeh (2011) in an EFL context aimed to explore the effectiveness of collaborative writing. Twenty students were asked to write individually while 18 students wrote in pairs. Writings were analysed in terms of content, organization, grammar, vocabulary and mechanics, using a rubric with ratings. The results showed that collaborative activities done over a prolonged period of time improved students' writings in content, grammar and vocabulary, even with students at low proficiency levels. Shehadeh (2011) pointed out that “the results of the study showed that collaborative writing had an overall significant effect on students’ L2 writing” (p.286). In fact, he found out that working collaboratively over a prolonged period of time enhanced the quality of students' writings.

In another study, Nassaji and Tian (2010) compared individual and collaborative work using two different tasks, a cloze task and an editing task, in an English as a L2

context. Learners working in pairs completed the tasks more accurately than learners working alone. However, the results of the vocabulary pre- and post-tests did not provide clear evidence of greater knowledge gains for the collaborative condition.

Fernández Dobao (2012) was the first study to compare group, pair and individual work in collaborative writing tasks in the L2 classroom. The researcher examined whether the number of participants had an effect on the fluency, complexity and accuracy of the written products and on the frequency and nature of the oral interaction produced in pairs and groups. Her findings showed that groups produced more LREs than pairs and also a higher number of correctly solved LREs. Consequently, texts written in groups were not only more accurate than those written individually, but they were also more accurate than those written in pairs. Although group work offered learners fewer opportunities to actively participate in the conversation, it led to better and more accurate results.

More recently, Fernández Dobao (2014) focused on vocabulary learning in collaborative writing tasks and compared pair and small group work. The participants were those in the 2012 study. Results showed that small groups produced more lexical LREs than pairs and that more LREs were correctly solved. It was also found that although learners had fewer opportunities to contribute to the conversation when working in small groups, there did not seem to be a negative effect on the learners' rate of retention of the lexical knowledge which was co-constructed in conversation. Learners seemed to benefit from the LREs when they were actively involved in the conversation, as well as when they were acting as observers. Therefore, Fernández Dobao concluded that small group interaction led to significantly more opportunities for L2 vocabulary learning than pair interaction.

Besides analysing the writing processes and the final product, some researchers have aimed to explore students' perceptions about collaborative writing. In the study by Shehadeh (2011), most of the participants expressed a positive attitude toward collaborative writing and they enjoyed it, although they were novices at such an experience. They also acknowledged that writing collaboratively was beneficial as it provided them with opportunities to discuss and plan their writing, to generate ideas, to create texts, to give immediate peer feedback and to polish up their texts. Furthermore,

the students reported that collaborative writing had helped them have self-confidence and had improved their speaking abilities. A similar impression was reported by the participants in the study by Fernández Dobao (2013), the students working collaboratively enjoyed the experience and had an overall positive attitude towards this type of methodology.

In a previously mentioned study carried out by Storch (2005), the students were interviewed after the study and 16 out of 18 from those who had worked in pairs also expressed very positive attitudes to writing in pairs as it helped them to pool their resources, observe each other and learn how to express the same ideas in different ways. However, it is worth highlighting that there were five students among those who expressed a positive attitude who expressed some reservation because they didn't feel comfortable enough with their language skills and were concerned about criticising others. In Storch's study (2005), two students said that writing is an individual activity and that paired work is good for oral activities.

From the results of the studies conducted so far, it can be concluded that collaboration has a positive effect on task performance, and that small group work points to better and more accurate results compared to pair and individual work.

## 2.2 Peer review

One of the intended aims of education is to enhance learners' thinking skills. Peer review, also known as peer response, peer feedback or peer assessment, is defined as a collaborative activity in which students read, criticize and give feedback on each other's writings to improve writing competence through mutual scaffolding (Hu, 2005; Tsui & Ng, 2000; Zhu, 2001). Therefore, it makes students independent in the judgments of their own work and that of their colleagues. Peer assessment is based on the assumption that students can learn as much from each other as they can from the teacher (Ashraf & Mahdinezhad, 2015).

For peer review to be successful, schemes and the list of criteria that will be used to assess the final products have a great importance, together with how the process is set-up and managed. Several authors have provided guidelines for the management of peer-assessment (Race 1999; Magin & Helmore 2001; Stefani 1994).

There are a growing number of pedagogical and practical arguments that have been advanced to support peer assessment for students in higher education (Falchikov 1995; Magin & Helmore 2001) mainly because it emphasizes learner's autonomy and cooperation.

The Vygotskian theory (1962, 1978) of language learning firmly supports the use of peer review. For Vygotsky, learning is a cognitive activity that occurs in social interaction and is mediated by it. Therefore, at a theoretical level, peer interaction is vital to language development because it allows students to construct knowledge through social sharing and interaction (Liu, Lin, Chiu, & Yuan, 2001). Consequently, peer review is also built on the notion of collaboration, which assumes that learning emerges through the shared understanding of multiple learners, and that learning effectively occurs within interactive peer groups (Asberg & Nulden, 1999.; Leidner & Jarvenpaa, 1995).

Peer feedback has garnered increasing attention in L2 writing classrooms, as the activity promises to encourage negotiation about and construction of meaning and to help students develop new perspectives on writing (Ferris, 2003; Liu & Hansen, 2002; Liou, 2009). Different studies conducted by Coniam & Lee (2008) and Lin & Yang (2011) have supported the advantages of peer feedback, which has been shown to help students improve their writing quality and to enhance their writing confidence.

Many investigators have argued that in L1 instructional settings, the peer interactions that occur during peer reviews have cognitive benefits because they provide students with opportunities to assume a more active role in their own learning (Barnes, 1976; Brief, 1984; Carl, 1981; Forman & Cazden, 1985). And, as suggested by Mangelsdorf (1989), peer interaction helps L2 students communicate their ideas and can enhance the development of L2 learning in general. Peer interactions during peer reviews, she concludes, give students “more ways to discover and explore ideas, to find the right words to express these ideas, and to negotiate with their audience about these ideas—all of which are critical in second language acquisition and cognitive growth” (p.143).

Many researchers have concluded that the implementation of peer assessment in the curriculum was beneficial for the learning goals. For instance, Cutler and Price



(1995), Freeman (1995), Horgan et al. (1997), and Sluijsmans et al. (2002) reported an increase in the quality of learning due to peer assessment tasks. Moreover, several studies reported that peer assessment tasks exposed students to the skills of critical reflection and analysis (Birenbaum, 1996; Sambell & McDowell, 1998), and they increased students' confidence in their ability to perform according to specified criteria (Cutler & Price, 1995).

In a study carried out by Mendonca and Johnson (1994) in an ESL writing class at a major university in the north of the United States, they examined the negotiation patterns of graduate student learners of English working in pairs, and analysed audio-taped peer review sessions and learners' written drafts. Five types of peer review negotiations were identified: asking questions, giving explanations, making restatements, offering suggestions, and correcting grammar. The analysis showed that during peer review learners focused on both local and global issues in their writings and that after negotiation they appeared to have a better understanding of the strengths and weaknesses of their writings. More important, learners developed audience awareness through peer review activities. The authors concluded that the learners in this study found peer review to be beneficial. In addition, peer review was found to "enhance students' communicative power by encouraging learners to express and negotiate their ideas" (p. 765-766).

However, there are controversial arguments about the true efficacy of peer feedback. On the one hand, peer feedback helps to promote language learner autonomy in process approaches to writing (Ekşi, 2012; Yang, Badger, & Yu, 2006), creates a friendly and secure environment for language learners (Sato, 2013; Yang et al., 2006), and develops learners' writing skills in subsequent writing drafts (Diab, 2010; Lundstrom & Baker, 2009). On the other hand, popular concerns about the true efficacy of peer feedback relate to students' limited knowledge of the language, the trustworthiness of feedback provided by peers on a wide range of errors and students' inappropriate attitude towards peer response (Hu, 2005).

Regarding the above, Cho, Schunn, & Wilson (2006), investigated the validity and reliability of peer review in writing and they demonstrated that the aggregate ratings of at least 4 peers on a piece of writing made the grades both highly reliable and as valid

as instructors' ratings while (paradoxically) producing very low estimates of reliability and validity from the students' perspective. The results suggest that instructor concerns about peer evaluation's reliability and validity should not be a barrier to implementing peer evaluations, at least with appropriate scaffolding.

Soleimani and Rahmanian (2014) studied the impact of self-assessment, peer assessment and teacher's feedback on –exclusively – the CAF writing abilities of their learners. The study was designed to compare the efficacy of self-assessment, peer assessment, and teacher assessment for CAF development. The results were as follows: self-assessment was effective for the short-term development of accuracy and fluency, but its impact declined in the delayed post-test; peer assessment led to significant improvements in complexity, accuracy, and fluency, all in the immediate post-test only; and the teacher assessment group experienced an improvement in fluency level in both post-tests, no significant gain in complexity, and only short-term progress in the accuracy domain.

Richer (1992) compared the effects of peers' feedback and teacher's feedback on college students' writing proficiency. The pre/post measures of students' essays revealed that greater gains in writing proficiency were obtained by the peer feedback group. Ramsden (1992) found that students could often learn more from formal or informal assessment by their peers. A quasi-experimental study by Plutsky and Wilson (2004) also revealed that peer review helped students become proficient writers.

A similar subsequent study from Ghahari & Farokhnia (2017) stated that the results of within-group comparisons revealed that both peer assessment and teacher assessment groups experienced significant improvement in terms of accuracy and fluency. But no significant improvement was observed in the complexity domain in either treatment group. In fact, regarding complexity, the type of feedback might be more important than the source as Sheppard (1992) demonstrated. He investigated the effect of unfocused teacher feedback on the written complexity and accuracy of ESL learners and reported that the group which received holistic comments, where the teacher was more comprehensive and did not only focus on one aspect of the writing but addressed the overall production, outperformed the group that received corrective feedback on accuracy. On the contrary, the corrective feedback group which received

comments based on what students did well and also what they needed to work on in order to improve future work seemed to regress over time by avoiding the use of complex structures.

As can be seen from the aforementioned studies, peer review in writing helps to develop students' thinking skills and makes students become better language learners when reflecting about the language they are using (Schwartz, 1989). Therefore, and taking into consideration those findings, in the present study, the effect of peer review will be analysed and compared to collaborative writing in order to discover the effect of these two approaches on a final individual writing.

### 3. Research questions and expected results

Even though there are many studies that have analysed the effects of collaborative writing and many others about the impact of peer review in students' writings production, to the best of my knowledge, there are not any researches that have compared these two approaches. This is the reason that prompts the present study.

Therefore, this research seeks to analyse the differences between two intact groups in a final individual writing after each of the groups received a different treatment: collaborative writing vs. peer review. After the treatment, subsequent individually written texts were analysed in order to find an answer to the following questions:

- Do collaborative writing and peer review help to improve EFL students' writing skills?
- Are there any differences in the final individual product after having experienced either collaborative writing or peer review in terms of fluency, lexical diversity, accuracy and complexity?

Regarding the first research question, it is expected that both groups of students, those who had worked in collaborative writing and those who had carried out a peer review on previous writings, will improve their subsequent final individual writing.

Therefore, the subsequent individual writings results are expected to be positive but similar, since both methods help students to focus on the language they use. On the one hand, researchers have found that collaborative writing contributes to an increased complexity as well as increased grammatical accuracy and overall quality of writing (Storch, 2005). On the other hand, peer feedback seems to be associated with better student academic achievement.

## 4. Methodology

### 4.1 Participants

The present study was conducted in two lower-intermediate English classes in a semi-private school in Pamplona, Navarre. A total of 36 students (21 females and 15 males) participated in the project; all of them were doing the third year of secondary school. Their ages ranged from 14 to 15. All participants were Spanish native speakers except a student coming from a European country but whose first language was neither Spanish nor English and who had a native-like command of Spanish. All the participants in the study had had 4 hours of English a week during the scholar-year.

In this school in the third year, students are divided into three different groups (A, B, C). According to the teachers, the two classes taking part in this research (A, B) had a similar English level whereas the third group (C) was more proficient in English. As mentioned by Storch (2005), collaborative situations are more likely to happen among students with similar language level. And as the purpose of the research was to compare the effect of collaborative writing vs. peer review, these two intact groups, A and B, were chosen for the study.

A total of 7 students were excluded from the present research. In group A (collaborative writing treatment), 2 students couldn't take part in the study because 1 participant was missing during the treatment intervention and another in the post-test session. In group B (peer review treatment), a total of 5 participants were discarded due to several reasons: 1 student was rejected because of his illegible handwriting, 3 participants were missing in either the pre-test or post-test session and 1 student couldn't attend the treatment session. Hence, group A was formed by 16 students and group B by 13 students which makes a total of 29 participants in the study.

## 4.2 Instruments and materials

Data was collected by means of two different instruments: 1) a pre-test which was a first news article written individually; 2) a post-test which was a third news article written individually. The intervention was carried out along three weeks.

The two pieces of writing (pre- and post-test) were analysed using two different approaches: a) quantitative ratings of fluency, lexical diversity, accuracy and complexity; and b) a holistic rating of fluency, cohesion, adequacy, mechanics and language control using a rubric (see Appendix 5).

Moreover, participants received a treatment session where they were taught the news article format. The researcher also collected the second writings done collaboratively (group A) and done individually and peer reviewed (group B) but they were not analysed.

The rubric group B used to carry out peer review (see Appendix I) was designed taking into consideration different aspects of writing: grammar (tenses, third person singular -s, etc.), the lay-out used for this format, the lexical variety used by the writer, the content etc.

## 4.3 Procedure

On the day of the pre-test, students didn't receive any input regarding a news article format or style. Students were asked to produce a news article based on the instructions provided (see Appendix II). They were given 30 minutes in order to complete the task. Afterwards, the writings were collected.

During the second week, the treatments were carried out. Separately, the students from the different groups received a 30 minutes master class on news articles (see Appendix III) led by the researcher in both cases. In this session, the format of a news article was explained as well as its style. Furthermore, students were told the procedure of the activity. Group A had to write in pairs and group B individually with a subsequent peer review. After that, 60 minutes (two 30 minutes sessions) were assigned for the task, taking into account they had to do some research on the topic.

The students from group A, the collaborative writing group, were set in pairs as heterogeneous as possible as the instructor can facilitate learning by preventing homogenous pairing (Zhu, 2001), so even though students were all at a similar level, the teacher tried to make the pairs as heterogeneous as possible. Group A had 10 more minutes to finish their collaborative writings because pair work has been reported to take longer (Fernández Dobao, 2012; Storch, 2005).

Regarding group B, the peer review group, the students had the assigned writing time (30 minutes) in addition to 15 minutes to carry out peer assessment. The students were provided with a rubric (see Appendix I) for assessing their assigned classmate's work. Each student reviewed someone else's work and graded the different items from 1 to 10, which was the highest mark.

As a post-test, a final individual writing was done (see Appendix IV). The procedure was similar to the pre-test: the students were given instructions and following those instructions, they had to produce their news articles. Once again, they had 30 minutes to complete the task.

Students carried out a total of three written tasks, all of them based on the news article format. This format was chosen because it was part of the syllabus of the third year and it hadn't been worked so far. The topics of the writings were selected taking into account the students' interests and motivations. That could be possible thanks to the teacher who knew her students very well and could advise me regarding the selection of topics.

The pre-test (first writing essay) was a news article about a Spanish teenager who became a famous singer after winning a talent show. For the treatment (second writing essay), the students were asked to write a news article based on an event related to their school, they were allowed to interview people and to search for information on the web. Finally, in the post-test (third writing essay) the students were told to write a news article about the school's day.

## 4.4 Data analysis

### Analysis of the written texts: Complexity, Accuracy and Fluency (CAF)

In the present study, following similar previous research (Fernández Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009), only the pre-test and post-test written texts were analysed for accuracy, fluency, and syntactic complexity. To answer the research questions, the four parameters used were based on the concept of T-unit which was first developed by Hunt (1965) as a measure of syntactic maturity in the writing of schoolchildren. Hunt (1970) defined it as “a main clause plus all subordinate clauses and nonclausal structures attached to or embedded in it” (p.189). Nowadays, T-units are usually used to analyse written and spoken discourse because it has been proven that the number of T-units are strongly correlated to language proficiency (Ellis & Barkhuizen, 2005; Wolfe-Quintero, Inagaki, & Kim, 1998).

The following analysis was carried out:

1) **Accuracy** was measured as the proportion of error-free T-units to total T-units in terms of **lexical**, **grammatical** and **mechanical** errors. Most previous research has focused on grammatical and lexical errors, ignoring spelling and punctuation problems. However, since the study of LREs has found that learners working collaboratively discuss mechanical as well as grammatical and vocabulary problems (e.g., Storch, 2007, 2008; Storch & Wigglesworth, 2007, 2010; Wigglesworth & Storch, 2009), in the present study all three types of errors have been identified.

Three measures of accuracy were used:

- a) **The ratios of error-free T-units to total T-units.** This measure consisted in counting all T-units and excluding all T-units with errors. Then, the error-free T-units were counted and divided by to total amount of T-units.
- b) **The ratios of error-free clauses to total clauses.** This measure consisted in counting all clauses and excluding all clauses with errors. Then, the error-free clauses were counted and divided by to total amount of clauses.

c) **Errors to words.** All the words were counted and analysed in order to detect any error, if so, excluded. Then, the error-free words were counted and divided by the total amount of words.

These three measures of accuracy were selected in order to make the results comparable to those of previous research (e.g., Fernández Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009).

The following extracts are examples of the type of errors. Although in some examples there is more than one type of error, the error referred to appears underlined.

(i) **Grammatical errors** include syntactical errors (missing elements and errors in word order) and morphological errors (errors in subject-verb agreement, errors in the use of articles and prepositions and verb tenses).

Example 1. **Missing elements:**

S3a: *First the pre-school students  $\emptyset$  at the chapel, act done by Natalia.* [First, the pre-school students went to the chapel and the act was done by Natalia.]

Example 2: **Errors in word order:**

S9b: *When she go out she was the spain queen.* [When she got out she was the queen of Spain.]

Example 3: **Errors in subject-verb agreement:**

S6a: *They was celebrating FEC's day.* [They were celebrating FEC's day.]

Example 4: **Errors in use of articles:**

S8b: *She will win the Eurovision.* [She will win Eurovision]

Example 5: **Errors in use of preposition:**

S13b: *In Friday 13th of April...* [On Friday 13th of April...]

Example 6: **Errors in use of verb tense:**



S7a: *This girl has make that all Pamplona are in love with her...* [This girl has made all Pamplona fall in love with her.]

(ii) **Lexical errors** include confusion of word choice (words from other languages or borrowings).

Example 7. **Transfers:**

S5a: *She release a new album and alcanza more than 3000 tweets.* [She released a new album and reached more than 3000 tweets.]

Example 8. **Borrowings:**

S4b: *Her actual boyfriend and the other concursants.* [Her current boyfriend and other contestants.]

(iii) **Mechanical errors** include spelling, punctuation and capitalization.

Example 9. **Spelling errors:**

S1a: *The people who was whit...* [The people who were with Amaia...]

Example 10. **Punctuation errors:**

S11a: *...the teacher from Vedruna school  $\emptyset$  in this day it has a three different events...* [...the teachers from Vedruna school. That day, there were three different events...]

Example 11. **Capitalization:**

S11a: *The singer that will represent spain in Eurovision.* [The singer that will represent Spain in Eurovision.]

**2) Complexity:** Clauses and T-units were identified, and three different measures of syntactic complexity were calculated:

**a) Number of words per clause.** Every clause was identified and all the words within the clause were counted. Then, to find the mean of the composition, the

total number of words was added per clause and divided by the number of clauses.

**b) Number of words per T-unit.** Every T-unit was identified and all the words in them were counted. Then, to find the mean of the composition, the total number of words per T-unit was added and divided by the number of T-units.

**c) Number of clauses per T-unit.** Every T-unit has were identified and all the clauses in them were counted. Then, to find the mean of the composition, the total number of clauses per T-unit was added and divided by the number of T-units.

**3) Fluency** was measured by the total number of words produced. Besides that, and differently from previously mentioned studies (e.g., Storch, 2007, 2008; Storch & Wigglesworth, 2007, 2010; Wigglesworth & Storch, 2009), lexical diversity was also calculated by counting the number of different words divided by the number of total words as considering only number of words as the only parameter to measure fluency was considered insufficient.

#### Holistic analysis

Besides the quantitative analysis, a holistic analysis was carried out in order to have a qualitative perception of the texts. The researcher designed a holistic rubric to assess the writing of the participants covering four points scale (see Appendix V) being 1 the lowest mark and 4 the highest. The five items were fluency, considered as the number of words in the text and lexical variety; cohesion, which evaluated the development of ideas; adequacy, which analysed whether the objective was fulfilled, the appropriateness of the length of the text and its organization in terms of task completion; mechanics, that assessed the spelling, punctuation and capitalization; and language control, which measured the use of agreement, number, tense, word order, pronouns, articles, prepositions and negation.

## 5. Results

The following section is aimed to present the results of the analyses conducted in relation to the two research questions that prompted the present study. (1) Does collaborative writing and peer review help to improve EFL students' writing skills? And

(2) are there any differences in the individual final product between the collaborative writing treatment group and the peer review treatment group regarding fluency, lexical diversity, accuracy and complexity?

## 5.1 Improvement of EFL students' writing skills

### 5.1.1 Global quantitative data analysis

The first research question aimed to investigate, on the one hand, if participating in collaborative writing helps to improve EFL students' individual writing skills and, on the other hand, if doing and receiving peer review on individual writing improves subsequent students' writing productions.

As can be seen in Table 1 and according to the parameters analysed, the mean scores indicate that students participating in collaborative writing improved their fluency and complexity, but not their lexical diversity and accuracy as their mean scores decreased in these two aspects.

*Table 1. General results collaborative writing group.*

		<b>COLLABORATIVE WRITING</b>	
		<b>Pre-test</b>	<b>Post-test</b>
<b>MEAN</b>	FLUENCY	<b>92,25</b>	<b>106</b>
	LEXICAL DIVERSITY	<b>0,75</b>	<b>0,67</b>
	ACCURACY	<b>0,76</b>	<b>0,73</b>
	COMPLEXITY	<b>10,79</b>	<b>11,28</b>

As shown in Table 2, as regards those students who carried out and received peer review on their writings, they did not seem to have improved their overall results as their mean scores show that they obtained better results in the pre-test than in the post-test except in accuracy where the mean score slightly increased.

**Table 2.** General results individual peer reviewed writing group.

		PEER REVIEW	
		Pre-test	Post-test
MEAN	FLUENCY	108	91,17
	LEXICAL DIVERSITY	0,74	0,68
	ACCURACY	0,74	0,75
	COMPLEXITY	10,41	9,40

Therefore, the global results suggest that the parameters where students improved were different from one group to another. In group A, the collaborative writing treatment group, the students obtained better global results in fluency and complexity, whereas group B, the peer review treatment group, improved in the global measures of accuracy.

### 5.1.2 Holistic data analysis

**Table 3.** Results of holistic measures.

HOLISTIC ANALYSIS	COLLABORATIVE WRITING		PEER REVIEW	
	Pre-test	Post-test	Pre-test	Post-test
	MEAN	6,47	7,03	6,79

The holistic analysis considered the content, organization, language use, vocabulary and mechanics. The results obtained from the correction done following the different aspects included in the holistic rubric show that both groups improved their mean in the post-test (see Table 3). Even though students in the collaborative writing treatment group obtained lower marks, the analysis of the data suggests that this group did slightly better in the post-test increasing their mean from 6,47 to 7,03 which makes a difference of 0,56 points, while the peer review group increased their mean from 6,79

to 7,29 which makes an improvement of 0,50 points, slightly lower than the other group.

## 5.2 Differences in subsequent individual writings

The second research question sought to examine the differences between students who had carried out collaborative writing and peer review regarding fluency, lexical diversity, accuracy and syntactic complexity.

### 5.2.1 Quantitative data analysis

#### 5.2.1.1 Fluency and lexical diversity

To examine fluency and lexical diversity, the total amount of words produced were counted and analysed. As can be seen in Table 4, students who had been writing collaboratively in the treatment obtained a mean of 92,25 words in the pre-test and 106 words in the post tests, which implies a positive difference of 13,75 points. Meanwhile, students who had been writing individually and then carrying out peer review obtained a mean score of 108 words in the pre-test and a mean of 91,17 words in the post test. Consequently, this group obtained a negative difference of 16,83 points.

Therefore, and contrary to our expectations, the fluency of the individually written texts was lower in the participants who had only worked individually and had peer reviewed their classmates' essays. The results obtained from this fluency measure suggest that the effect of writing collaboratively was greater than the effect of peer review.

**Table 4.** Results of the fluency and lexical diversity measures.

		COLLABORATIVE WRITING		PEER REVIEW	
		Pre-test	Post-test	Pre-test	Post-test
MEAN	TOTAL WORDS	92,25	106,00	108	91,17
	LEXICAL DIVERSITY	0,74	0,67	0,74	0,68

Lexical diversity was measured calculating the total number of different words by the total number of words. The mean scores, as can be seen in Table 4, reveal that there is almost no difference between the students who had carried out peer review and the collaborative writers with regards to lexical diversity. Regarding the ones who had done collaborative writing, students obtained a mean of 0,74 in the pre-test and 0,67 in the post test which makes a negative difference of 0,07 points. Something similar happened to the students in the second group, those who had written individually and then carried out a peer review scored 0,74 in the pre-test and 0,68 in the post-test implying a negative difference of 0,06.

This finding seems to point out that neither carrying out collaborative writing nor peer assessment had an important effect on the lexical diversity of subsequent L2 learners' productions as neither group improved their lexical variety in the post-test.

#### 5.2.1.2 Accuracy

In Table 5 we can see the global mean for accuracy in the pre-test and post-test in both groups.

*Table 5. Results of the global accuracy measures.*

<b>ACCURACY</b>			
<b>COLLABORATIVE WRITING</b>		<b>PEER REVIEW</b>	
<b>Pre-test</b>	<b>Post-test</b>	<b>Pre-test</b>	<b>Post-test</b>
<b>0,76</b>	<b>0,73</b>	<b>0,74</b>	<b>0,75</b>

As can be seen in Table 4, the collaborative writing group decreased their mean from 0,76 to 0,73 which makes a negative difference of 0,03 points while the peer review group improved their global mean from 0,74 to 0,75 which makes a slight improvement of 0,01 points.

Thus, regarding global accuracy measure the data suggests that those students who had worked on peer review obtained better results while those who had worked on collaborative writing does not seem to improve their productions.

As in the previous section, group comparisons were performed in order to investigate accuracy. Three parameters (grammar, lexical and mechanical) were considered on the scores of the three accuracy measures (T-units, clauses and words).

**Table 6.** *Results of grammar accuracy measures.*

<b>GRAMMAR ACCURACY</b>					
		<b>COLLABORATIVE WRITING</b>		<b>PEER REVIEW</b>	
		<b>Pre-test</b>	<b>Post-test</b>	<b>Pre-test</b>	<b>Post-test</b>
<b>MEAN</b>	<b>ERROR FREE T-UNITS / TOTAL T-UNITS</b>	<b>0,28</b>	<b>0,23</b>	<b>0,33</b>	<b>0,36</b>
	<b>ERROR FREE CLAUSES / TOTAL CLAUSES</b>	<b>0,53</b>	<b>0,47</b>	<b>0,45</b>	<b>0,43</b>
	<b>ERRORS / WORDS</b>	<b>0,95</b>	<b>0,89</b>	<b>0,95</b>	<b>0,88</b>

As can be seen in Table 6, in error free T-units to total amount of T-units, the group who had carried out collaborative writing obtained in the pre-test 0,28 and in the post-test 0,23 which makes a negative difference of 0,05 points meaning that in the post-test the writers did worse in this measure, while the group who had carried out peer review improved their productions from 0,33 to 0,36 which makes an increment of 0,03 over the pre-test. With regard to the second accuracy measure (error free clauses), both groups obtained worse results in the post-test, the group who had carried out peer review went from 0,45 to 0,43 which makes a negative difference of 0,02 points, while the group carrying out collaborative writing went from 0,53 in the pre-test to 0,47 in the post test, obtaining a negative difference of 0,06 points. This implies that, although neither group improved their scores, the peer review group obtained better results than the collaborative writing group.

Similarly, as Table 6 indicates, the mean score of the third accuracy measure (errors to words) shows that the participants in the collaborative group obtained a mean of 0,95 in the pre-test and 0,89 in the post test, so there was an improvement of 0,06 as having a lower score indicates fewer errors and, thus, improved performance. In the case of peer review, students also performed better in the post-test, obtaining a positive difference of 0,07 from an initial score of 0,95 to a final score of 0,88. Therefore, regarding errors to words, both groups obtained similar results improving their mean in the post-test.

Thus, the peer review group improved in two of the components of the grammar accuracy, error free T-units to total T-units and errors to words. While the collaborative writing group only improved in one of the components, errors to words. Consequently, the peer review group obtained better results in grammar accuracy.

*Table 7. Results of lexical accuracy measures.*

<b>LEXICAL ACCURACY</b>					
		<b>COLLABORATIVE WRITING</b>		<b>PEER REVIEW</b>	
		<b>Pre-test</b>	<b>Post-test</b>	<b>Pre-test</b>	<b>Post-test</b>
<b>MEAN</b>	<b>ERROR FREE T-UNITS / TOTAL T-UNITS</b>	<b>0,71</b>	<b>0,76</b>	<b>0,73</b>	<b>0,84</b>
	<b>ERROR FREE CLAUSES / TOTAL CLAUSES</b>	<b>0,86</b>	<b>0,90</b>	<b>0,83</b>	<b>0,93</b>
	<b>ERRORS / WORDS</b>	<b>0,99</b>	<b>0,98</b>	<b>0,99</b>	<b>0,98</b>

Table 7 reports the results of the lexical accuracy analysis. The first measure, error free T-Units, presents improvements in the peer review group since in the pre-test the students obtained 0,73 and in the post-test 0,84 which implies an improvement of 0,11 points. In the collaborative writing group, the participants also obtained better results as they improved from 0,71 to 0,76 in the post-test although that improvement (0,05) was smaller than the one obtained by the peer review group.

Similarly, with regard to the second measure, error free clauses, the mean score obtained in the peer review group (0,83 in the pre-test and 0,93 in the post-test)



indicates that results were marginally better since they improved their mean in 0,10 while those of students who had done the previous writing collaboratively (0,86 in the pre-test and 0,90 in the post-test) improved their result in 0,04.

With respect to the third measure, error to words, both groups obtained the same result: a mean of 0,99 in the pre-test and 0,98 in the post test which makes a very slight improvement of 0,01 in both cases.

Therefore, according to the results retrieved, the data collected shows a positive effect in lexical accuracy in all the parameters analysed (error free T-units to total T-units, error free clauses to total clauses and errors to words) in both groups. Even though both groups improved their production in the post test, the improvement in the peer review group were higher than in the collaborative group in error free T-units and error free clauses.

*Table 8. Results of mechanical accuracy measures.*

<b>MECHANICAL ACCURACY</b>		<b>COLLABORATIVE WRITING</b>		<b>PEER REVIEW</b>	
		<b>Pre-test</b>	<b>Post-test</b>	<b>Pre-test</b>	<b>Post-test</b>
<b>MEAN</b>	<b>ERROR FREE T-UNITS / TOTAL T-UNITS</b>	<b>0,79</b>	<b>0,61</b>	<b>0,72</b>	<b>0,66</b>
	<b>ERROR FREE CLAUSES / TOTAL CLAUSES</b>	<b>0,85</b>	<b>0,82</b>	<b>0,75</b>	<b>0,80</b>
	<b>ERRORS / WORDS</b>	<b>0,99</b>	<b>0,98</b>	<b>1,00</b>	<b>0,97</b>

Table 8 provides an overview of the results of mechanical accuracy. As presented in it, the mean score of the first mechanical accuracy measure, error free T-Units, shows a worsening of the mean score in both groups in the post test. However, the peer review group obtained marginally better results since they went from 0,72 to 0,66 (which makes a negative difference of 0,06 points) while the collaborative writing group obtained 0,79 in the pre-test and 0,61 in the post test, implying a negative difference of 0,18 points. These results indicate better scores for the peer review group in the first mechanical accuracy measure.

The mean score of the second mechanical accuracy measure, error free clauses, shows that the peer review group improved their writings (a mean of 0,75 in the pre-test and 0,80 in the post-test) in 0,05 points while the collaborative writing group worsened their mean score in 0,03 points as compared to the pre-test (mean of 0,85 in the pre-test and 0,82 in the post-test).

As can be seen in Table 8, the mean score for the third accuracy measure (errors to words) was quite similar. The participants in the collaborative group obtained a mean of 0,99 in the pre-test and 0,98 in the post test which makes a slight positive difference of 0,01. Meanwhile, the peer review group obtained a mean of 1 in the pre-test and 0,97 in the post-test implying an increment of 0,03 points.

Hence, the mechanical accuracy mean scores suggest that the group who had worked on peer review obtained better results regarding error free clauses to total clauses and errors to words. While the collaborative group only improved in errors to words. And both groups did worse in error free T-units to total T-units.

#### 5.2.1.3 Complexity

Our study also explored if there were differences in regard to complexity in the final product of both groups. The mean scores of the two groups with respect to the three complexity measures are presented in Table 9.

*Table 9. Results of complexity measures.*

<b>COMPLEXITY</b>					
		<b>COLLABORATIVE WRITING</b>		<b>PEER REVIEW</b>	
		<b>Pre-test</b>	<b>Post-test</b>	<b>Pre-test</b>	<b>Post-test</b>
<b>MEAN</b>	<b>CLAUSES/T-UNITS</b>	<b>2,23</b>	<b>2,47</b>	<b>1,95</b>	<b>1,95</b>
	<b>WORDS/T-UNITS</b>	<b>20,71</b>	<b>22,28</b>	<b>19,28</b>	<b>17,29</b>
	<b>WORDS/CLAUSE</b>	<b>9,45</b>	<b>9,09</b>	<b>10,00</b>	<b>8,98</b>

Three measures were utilized to analyse the complexity of the language used in participants' written products: number of words per clause, number of words per T-unit and number of clauses per T-unit. As shown in Table 9, the mean scores in clauses per

T-units of the collaborative writing group (2,23 in the pre-test and 2,47 in the post-test) increased in 0,24 points while the peer review group (1,95 in both tests) did not improved neither worsened their mean in this parameter.

The data demonstrated an improvement regarding number of words per T-unit in the collaborative writing treatment group. This group obtained an increase of 1,57 points in the post-test over the pre-test, the initial score was 20,71 and the final score was 22,28. On the other hand, the peer review treatment group worsened their mean score in 1,99 points obtaining 19,28 in the pre-test and 17,29 in the post-test.

With regard to the third measure of complexity (words per clause), both groups worsened their results from the pre-test to the post-test. Collaborative writing participants obtained a mean score of 9,45 in the pre-test and 9,09 in the post-test which implies a negative difference of 0,36 points. In the peer review group, students obtained a mean score of 10 in the pre-test and 9,40 in the post test, worsening their results in 1,02 points.

The results obtained from the complexity measures suggest that the effect of writing collaboratively had a greater impact on students' productions in two of the analysed components (clauses to T-units and words to T-units) while the peer review group did not improve in any parameter.

## 6. Discussion

This study compared the differences between two intact groups in a final individual writing production after the students had had as a treatment either collaborative writing or peer review. The aspects analysed were fluency, lexical variety, accuracy and complexity using T-unit based measures.

The first research question investigated how students' written production were overall affected after the treatment session which was, in group A, collaborative writing, and in group B, peer assessment.

Regarding the group that received collaborative writing as a treatment, the global scores suggest that students had a beneficial effect on fluency, meaning that, after the treatment session, students produced markedly longer texts. These results contradict

findings in previous studies which stated that collaboratively writing productions does not seem to improve regarding fluency (Fernández Dobao, 2012; Kim, 2008; Shehadeh, 2011; Storch, 2005; Storch & Wigglesworth, 2007; Swain & Lapkin, 2001; Wigglesworth & Storch, 2009). Similar results were obtained regarding complexity, those participants who had received collaborative writing as a treatment showed an increased mean in the post-test. And, as regards accuracy, and contrary to the results found in the previous studies mentioned above (Fernández Dobao, 2012; Storch, 2005), the students who had worked in collaborative writing decreased their accuracy mean scores (error free T-Unit, clauses and words) in the post-test.

These students were lower-intermediate learners and accuracy is not the main focus in that level which tends to focus more on communicative competence, hence, and according to Storch (2005), it could be argued that with respect to accuracy low-proficiency students may not benefit from collaborative tasks.

Regarding those students who had received peer review in the treatment session, they did not show any increase in their fluency nor in their complexity scores but the measures present an overall improvement in accuracy and lexical variety. The results were worse than expected since as Soleimani and Rahmanian (2014) found out in their research, it was expected that the participants would increase their means in complexity, accuracy and fluency.

Therefore, the first hypothesis suggests that those students who received collaborative writing as a treatment obtained better results in fluency and complexity in the post-test, while those who carried out peer review did better in lexical diversity and accuracy.

The second research question was aimed to compare in more detail different aspects such as fluency, lexical variety, accuracy and complexity in both groups.

Fluency was measured by the total number of words produced and lexical diversity was calculated by counting the number of different words per the number of total words. The findings with regard to fluency indicated that the texts written by participants who had received collaborative writing as a treatment were a little more fluent than the texts written by those who had carried out peer review. Contrary to the

findings of Soleimani and Rahmanian (2014) the texts reviewed by a peer did not lead to improvements in fluency in the immediate post-test.

Regarding lexical variety, the results retrieved from the data suggest that no great differences could be found between both groups in this aspect. Both decreased their means in the post-test, even though the peer review group used a slightly wider variety of words.

Accuracy was measured calculating the proportion of error-free T-units to total T-units in terms of lexical, grammatical and mechanical errors. Considering the accuracy of the writings, results indicated that those students who had received collaborative writing as a treatment produced less accurate compositions in comparison to those who had carried out peer review.

One issue with regard to accuracy is that collaboration afforded students the opportunity to provide and get immediate feedback on language, an opportunity which is absent when learners work individually, this may justify why learners when writing collaboratively tend to write better texts in terms of accuracy (Fernández-Dobao 2012). Nevertheless, the results retrieved from the present study do not confirm previous findings from Fernández-Dobao (2012). Contrarily, our findings suggest that, regarding accuracy, the peer review group obtained better results than the collaborative writing group. However, this cannot be compared to similar studies since, to the best of my knowledge, and as mentioned in previous sections, there have not been any studies where both approaches have been compared.

In terms of grammar accuracy, the group who had worked on peer review obtained better global results compared to those who had worked on collaborative writing. Regarding lexical accuracy, both groups showed a positive effect in their global means in the parameters analysed, still the students who had worked on peer review obtained better results than those of the collaborative writing treatment. Apparently, something similar happened regarding mechanical accuracy, where those who had worked on peer review improved their global mean while participants who had received collaborative writing as a treatment only improved in one out of three aspects. It can, thus, be suggested that, in terms of accuracy, those participants who had carried out peer review obtained better results than those who had worked on collaborative writing.

Complexity was calculated by identifying clauses and T-units and applying three different measures of syntactic complexity (number of words per clause; number of words per T-unit; number of clauses per T-unit). The results for complexity are somehow in disagreement with the findings of Storch and Wigglesworth (2007, 2009). In 2009, these authors found that collaboration had no impact on grammatical complexity and that there were no great differences between the texts produced collaboratively and those produced individually. Nevertheless, in this study, the participants who had received collaborative writing in the treatment session revealed a slight positive difference compared to those who had worked individually and had carried out a peer review in the treatment session. Therefore, this finding seems to support the results of the studies which found that collaboration leads to more complex language use (Storch, 2001, 2005) and the findings from Soleimani and Rahmanian (2014) that confirm that peer assessment failed to improve the language complexity level of the writers.

The development of complexity, however, is typically postponed until a basic command of language proficiency is achieved since it represents the relative linguistic and cognitive difficulty of a task or an utterance and serves as a marker of linguistic sophistication (Norris & Ortega, 2009; Palloti 2015). And as the sample in this study was composed of lower-intermediate language learners, facilitating the development of accuracy and fluency seems more feasible than complexity, which typically requires a certain language control and suits the ability of advanced language learners.

Regarding the holistic measures, both groups revealed an overall improvement in the post-test. The global holistic results showed that students who had received either collaborative writing or peer review as a treatment produced better structured and organized texts and the ideas were more clearly exposed. Even though, no great improvements could be appreciated regarding vocabulary and grammar. These results are important because, the way of measuring students' writing competence in schools is through rubrics since they are considered to be one of the most objective tools. The written competence of the students is measured using holistic rubrics where not only grammar and vocabulary are measured but also different aspects of the written text such as cohesion and coherence, lay out, register, the organization of ideas into paragraphs, etc.

All in all, from the obtained findings regarding the CAF measures, it can be concluded that both collaborative writing and peer review are an effective approach that can be used to improve their writing skills of EFL students since, as it has been seen, it improves students' productions in many different aspects. Furthermore, it can be appreciated that writing does not necessary have to be an individual act but it can be done or complemented with another students' contribution. Moreover, with these techniques, teachers encourage pupils to interact creating a positive social atmosphere where they discuss about the language, correct each other and solve grammatical and lexical misunderstandings.

## 7. Conclusions

The motivation of the present study was to get a new perception of written tasks and to know in detail the effects of two writing approaches. The main goal of this research was to analyse the effects of collaborative writing and peer review on a subsequent final individual writing of EFL students. The data was analysed using fluency, lexical diversity, accuracy and complexity as parameters.

First, considering the effect of the treatment on the fluency and lexical diversity of the written productions, it can be concluded that there were no great differences between groups. A slight improvement in the collaborative writing group and a decrease in group the peer review group can be observed, even though this increment cannot be considered important. Therefore, it can be inferred that collaborative writing might be moderately advantageous for producing more fluent subsequent written texts. Regarding lexical accuracy, the data suggests that no great improvements can be observed in either group, even though, those students who had worked on peer review produced little more varied number of words in the post-test.

Second, the findings obtained from the accuracy measures suggested that the participants from the peer review group obtained better results than those from the collaborative writing group but not many great differences could be found between groups. The improvements of the means for accuracy measures were more important in

those students who had worked on peer review, specifically grammar and mechanical accuracy measures.

Third, with respect to complexity, deductions can be made that the students who had received collaborative writing as a treatment produced more complex texts than those who had received peer review as a treatment. That is, collaborative writing had a positive effect on the complexity measures of the written texts.

The most important limitation for the current study lies in the short time devoted to the process of data compilation, which was gathered within a month. Only two writings were collected in order to analyse them. The students in the present study produced a pre-test and a post-test writing which were analysed. The participants also received one-session treatment, either collaborative writing or individual peer reviewed writing. However, only two writings might not be able to reveal significant effects in improving students' writings. Therefore, in further studies, the period of treatment should be extended, more writing tasks should be carried out and more texts should be collected. Further aspects to be taken into consideration are increasing the number of participants, incorporating delayed post-tests for analysing long-term retention effects, observing the students' interactions and reactions during collaborative writing and peer correction tasks, and interviewing participants to retrieve data related to their experiences on the interventions.

## 8. Acknowledgments

First of all, I would like to express my gratitude to Dr. Camino Bueno Alastuey who has been guiding me through all this study and giving me her significant and fundamental advice for the accomplishment of the research work. Thank you for your patience, quick response to all my enquiries and for conveying your positive energy.

It's a must to thank Vedruna FEC school and its staff for letting me implement the study and special thanks to Cristina Isturiz for allowing me performing the intervention in her class finding a spot in her tight schedule and, I couldn't forget Nerea Ugalde who has provided me with good advices for carrying out the study and also has let me do the intervention with her group of students.



Finally, I would like to thank the participants who have taken part in this study and for having done it willingly and disinterestedly. Without their help this research wouldn't have been possible.

## 9. References

- Asberg, P., & Nulden, U. (1999). *Making peer review in large undergraduate courses an option*. The Viktoria Institute and Department of Informatics, School of Economics and Commercial Laws, Goteborg University, Sweden.
- Ashraf, H., Mahdinezhad, M. (2015). The role of peer-assessment versus self-assessment in promoting autonomy in language use: a case of EFL learners. *Iranian Journal of Language*, Vol.5 (2): 110.
- Barnes, D. (1976). *From communication to curriculum*. London: Penguin Books.
- Birenbaum, M. (1996). Assessment 2000: towards a pluralistic approach to assessment in: M. Birenbaum & F. Dochy (Eds) *Alternatives in assessment of achievements, learning processes and prior knowledge*, (pp. 3–29), Boston, MA: Kluwer Academic Press.
- Bygate, M., Skehan, P., & Swain, M. (Eds.). (2001). *Researching pedagogic tasks: Second language learning, teaching, and testing*. London: Pearson.
- Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), (pp. 891-901).
- Coniam, D., & Lee, M. W. K. (2008). Incorporating wikis into the teaching of English writing. *Hong Kong Teachers' Centre Journal*, 7, (pp. 52-67).
- Cutler, H. & Price, J. (1995). The development of skills through peer assessment, In A. Edwards & P. Knight (Eds) *Assessing competence in higher education*, (pp. 150–159). Birmingham, Staff and Educational Development Series.
- Davis, S. (1998). *Creative Writing*. Forum V 36 (4), (p. 44).

- Diab, N. M. (2010). *Effects of peer- versus self-editing on students' revision of language errors in revised drafts*. *System*, 38, (pp. 85-95).
- Ekşi, G. Y. (2012). Peer review versus teacher feedback in process writing: how effective? *International Journal of Applied Educational Studies*, 13(1), (pp. 33-48).
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford: Oxford University Press.
- Falchikov, N. (1995). Peer feedback marking: developing peer-assessment. *Innovations in Education and Training International*, 32, (pp. 175-187).
- Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair and individual work. *Journal of Second Language Writing*, 21 (1), (pp. 40-58).
- Fernández Dobao, A. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, 18, (pp. 497–520).
- Ferris, D. R. (2003). *Response to student writing implications for second language students*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Freeman, M. (1995). *Peer assessment by groups of group work, Assessment and Evaluation in Higher Education*, 20, (pp. 289–300).
- García Mayo, M. P. (Ed.). (2007). *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters.
- Harmer, J. (2007). *How to Teach English*. New York: Longman.
- Horgan, D. D., Bol, L. & Hacker, D. (1997). *An examination of the relationships among self, peer, and instructor assessments*. Athens, Greece.
- Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9(3), (pp. 321-342).

- Hunt, K. (1965). Grammatical structures written and three grade levels. NCTE Research Report, 3. Urbana, IL: *The National Council of Teachers of English*.
- Hunt, K. (1970). Syntactic maturity in schoolchildren and adults. *Monographs of the society for research in child development*, 35, (pp. 134, 1-67).
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. Urbana, IL: The National Council of Teachers of English.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), (pp. 590-619).
- Liou, H. C. (2009). A case study of web-based peer review for college English writing. *Curriculum & Instruction Quarterly*, 13(1), (pp. 173-208).
- Liu, E. Z., Lin, S. S., Chiu, C. H., & Yuan, S. M. (2001). Web-based peer review: The learner as both adapter and reviewer. *IEEE Transactions on Education*, 44(3), (pp. 246-251).
- Liu, J., & Hansen, J. G. (2002). *Peer response in second language classroom*. Ann Arbor, MI: University of Michigan Press.
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18, (pp. 30-43).
- Magin, D. & Helmore, P. (2001). Peer and teacher assessments of oral presentations: how reliable are they? *Studies in Higher Education* 26, (pp. 287-298).
- Mangelsdorf, K. (1989). Parallels between speaking and writing in second language acquisition. In D. M. Johnson & D. H. Roen (Eds.), *Richness in writing: Empowering ESL students* (pp. 134-145). White Plains, NY: Longman.
- Mendonca, C. O., & Johnson, K. E. (1994). *Peer review negotiations: Revision activities in ESL writing instruction*. *TESOL Quarterly*, 28 (4), (pp. 745-769).

- Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, 14, (pp. 397–419).
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, (pp. 555–578).
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.
- Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51-78). Oxford: Oxford University Press.
- Pallotti, G. (2015). *A simple view of linguistic complexity*. *Second Language Research*, 31, 117–134.
- Plutsky, S., & Wilson, B. A. (Winter 2004). Comparison of the three methods for teaching and evaluating writing: A quasi-experimental study. *The Delta Pi Epsilon Journal*, 46(1), 50- 61.
- Race, P. (1999). *2000 Tips for lecturers*. Kogan Page, London.
- Ramsden, P. (1992). *Learning to teach in higher education*. London, Routledge.
- Richer, D. L. (1992). *The effects of two feedback systems on first year college students writing proficiency*. *Dissertation Abstracts International*, 53(08A), 2722.
- Ruegg, R. (2014). The effect of peer and teacher feedback on changes in EFL students' writing self-efficacy. *The Language Learning Journal*, (pp. 1-18).
- Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, 29, (pp. 73-82).
- Sambell, K. & McDowell, L. (1998). *The construction of the hidden curriculum: messages and meanings in the assessment of student learning, assessment and evaluation in higher education*, 23 (4), (pp. 391–402).

- Sato, M. (2013). Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *Modern Language Journal*, 97(3), (pp. 611-633).
- Schwartz, M. (1989). *Research Associate, for the Learning and Teaching Office*, Retrieved from: <http://www.ryerson.ca/lt>.
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20(4), (pp. 286–305).
- Shima G. & Farzaneh F. (2017). Peer versus teacher assessment: Implications for CAF triad language ability and critical reflections. *International Journal of School & Educational Psychology*, DOI.
- Sluijsmans, D., Brand-Gruwel, S. & Van Merriënboer, J. (2002). Peer assessment training in teacher education, *Assessment and Evaluation in Higher Education*, 27(5), (pp. 443–454).
- Soleimani, H., & Rahmanian, M. (2014). Self-, peer-, and teacher-assessments in writing improvement: A study of complexity, accuracy, and fluency. *Journal of Research in Applied Linguistics*, 5, (pp.128–148).
- Sotillo, S. (2002). Constructivist and collaborative learning in a wireless environment. *TESOL Journal*, 11(3), (pp.16-20).
- Stefani A.J. (1994) Self, peer and group assessment procedures. In: *An Enterprising Curriculum: teaching innovations in higher education*. Eds I.Sneddon and J.Kramer. (pp. 24-46). HMSO, Belfast.
- Storch, N. (1999). *Are two heads better than one? Pair work and grammatical accuracy*. *System*, 27, (pp. 363–374).
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(3), 153–173.
- Storch, N., & Wigglesworth, G. (2007). *Writing tasks: Comparing individual and collaborative writing*. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–177). London: Multilingual Matters.

- Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82, (pp.320–337).
- Swain, M. (2001). Integrating language and content teaching through collaborative tasks. *The Canadian Modern Language Review*, 58, (pp. 44–63).
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, M.I.T. Press.
- Vygotsky, L.S. (1978). *Mind in Society: The Development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Watanabe, Y. (2014). *Collaborative and independent writing: Japanese university English learners' processes, texts and opinions*. (Unpublished doctoral dissertation). University of Toronto, Canada.
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15, (pp. 179-200).
- Zhu, W. (2001). Interaction and feedback in mixed peer response groups. *Journal of Second Language Writing*, 10(4), (pp. 251-276).

## Appendix I

**Write a news article using the following information (120-150 words):**

Album signing  
Amaia Romero  
Friday, 9<sup>th</sup> march at 5pm  
Baluarte square, Pamplona  
Winner of the singing show *Operación Triunfo*; the singer who represents Spain in Eurovision.  
More than 3000 tweets; thousands of fans; long queues; posters to support her; awaiting to take picture;

## Appendix II

1. Read the following news article:



2. Follow the steps to write your own news article:

### Step 1:

a) Decide the topic of your article. Try to choose something related to the school (the exchange, the school's voluntary service, any students' attractive story, school trip...).

b) Research: Where will you get the information from? Will you have to interview anyone? Only use correct information.

c) Watch your language! Make sure you use the **third person**. You can use different tenses, newspaper often use the **present perfect** to say what has happened and the **past simple** to talk about when and where it happened.



## **Step 2:** Start planning

Answer the following questions:

**Who?**

**What?**

**Where?**

**When?**

**Why?**

**How?**

Now add:

**Quotes:**

**More information:**

**Think of a catchy lead paragraph:**

## **Step 3:** Take a sheet of paper and...Create your news!

- Write a byline
- Create a placeline
- Create a catchy lead paragraph
- Write in the 3rd person
- Write the body (1-3 paragraphs)
- Be ensure to include a quotation

## Appendix III

### Step 4: Peer Review

Correct your classmate's news. Grade the different sections from 1 to 10, being 1 the lowest and 10 the highest mark.

CONTENT	
It has all the parts of a news article (headline, byline, lead, body...).	
The headline and the body of the article are related.	
It answers the 5 W's (who, what, where, when, why).	
It includes quotations.	
It has plenty of detail.	
GRAMMAR	
It's written in third person (he, she, it, they).	
The grammar tenses used are accurate.	
The sentences are neither too short nor long.	
VOCABULARY	
The words haven't got any spelling mistake.	
The vocabulary used is varied.	
It is used the rephrasing to avoid repetitions.	
The text includes connectors.	
LAY OUT	
The handwriting is clear and legible.	
It is divided into paragraphs.	
The sentences have correct punctuation and begin with capital letter.	
It respects the margins.	

## Appendix IV

Use the information below to write a **news article (120-150 words)**.

- Celebrating FEC's day
- Students and teachers from Vedruna School
- Friday 13<sup>th</sup> of April 2018
- Pamplona
- 21 FEC schools in Spain; Special morning prayer in the classroom; three different festive events: pre-school at the chapel, act done by Natalia; primary school students at the chapel; secondary students at the gym. Hand-crafts; community work; watch videos from last year; sing song "Somos".

Appendix V

Measures	4	3	2	1
<b>Fluency</b>	The essay includes at least four, well developed paragraphs. The sentences are easy to follow and help to make the essay understandable.	The essay includes four paragraphs, but they are not well developed or the language is choppy.	The essay does not include four paragraphs or the choppieness of the language distracts from the meaning.	The essay does not have four paragraphs and it's unclear and incomprehensible.
<b>Cohesion</b> (Linguistic complexity)	A variety of sentence lengths of varying linguistic complexity. Tight cohesion and organization. The transitions between examples and paragraphs make for smooth reading.	Simple and expanded sentences that show emerging complexity used to provide detail. Some transitions are evident that make the essay easier to read.	Phrases and short sentences; some attempt at organization may be evidenced. The transitions between examples and between paragraphs are choppy or missing and may distract from the meaning.	Single words, set phrases or chunks of simple language. Not clear organization.
<b>Adequacy</b>	The essay includes body paragraphs in the most appropriate order to support the claim and the reasons. The length of the text is appropriate.	The essay includes paragraphs and is organized in a reasonable order to support the claim and reasons. The length of the text is not appropriate.	The essay does not include organized body paragraphs in a particular order to provide clear support for the claim and the reasons. The text is too short and ideas are not well developed.	The essay is disorganized. It may be missing indented paragraphs. It may be missing an introduction and/or conclusion. The instructions given are not followed and many parts of the news article are missing.
<b>Mechanics</b> (spelling, punctuation, capitalization)	Excellent capitalization, punctuation and spelling. Less than 2 errors.	Good use of capitalization, punctuation and spelling. Less than 5 errors.	Sometimes uses capitalization and punctuation. There are some errors on spelling. Less than 8 errors.	Incorrect use of capitalization and punctuation. There are many mistakes in spelling. More than 8 errors.
<b>Language control</b>	The essay is essentially error free or minor grammar errors. Good command of grammar. There is a sense of careful editing. No redundancy.	The essay includes occasional grammar errors. Fair command of grammar. Some sense of editing is obvious. Some redundancy might be present.	The essay includes multiple errors. Little or no sense of editing. Poor command of grammar. Multiple redundancies might be present.	The essay includes serious and many errors in spelling, capitalization and punctuation that often interfere the meaning. Very poor command of grammar.