The effects of using collaborative writing vs. peer review treatments on subsequent individual writings

Mª Camino Bueno-Alastuey
I-COMMUNITAS. Instituto de Investigación Social Avanzada (Public University of Navarre)
camino.bueno@unavarra.es

Seila Rodero Albaiceta
Public University of Navarre

Abstract: Previous research has investigated separately the effects of collaborative writing and peer review. However, there has not been any research comparing both approaches. This study is aimed to analyse the effects of those two treatments on a subsequent individually written production in terms of complexity accuracy, and fluency (CAF measures), as well as quality. The participants were 29 students of lower-intermediate English proficiency level, 16 in the collaborative writing group and 13 in the peer review group. Two compositions were analysed as pre-test and post-test using CAF measures and an analytic rubric. The results suggest that students in the collaborative writing group produced longer writings and more complex language, while those in the peer review group improved their final individual writing in terms of lexical variety and accuracy.

Keywords: collaborative writing; peer review; complexity; accuracy and fluency (CAF measures); lexical variety; EFL.

Resumen: Los efectos tanto de la escritura colaborativa como de la revisión por pares se han investigado por separado sin comparar ambos enfoques. Este trabajo tiene como objetivo comparar el efecto que estos dos tratamientos tienen en una redacción posterior tomando en consideración la complejidad, la corrección, y la fluidez (Medidas CAF), y además la calidad. 29 estudiantes de nivel de inglés intermedio bajo participaron en este estudio, 16 en el grupo de escritura colaborativa y 13 en el grupo de revisión por pares. Se analizaron dos redacciones, una como pre-test y otra como post-test, utilizando las medidas CAF y una rúbrica analítica. Los resultados sugieren que los estudiantes de la escritura colaborativa escribieron textos más largos y usaron un lenguaje más complejo, mientras los de la revisión por pares mejoraron sus redacciones en variedad léxica y corrección.

Palabras clave: escritura colaborativa; revisión por pares; corrección; complejidad y fluidez (Medidas CAF); variedad léxica.

I. Introduction

To be proficient in a language, learners need to be competent in reading, speaking, listening and writing, which are interrelated since the use of a 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, p. 1979).

Writing is one of the productive skills, which requires creating language. The same as when speaking, writing is used to share ideas, feelings, to convince others, etc. As Davis (1998) stated, «writing is essentially a creative process and good writers must learn to communicate their ideas clearly to an unseen audience» (p.25). It is considered one of the most arduous tasks to be performed due to the mental exertion it requires.

When second language learners are writing, they are forced to retrieve 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).

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). The benefits of both collaborative writing and peer-review are well-known since these approaches, rooted in sociocultural approaches to language learning, have been shown to help students' focus on language, promote noticing and enhance learners' thinking skills, especially when assessing their classmates (Lin & Yang, 2011; Villarreal & Gil Sarratea, 2019).

Both kinds of scaffolded assistance, collaborative writing and peer review, have been studied to analyse their effect on students' writing development (Bueno-Alastuey & Martinez de Lizarrondo, 2017; Diab, 2010; Lundstrom & Baker, 2009; Villarreal, & Munarriz-Ibarrola, in press). Nevertheless, to the best of the authors' knowledge, there are no studies comparing the effect both approaches might have on subsequent writings.

II. Literature Review

1. *Collaborative writing*

As described by Swain (2001), collaborative tasks are communicative tasks, which 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, p. 10). They demand communication, and when communicating, there are moments when students focus on language to solve misunderstandings, to enquire about usage doubts, or to correct each other. Those instances have been termed Language-related episodes (LRE) (Swain & Lapkin, 1998). Collaborative tasks also require that learners work together, sharing ideas and pooling their knowledge to achieve one common goal (Fernández Dobao, 2012).

In the last decade, the benefits of collaborative writing have been investigated and compared to individual writing. The majority of studies have been conducted from a cognitive perspective, focusing on the effect of task –measured in terms of complexity, fluency and accuracy (CAF)– on L2 production (see, for instance, Bygate, Skehan, & Swain, 2001; Ellis, 2003; García Mayo, 2007; Villarreal, Bueno-Alastuey, & Sáez-León, in press).

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, and the authors concluded that it was due to the LREs that had taken place.

A 16-week quasi-experimental study conducted by Shehadeh (2011) in an EFL context 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). Nassaji and Tian (2010) confirmed those findings comparing 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.

Fernández Dobao (2012) was the first study to compare group, pair and individual work in collaborative writing tasks and 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 were more accurate, produced more LREs and a higher number of correctly solved LREs than pairs and individual productions, and pairs better than individual productions.

More recently, Fernández Dobao (2014) focused on vocabulary learning in collaborative writing tasks, and compared pair and small group work. 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.

The same benefits including more complex and accurate texts, and of a higher subjective quality have been reported in secondary school settings (Villarreal & Gil Sarratea, 2019; Villarreal & Munarriz, in press). In these settings, triads have also been confirmed to produce more accurate, fluent and syntactically complex texts than both pairs and individuals (Bueno-Alastuey & Martínez de Lizarrondo, 2017).

Apart from benefits in writing quality and vocabulary learning, collaborative learning is also perceived by students as an enjoyable activity (Fernandez Dobao, 2013), which can provide opportunities to discuss and plan their writing, to generate ideas, to create texts, to give immediate peer feedback and to polish up their texts (Shehadeh, 2011). Furthermore, students have reported that collaborative writing have helped them have self-confidence and had improved their speaking abilities (Storh, 2005).

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. *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 to judge their own work and that of their colleagues. It 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, a series of guidelines have been provided including schemes and the list of criteria that will be used to assess the final products as well as recommendations to set-up and manage the process (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 Vygotskyan 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 and Lee (2008) and Lin and 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). This has also been suggested to happen in L2 learning as peer interaction helps L2 students communicate their ideas and can enhance the development of L2 learning in general, as students find «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» (Mangelsdorf, 1989, 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, Bol and Hacker (1997), and Sluijsmans, Brand-Gruwel and Van Merriënboer (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» (pp. 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, Badger and Yu, 2006), and develops learners' writing skills in subsequent writing drafts (Diab, 2010; Lundstrom and 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 and 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 results indicated that 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; while teacher assessment produced 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 and 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.

As can be seen from the aforementioned studies, peer review in writing helps to develop students' thinking skills, contributes to improvements in complexity, accuracy and fluency, and makes students become better language learners when reflecting about the language they are using (Schwartz, 1989).

III. Research questions

Considering both collaborative writing and peer feedback have shown to have benefits for the development of writing, this study aims at comparing the effect

of both treatments in a subsequent written production. The research questions which guided this investigation were:

- Do EFL students' writing skills improve after doing collaborative writing and peer review?
- Are there any differences in a subsequent individual production after having experienced either collaborative writing or peer review in terms of complexity, accuracy, and fluency?

IV. Methodology

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 education. Their ages ranged from 14 to 15. All participants were Spanish native speakers except a student whose first language was neither Spanish nor English, but had a native-like command of Spanish. All the participants in the study had had 4 hours of English a week during the school year.

The students had been divided into three different groups (A, B, C) depending on their level of English proficiency by their teachers. Groups A and B had a similar English level whereas the third, Group C, was more proficient in English. Taking into consideration that collaborative situations are more likely to happen among students with similar language level (Storch, 2005), and that the purpose of this research was to compare the effect of two treatments, similar level groups, Groups A and B, were chosen for the study.

From the initial 36 participants, 6 students had to be excluded because they were not present in some of the sessions, and one because of his illegible handwriting. Consequently, the final participants were 29 students: 16 in Group A, and 13 in Group B.

2. *Instruments and procedure*

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, thus, this would give ecological validity to the study. The students had not written this type of text previously. The topics were selected based on students' interests and motivations as reported by their teacher.

Data was collected by means of two different instruments: 1) a pre-test, which was a news article written individually; and 2) a post-test, which was another news article written individually. The second writings done collaboratively (group A), and individually and peer reviewed (group B) were also collected, but they were not analysed.

The intervention lasted three weeks. First, on the day of the pre-test, students were asked to produce a news article based on the instructions provided (see Appendix I) without having received any initial training. They were given 30 minutes to complete the task, and then, the writings were collected.

During the second week, the treatments were carried out. The students in each of the groups received a 30 minutes master class on news articles (see Appendix II) led by one of the authors. In this session, the format and the style of a news article was explained and 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 III) for assessing their assigned classmate's work, which was designed taking into consideration different aspects of writing such as grammar (tenses, third person singular -s, etc.), the lay-out used for this format, the lexical variety used by the writer, the content etc. 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, and the time given to complete the task were the same as in the pre-test.

3. *Data analysis*

The 58 compositions collected were analysed quantitatively using CAF measures, and holistic quality ratings based on an analytic rubric created by one of the authors.

Regarding the quantitative measures for complexity, accuracy and fluency (CAF), complexity was measured taking into consideration lexical diversity and grammatical complexity. Lexical diversity was calculated by counting the number of different words divided by the number of total words, while grammatical complexity included the number of words per clause, the number of words per T-unit and the number of clauses per T-unit. To calculate the three measures, clauses and T-units were identified and computed. A clause was codified as any unit consisting of a subject (visible or implied) plus a predicate, i.e. a construction with a finite or a non-finite verb as its head (Bulté & Housen, 2012). A T-unit was defined as «a main clause plus all subordinate clauses and nonclausal structures attached to or embedded in it» (Hunt, 1970, p.189).

Accuracy was measured as the proportion of error-free T-units to total T-units, the ratio of error-free clauses to total clauses, and the number of errors to 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). Three types of errors were taken into consideration: lexical, grammatical and mechanical errors. 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 <u>pre-school students ∅</u> at the chapel, act done by Natalia.* [First, the pre-school students went to the chapel and Natalia did the act.]

Example 2: **Errors in word order**:
S9b: *When she go out she was the <u>spain queen</u>.* [When she got out, she was the queen of Spain.]

Example 3: **Errors in subject-verb agreement**:
S6a: *<u>They was</u> celebrating FEC's day.* [They were celebrating FEC's day.]

Example 4: **Errors in use of articles**:
S8b: *She will win <u>the</u> Eurovision.* [She will win Eurovision.]

Example 5: **Errors in use of preposition**:
S13b: *<u>In</u> Friday 13th of April...* [On Friday 13th of April...]

Example 6: **Errors in use of verb tense**:
S7a: *This girl <u>has make</u> 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 ∅ 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.]

Fluency was measured by the total number of words produced, as in previous studies (e.g., Storch, 2007, 2008; Storch & Wigglesworth, 2007, 2010; Wigglesworth & Storch, 2009; Villarreal & Gil Sarratea, 2019).

Besides the quantitative analysis, a holistic analysis was carried out in order to have a qualitative perception of the quality of the texts. The texts were scored using a five factor analytic rubric with a four points scale (see Appendix V), where 1 represented the lowest mark and 4 the highest. The five factors rated in the rubric were i) fluency, considered as the number of words in the text, and lexical variety; ii) cohesion, which evaluated the development of ideas; iii) adequacy, which analysed whether the objective was fulfilled, the appropriateness of the length of the text and its organization in terms of task completion; iv) mechanics, that assessed the spelling, punctuation and capitalization; and v) language control, which measured the use of agreement, number, tense, word order, pronouns, articles, prepositions and negation.

V. Results

1. *Improvement of EFL students' writing skills*

a) *Global quantitative data analysis*

The first research question aimed to investigate whether participating in collaborative writing or in peer feedback practices helps to improve EFL students' individual writing skills.

As can be seen in Table 1 and according to the parameters analysed, the mean scores indicate that students participating in the collaborative writing condition improved their fluency (from 92.25 words to 106) and grammatical complexity (10.79 vs 11.28), but not their lexical diversity (0.75 vs 0. 67) and accuracy (0.76 vs 0.73).

Table 1. General results collaborative writing group

	Pre-test	Post-test
Complexity	10.79	11.28
Lexical diversity	0.75	0.67
Accuracy	0.76	0.73
Fluency	92.25	106

As can be seen in Table 2, those students who carried out and received peer review on their writings did not show any improvement in their mean results as they obtained better results in the pre-test than in the post-test, except for a slight increase in accuracy (0.74 vs 0.75).

Table 2. General results individual peer reviewed writing group

	Pre-test	Post-test
Complexity	10.41	9.40
Lexical diversity	0.74	0.68
Accuracy	0.74	0.75
Fluency	108	91.17

These global results suggest that students' improvements were different depending on the treatment. 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 accuracy.

b) *Holistic data analysis*

Table 3. Results of holistic measures

	Pre-test	Post-test
Collaborative writing	6.47	7.03
Peer review	6.79	7.29

The holistic analysis took into consideration the quality of the tests produced and analysed the content, organization, language use, vocabulary and mechanics subjectively. The results obtained from the scoring of those aspects using an analytic rubric showed 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, they improved slightly more in the post-test increasing their mean from 6.47 to 7.03 (+ 0.56), while the peer review group increased their mean from 6.79 to 7.29 (+ 0.50).

2. *Differences in subsequent individual writings*

The second research question sought to examine the effect of both treatments in a subsequent writing regarding complexity, accuracy and fluency.

a) *Quantitative data analysis*

- *Complexity*

In this study, four measures were utilized to analyse the complexity of the language used in participants' written texts: a lexical measure, diversity; and three grammatical measures – number of words per clause, number of words per T-unit, and number of clauses per T-unit.

Table 4. Results of complexity measures

	Collaborative writing		Peer review	
	Pre-test	Post-test	Pre-test	Post-test
Lexical diversity	0.74	0.67	0.74	0.68
Grammatical complexity	10.79	11.28	10.41	9.50
CLauses/T-Units	2.23	2.47	1.95	1.95
Words/T-Units	20.71	22.28	19.28	17.29
Words/Clause	9.45	9.09	10.00	8.98

As can be seen in Table 4, scores of lexical diversity were quite similar for both groups. Students who had done collaborative writing obtained a mean of 0.74 in the pre-test and of 0.67 in the post-test (-0.07). Something similar happened to the students who had done peer review. They scored 0.74 in the pre-test, and 0.68 in the post-test (-0.06). These results seem to indicate that neither carrying out collaborative writing nor peer assessment had a positive effect on the lexical diversity of subsequent productions.

Regarding grammatical complexity (see Table 4), the group who had written collaboratively improved from the pre-test to the post-test (10.79 vs 11.28), while the group providing peer feedback worsened from the pre-test to the post-test (1.40 vs 9.50).

Analysing the three measures which form grammatical complexity, the mean scores in clauses per T-units of the collaborative writing group increased from 2.23 to 2.47 (+0.24), while the peer review group had the same mean in both (1.95). An improvement could also be observed regarding number of words per T-unit in the collaborative writing treatment group, who increased their score from 20.71 to 22.28 (+1.57). On the contrary, the peer review treatment group worsened their mean score from 19.28 to 17.29 (-1.99). With regard to the third measure of complexity (words per clause), both groups worsened their results from the pre-test to the post-test. The collaborative writing participants scored 9.45 in the pre-test and 9.09 in the post-test (-0.36), while the peer review group worsened their results from 10 to 9.40 (-1.02).

These results suggest that none of the methodologies seem to benefit lexical diversity, but writing collaboratively helped improve two measures of grammatical complexity (clauses to T-units and words to T-units).

- *Accuracy*

As can be seen in Table 5, the collaborative writing group decreased their mean in global accuracy from 0.76 to 0.73 (-0.03), while the peer review group improved their global mean from 0.74 to 0.75 (+0.01).

Table 5. Results of the global accuracy measures

	Pre-test	Post-test
Collaborative writing	0.76	0.73
Peer review	0.74	0.75

Three measures of accuracy (grammar, lexical and mechanical) were considered to measure global accuracy. Regarding grammar accuracy and 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 (-0.05), while the group who had carried out peer review improved their productions from 0.33 to 0.36 (+ 0.03). With regard to the second measure (error free clauses), both groups obtained worse results in the post-test. The group carrying out collaborative writing experienced a negative difference of 0.06 points (from 0.53 to 0.47) and the group who had carried out peer review of 0.02 points (from 0.45 to 0.43) This implies that, although neither group improved their scores, the peer review group obtained better results than the collaborative writing group.

Table 6. Results of grammar accuracy measures

	Collaborative writing		Peer review	
	Pre-test	Post-test	Pre-test	Post-test
Error Free T-Units / Total T-Units	0.28	0.23	0.33	0.36
Error Free Clauses / Total Clauses	0.53	0.47	0.45	0.43
Errors / Words	0.95	0.89	0.95	0.88

Finally, as Table 6 indicates, the mean score of the third grammar accuracy measure (errors to words) showed that the participants in the collaborative group had a mean of 0.95 in the pre-test and of 0.89 in the post-test, so they improved 0.06 because having a lower score indicates fewer errors and, thus, better performance. Students in the peer review group also performed better in the post-test, obtaining a positive difference of 0.07 (from 0.95 to 0.88). Therefore, regarding errors to words, both groups improved their mean in the post-test.

The results obtained in grammar accuracy indicate that the peer review group improved in two of the components of grammar accuracy: error free T-units to total T-units, and errors to words. On the other hand, 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

	Collaborative writing		Peer review	
	Pre-test	Post-test	Pre-test	Post-test
Error Free T-Units / Total T-Units	0.71	0.76	0.73	0.84
Error Free Clauses / Total Clauses	0.86	0.90	0.83	0.93
Errors / Words	0.99	0.98	0.99	0.98

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 (+ 0.11). 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 (+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, 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 was higher than in the collaborative group in error free T-units and error free clauses.

Table 8. Results of mechanical accuracy measures

	Collaborative writing		Peer review	
	Pre-test	Post-test	Pre-test	Post-test
Error Free T-Units / Total T-Units	0.79	0.61	0.72	0.66
Error Free Clauses / Total Clauses	0.85	0.82	0.75	0.80
Errors / Words	0.99	0.98	1.00	0.97

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 (-0.06) while the collaborative writing group obtained 0.79 in the pre-test and 0.61 in the post-test (-0.18). 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 from 0.75 to

0.80 in the post-test (+ 0.05), while the collaborative writing group worsened their mean score from 0.85 to 0.82 in the post-test (-0.03).

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 (+ 0.01), while the peer review group obtained a mean of 1 in the pre-test and 0.97 in the post-test (+ 0.03).

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. In addition, both groups did worse in error free T-units to total T-units.

- *Fluency*

To examine fluency, the total amount of words produced were counted and analysed. As can be seen in Table 9, students who had written collaboratively obtained a mean of 92.25 words in the pre-test and 106 words in the post-tests (+ 13.75). Meanwhile, students who had written individually and then carried 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 (-16.83). Therefore, these results suggest that the effect of writing collaboratively was better than the effect of peer review.

Table 9. Results of fluency

	Pre-test	Post-test
Collaborative writing	92.25	106.00
Peer review	108.00	91.17

VI. 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 complexity, accuracy, and fluency 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 experienced a beneficial effect on fluency, as students produced markedly longer texts after the treatment. These results contradict findings in previous studies which stated that collaboratively writing productions does not seem to improve fluency (Fernández Dobao, 2012; Kim, 2008; Shehadeh, 2011; Storch, 2005; Storch & Wigglesworth, 2007; Swain & Lapkin, 2001; Wigglesworth & Storch, 2009), but support findings in secondary school settings (Bueno-Alastuey & Martínez de Lizarrondo, 2017). Similar results were obtained regarding complexity, those participants who had received collaborative writing as a treatment showed an increased mean in the post-test. Moreover, 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 those reported by previous studies (Soleimani & Rahmanian, 2014) in which the participants increased their means in complexity, accuracy and fluency.

Regarding the quality rating measures, both groups revealed an overall improvement in the post-test. The global 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. However, 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 analytic rubrics. They are considered one of the most objective tools, and they rate not only grammar and vocabulary, but also different aspects of the written text such as cohesion and coherence, lay out, register, the organization of ideas into paragraphs, etc.

Therefore, the answer to our first question seems to be that both treatments are beneficial for students as both methodologies help improve some aspect of subsequent writings. Collaborative writing as a treatment affected positively fluency and complexity, while peer review provoked an improvement in lexical

diversity and accuracy. Furthermore, both treatments improved the subjective quality of subsequent writings.

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

Complexity was calculated using a measure of lexis, lexical diversity, and three measures of syntactic complexity (number of words per clause; number of words per T-unit; number of clauses per T-unit). Regarding lexical diversity, the results 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.

Regarding grammatical complexity, our results are somehow in disagreement with the findings of Storch and Wigglesworth (2007, 2009). They reported that collaboration had no impact on grammatical complexity, and that there were no great differences between the texts produced collaboratively and those produced individually. On the contrary, our participants in the collaborative writing group had 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 other studies which have found that collaboration leads to more complex language use (Storch, 2001, 2005), and confirm that peer assessment failed to improve the language complexity level of the writers (Soleimani & Rahmanian, 2014).

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). 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.

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.

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. Something similar happened regarding mechanical accuracy, where those who had worked on peer review improved their global mean, while participants in the collaborative writing group only improved in one out of three aspects. Thus, it can 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.

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.

All in all, from the obtained findings regarding CAF measures, it can be concluded that both collaborative writing and peer review are effective approaches that can be used to improve the writing skills of EFL students since, as it has been seen, they improve 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, thus, introducing oral communication while writing. 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.

VII. 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 in a secondary EFL context. The main goal of this research was to analyse the effects of collaborative writing and peer review on a subsequent final individual writing. The writings were analysed using complexity, accuracy, and fluency as parameters.

First, with respect to lexical diversity and grammatical complexity, the data suggest that no great improvements in lexical diversity can be observed in either group, even though those students who had worked on peer review displayed more lexical diversity in the post-test. On the contrary, the students from the collaborative writing group produced more complex texts than those in the peer review group. That is, collaborative writing had a positive effect on the complexity measures of the written texts.

Second, the findings obtained from the accuracy measures indicate that the participants from the peer review group obtained better results than those from the collaborative writing group, but the differences were slight. 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, considering the effect of the treatments on the fluency 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 the peer review group can be observed. Therefore, it can be inferred that collaborative writing might be slightly advantageous for producing more fluent subsequent written texts.

The findings in this study need to be taken with caution, as there are some important limitations, which need to be taken into consideration. The most important limitation 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. 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. Another important limitation is the fact that no statistical analysis was carried out to see whether the differences were statistically significant. Further research should include inferential analysis of the data. 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.

Despite its limitations, this study has shown that both methodologies should be pursued in secondary school settings as they affect different aspects of writing development. Collaborative writing should be used to improve measures of complexity and fluency, and to introduce an oral component into written tasks, while peer review should be practiced to improve lexical diversity and accuracy and to develop students' higher order thinking skills.

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*, *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.

Bueno-Alastuey, M. C., & Martínez de Lizarrondo, P. (2017). Collaborative writing in the EFL Secondary Education classroom: comparing triad, pair and individual work. *Huarte de San Juan. Filología y Didáctica de la Lengua*, 17, 254-275.

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), 891-901.

Coniam, D., & Lee, M. W. K. (2008). Incorporating wikis into the teaching of English writing. *Hong Kong Teachers' Centre Journal*, 7, 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), 44.

Diab, N. M. (2010). Effects of peer- versus self-editing on students' revision of language errors in revised drafts. *System*, 38, 85-95.

Ekşi, G. Y. (2012). Peer review versus teacher feedback in process writing: how effective? *International Journal of Applied Educational Studies, 13*(1), 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*, 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), 40-58.

Fernández Dobao, A. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, *18*, 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, 289-300.

García Mayo, M. P. (Ed.) (2007). *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters.

Ghahari, S., & Farokhnia, F. (2018). Peer versus teacher assessment: Implications for CAF triad language ability and critical reflections. *International Journal of School & Educational Psychology*, *6*(2), 124-137.

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. *European Association for Research on Learning and Instruction*. Athens, Greece.

Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, *9*(3), 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, 134-167.

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), 590-619.

Liou, H. C. (2009). A case study of web-based peer review for college English writing. *Curriculum & Instruction Quarterly*, *13*(1), 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), 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, 30-43.

Magin, D., & Helmore, P. (2001). Peer and teacher assessments of oral presentations: how reliable are they? *Studies in Higher Education*, 26, 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), 745-769.

Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, 14, 397-419.

Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 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. (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*. London: Kogan Page.

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*, *46*(2) 1-18.

Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, 29, 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), 391-402.

Sato, M. (2013). Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *Modern Language Journal*, *97*(3), 611-633.

Schwartz, M. (1989). *Research Associate, for the Learning and Teaching Office*. Retrieved from: http://www.ryerson.ca/It.

Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, *20*(4), 286-305.

Sluijsmans, D., Brand-Gruwel, S., & Van Merriënboer, J. (2002). Peer assessment training in teacher education, *Assessment and Evaluation in Higher Education*, *27*(5), 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, 128-148.

Sotillo, S. (2002). Constructivist and collaborative learning in a wireless environment. *TESOL Journal*, *11*(3), 16-20.

Stefani, A. J. (1994). Self, peer and group assessment procedures. In I. Sneddon & J. Kramer (Eds.). *An Enterprising Curriculum: teaching innovations in higher education*. (pp. 24-46). Belfast: HMSO.

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System*, 27, 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, 320-337.

Swain, M. (2001). Integrating language and content teaching through collaborative tasks. *The Canadian Modern Language Review,* 58, 44-63.

Villarreal, I., & Gil Sarratea, N. (2019). The Effect of Collaborative Writing in an EFL Secondary Setting. *Language Teaching Research*. doi: https://doi.org/10.1177/1362168819829017.

Villarreal, I., Bueno-Alastuey, M.C., & Sáez-León, R. (in press). Computer-based collaborative writing with young learners: Effects on text quality. In M.P. García Mayo & A. Imaz, *Working collaboratively in second/foreign language learning.* Alemania: De Gruyter.

Villarreal, I., & Munarriz-Ibarrola, M. (in press). «Together we do better»: The effect of pair and group work on young EFL learners' written texts and attitudes. In M. P. García Mayo & A. Imaz, *Working collaboratively in second/foreign language learning.* Alemania: De Gruyter.

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, 179-200.

Zhu, W. (2001). Interaction and feedback in mixed peer response groups. *Journal of Second Language Writing*, *10*(4), 251-276.

Appendix I
Pre-test

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

Album signing
Amaia Romero
Friday, 9th 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
Treatments

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
Peer Review instructions

Step 4: Peer Review

Correct your classmate's news report. Grade the different sections on a scale from 1 to 10, where 1 represents 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
Post-test

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

- Celebrating FEC's day
- Students and teachers from Vedruna School
- Friday 13th 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
Rubric

Measures	4	3	2	1
Fluency	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 choppiness of the language distracts from the meaning.	The essay does not have four paragraphs and it's unclear and incomprehensible.
Cohesion (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.
Adequacy	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.
Mechanics (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.
Language control	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.