PublicacionesDidácticas
www.publicacionesdidacticas.com

The Effectiveness of Focused Direct Written Corrective Feedback and Metalinguistic Explanations on the Use of English Articles by Vocational Training Students (Part II)

Title: The Effectiveness of Focused Direct Written Corrective Feedback and Metalinguistic Explanations on the Use of English Articles by Vocational Training Students (Part II).

Abstract
The second part of this paper presents an empirical study in which the effects of Direct Written Corrective Feedback and Metalinguistic Explanations are compared, as measured by a Task Repetition test and an Error Correction Test, in order to test the participants' use of the English article. The results reveal the usefulness of both types of feedback, with an advantage for Metalinguistic Explanations. These results are reported and discussed, and pedagogical implications are suggested.
Keywords: Language Learning, Written Corrective Feedback, Metalinguistic Explanations

Título: La Efectividad del Feedback Directo y las Explicaciones Metalingüísticas en el Uso de Artículos por Alumnos de Formación Profesional (Parte II).

Resumen
La segunda parte de este artículo presenta un estudio empírico en el que se comparan el feedback directo y las explicaciones metalingüísticas, usando task repetition y un test de corrección de errores, con el fin de observar el uso que los participantes hacen de de los artículos en inglés. Los resultados revelan que ambos tipos son útiles para el aprendizaje y dan cierta ventaja a las explicaciones metalingüísticas en ambas pruebas. Estos resultados se discuten y se indican posibles implicaciones pedagógicas.
Palabras clave: Aprendizaje de Lenguas, Feedback, Correcciones, Explicaciones Metalingüísticas.

I. INTRODUCTION

The first part of the study offers the theoretical background in which the Language Learning Potential (LLP) of writing in an additional language (L2) is discussed, followed by the LLP of WCF, different types of WCF and Task Repetition. Part II presents an empirical study whose aim is to examine the effects of DWCF and ME on the texts produced by 22 vocational training learners and to determine whether one type of feedback is more effective than the other. To this end, a Task Repetition (TR) test and an Error Correction Test (ECT) were administered, in order to test the participants' use of the English article. The results reveal the usefulness of both types of WCF with an advantage for ME. These results are reported and discussed, and pedagogical implications are suggested.

II. THE CURRENT STUDY

II. 1. Aims and Research Questions

The aim of this empirical study is to analyze the extent to which written corrective feedback can be useful in improving learner's accuracy when writing and to look into the potential differential effects of Direct WCF (DWCF) and metalinguistic explanations (ME). To this end, this study employed a pretest / posttest design, with one week between sessions, and attempted to study the learners' knowledge and use of English articles, by using an Error Correction Test (ECT) and a narrative task, in an approach inspired by the one used in Shintani and Ellis (2013). This way, the two main research questions (RQ) may be enunciated as follows:

- ○ RQ1: Does the provision of WCF help L2 writers improve the accuracy in their use of English articles?
- ○ RQ2: Do different types of feedback differentially affect potential improvements in the use of English articles by L2 learners and, if so, which is more beneficial?

II. 2. Method

II.2.1. Participants and Context

All participants were students from a secondary school in Murcia. They belonged to two superior vocational training courses and had English as one of their compulsory subjects. According to the curriculum and the textbook they usually worked with, their level of proficiency was A2-B1, which is considered elementary-intermediate by the Common European Framework of Reference (CEFR). Nevertheless, they were also asked to rate their proficiency in the four main skills themselves in the background questionnaire they completed in the first session. They rated their receptive skills with a 2.9/5 for reading and 2.6/5 for listening. Not surprisingly, they rated their productive skills even lower, with 2.5/5 for both speaking and writing.

Since all participants were at least 18 years old, they signed a participant's consent form during the first session, in which 32 people participated. They were randomly divided into three different groups, one treatment group for each of the two WCF types studied and third one that worked as a control group, with a distribution of 11 / 11 / 10 participants in each. Nevertheless, due to experimental mortality, some of those participants did not return for the second session and were consequently eliminated from the study. Therefore, the final number of participants was 22 and the number of participants in each group was even, with 8 in the DWCF group, 8 in the ME group and 6 in the control one.

All participants were males and most of them were Spanish, with the exception of a Peruvian student. Spanish was their first language (L1), except for one participant who was Spanish but whose L1 was Arabic. The age of the participants ranged from 18 to 39 with a mean of 21.2 and a standard deviation (SD) of 4.6. Nevertheless, not counting the outlier who was 39 years old, the mean age was 20.5 with a SD of 2.8.

II.2.2. Target Structure

The present study is focused on English articles: the definite article *the*, the indefinite articles *a/an* and the zero article. As Shintani, Ellis and Suzuki (2014, citing Hawkins, 2001) explain, articles can convey different meanings via two binary features: "whether the article and associated noun phrase refer to a specific entry [±specific referent] and whether the article and associated noun phrase are deemed to be already known (from the previous discourse or from context) to the listener [±hearer knowledge]" (p. 111).

The reason for choosing this target structure is that articles are bound to appear in a writing composition, no matter the topic or the length of the text. Sheen (2007) also points out that *the* and *a* belong to the top five most frequently occurring words in English, with *the* being the clear cut first. He also explains that articles are sometimes not explicitly taught, which was the case of the participants of this study as well, and claims that "the SLA literature has clearly shown that English articles are considered to be a nonsalient feature [whose] misuse does not lead to communication breakdown" (p. 262). Apart from that, the multi-functional nature of English articles does not follow Andersen's one-to-one principle (1984, in Shintani & Ellis, 2013) which helps learners in the early stages of L2 acquisition and is essential in the development of their interlanguage, thus making it a difficult structure for them.

Previous research on articles has focused only on the use of the definite article *the* for first mention and the indefinite article *a/an* for anaphoric mentions (Sheen, 2007; Bitchener & Knoch, 2010) or studies have been restricted only to the use of *a/an* (Shintani & Ellis, 2013; Shintani et al., 2014). The present study, in contrast, focuses on all three cases: the definite article, the indefinite article and the zero article.

II.2.3. Treatment Conditions

- ○ Direct written corrective feedback group: they were provided with a photocopy of their first writing composition (the originals were not altered in any way) in the second session. It included explicit corrections about misuses of articles, as explained in section I.1.3.

- Metalinguistic explanation group: they were provided with a copy of their first writing composition in which errors concerning articles had been labeled 1 to 5. They were also given a handout with metalinguistic explanations of five cases in which articles are used and how they are used. This handout was in Spanish to ensure that they understood the explanations and it also included examples for better clarity. Their errors were marked and labeled but not explicitly corrected, as they had been in the other group.

- Control group: a smaller third group of students did not receive any form of feedback or correction on their compositions. Instead, they were given their original compositions back and were asked to try to improve them by themselves in the second try. Having a control group is a common practice in this kind of studies, since it allows us to check if they also improved without the provision of feedback, in which case the improvements of those participants who did receive any form of feedback could not be entirely attributed to the feedback alone but rather to the fact of having engaged in a form of task repetition.

II.2.4. Instruments and Materials

- **Participant's consent form**: For ethical reasons, all participants signed a consent form. It was a short form in which participants acknowledged having information about the study and being able to ask questions, as well as authorizing the use of their information in the study, always keeping their anonymity. All students were of age so they all signed the form themselves rather than having their parents do it.

- **Background questionnaire**: All participants completed a background questionnaire that included information such as gender, age, nationality, L1 and a self assessment of the four main skills.

- **Metalinguistic explanation handout** (see appendix A): The ME group received a handout with metalinguistic explanations for five cases in which English articles are used differentially. The handout was in Spanish to ensure that students understood the explanation and it also included examples. One of the main advantages of a handout like this is that participants were able to see all five cases and, time permitting, learn from them even if they had not had all five kinds of errors in their compositions. Furthermore, students expected to be more likely to identify errors and correct them if they know that errors labeled with the same number are similar in nature, rather than trying to correct the two errors separately and perhaps not making the connection. In addition, correcting one's own errors by using this kind of handout is also easy for teachers, who can identify errors and label them with a number instead of having to write down metalinguistic explanations for each error in each composition. Therefore, there were both theoretical and pedagogical reasons for the choice of metalinguistic explanations.

- **Error Correction Test** (see appendix B): It included twelve decontextualised sentences and its main aim was to check the participants' use of the target structure. The ECT included 6 errors about articles and 6 distractors, including errors such as verb tenses, personal pronouns or possessive determiners. The same ECT was completed by the participants as the pre test and the post test.

- **Narrative task**: The participants were provided with a sheet of paper with delimited margins and line spacing in order to facilitate correction later. The English taught in these two vocational training courses is aimed at their specialty (rather than having the usual topics such as nature, travelling or food) and I decided to choose a topic that was familiar to both courses, as suggested by their tutor. They were asked to write a 200-250 word essay describing what they usually do in a normal day at the workshop. They were allowed to use dictionaries since it would not affect their use of articles and my main concern was to avoid participants struggling with vocabulary and consequently writing short compositions. While they managed to get by, nobody reached the expected length in their compositions.

- **Exit questionnaire**: It included multiple choice questions about whether or not the participants believed their second composition would be better than the first, whether or not they would have done better if they had had more time to complete the task, or their opinion on the experience of participating on a research study like this. Apart from the control group participants, who made few changes to their texts, most of the students believed their second text was better and having more available time would not have made an impact on its quality, so no further investigation was carried out in relation to this questionnaire. It also included an open ended question for general comments which they all left unanswered.

II.2.5. Procedure

The study was carried out in two one-hour long sessions (S1 and S2) following the pretest / post test design. The first session opened greeting the participants during the first few minutes and explaining to them how the two sessions would come about. Then, they had 5 minutes to read and sign the consent form and another 5 minutes to complete the demographic questionnaire. After that, participants had 15 minutes to complete the ECT, in which they had to rewrite 12 sentences and correct the mistakes they contained. They devoted the last 30 minutes to writing their compositions (WR1) and time seemed not to be an issue. In addition, allowing 30 minutes seems to be sufficient time to produce an adequate sample of writing (Jacobs et al., 1981, cited in Coombe et al., 2007).

The second session took place one week after S1 and had a different structure. During the first 10 minutes of the session, participants from the DWCF and ME groups were given back their WR1 with their respective corrections and asked to revise and learn from their mistakes, whereas the control group was given their WR1 without any feedback and still had those 10 minutes to think of possible improvements in the use or articles by themselves. After that, the DWCF and ME groups handed back their corrected compositions and they were given their original unlabeled WR1. Then, all participants had 30 minutes to write a new version of the same text, trying to correct their previous errors. Consequently, none of the three groups had any corrections in front of them when writing the new version (WR2) but they had their original compositions so they could focus on form and not on remembering what they had written the first time. After they finished the WR2, they were asked to complete the same ECT they did in S1, in 15 minutes again. Finally, they devoted the last 5 minutes to completing the exit questionnaire.

II.2.6. Data Coding and Analysis

While other similar studies (Bitchener & Knoch, 2010; Sheen, 2007; Shintani & Ellis, 2013, Shintani et al., 2014) studied accuracy as a percentage of correct uses of the target structure in relation to occasions in which its use was obligatory, the present study counted the number of misuses of the target structure, articles in this case. The reasons for taking this approach include the fact that most participants, probably because of their low level of proficiency, used long lists of vocabulary in order to make their compositions longer. In my view, these lists distorted the results, since a student could add a list of eight items, for instance, and net eight correct uses to his score, thus obtaining a high percentage of accuracy that could mask other mistakes. Nevertheless, I could not count lists as a single item either since some participants would use both correct and incorrect articles in the same list, making it relevant to count every misuse separately. For instance, one participant wrote "I put on work clothes, a boots and a gloves", which counted as two errors. Conversely, a participant could write "I put on work clothes, boots and gloves" and obtain three correct uses out of three, thus increasing his percentage even though this participant may have not even thought about articles and was just making a list of words he or she knows. Additionally, most compositions were similar in length, reducing the statistical advantage of using percentages rather than counting instances of errors. For these reasons, I deemed adequate this approach. Errors in WR2 were classified as (i) uncorrected, (ii) new errors, (iii) corrected and (iv) left out errors.

The answers in the ECT were scored on a discrete item basis, giving one point to each sentence corrected, similar to what other authors have done in similar studies (Sheen, 2007; Shintani & Ellis, 2013). Both the sentences including the target structure and those including distractors were taken into account to check if the participants obtained different results in distractor sentences after receiving feedback only on articles.

II. 3. Results

This section is divided into two subsections, one devoted to the results obtained in the written composition and another one to those reporting the results of the ECT. On the one hand, the first subsection includes a table (Table 1) and a graph (Figure 1) indicating the results obtained in the writing compositions (WR1 and WR2), encompassing the mean of errors, the standard deviation and the range of the errors. On the other hand, the second subsection includes tables 2 and 3, as well as figures 2 and 3, in which similar information is displayed in relation to the parts of the ECT containing the target structure and the distractors, respectively.

II.3.1. Written composition

All participants, except for one, made either the same amount or fewer errors in WR2 in comparison to WR1 and the mean of errors of all participants was reduced by half. As can be seen in Table 1, the DWCF group started with the highest number of errors out of the three groups but also made the biggest improvement, going from 5.8 to 2.5 errors, less than half, even though the only participant who did not improve in WR2 belonged to this group. Additionally, 2 of the 8 members of the group managed to make 0 errors in their WR2. Participants were divided into the 3 study groups randomly and, even though the DWCF group obtained a higher mean of errors than the other groups, I decided to keep this distribution rather than modifying the groups manually. I gave more importance to distributing them randomly before knowing the results of the pre test, at the risk of them having different results in the pre test, since I had dealt with these students during my first practicum period and I did not want to make a biased distribution.

The ME group, in contrast, reduced their average number of errors from 3.6 to 0.9, a fourth of the original quantity. With that mean, they were the group with the best results in WR2 and 5 members made 0 errors in WR2. Finally, the control group obtained exactly the same figures in WR2, including a member who corrected one of their mistakes but made a new one.

Table 1: Results of the written compositions.

Written compositions	N	Mean: errors	Standard Deviation	Range
WR1: All participants	22	4.4	2.9	0 (2 times) - 9 (3 times)
WR2: All participants	22	2.2	2.9	0 (8) - 10 (1)
WR1: DWCF group	8	5.8	2.9	2 (2) - 9 (2)
WR2: DWCF group	8	2.5	3.3	0 (2) - 10 (1)
WR1: ME group	8	3.6	2.6	0 (1) - 8 (1)
WR2: ME group	8	0.9	1.7	0 (5) - 5 (1)
WR1: Control group	6	3.6	3.3	0 (1) - 9 (1)
WR2: Control group	6	3.6	3.3	0 (1) - 9 (1)

(Lower is better)

No errors were left out (one of the options contemplated in this study), which was to be expected since participants had an uncorrected copy of their WR1 when working on WR2 and they copied it word by word, with the exception of those errors they corrected or the few new errors they had. There was a mean of only 0.3 new errors per participant but only 4 of the 22 participants had new errors and, since they belonged to the three groups and the number was so low, it was deemed insignificant.

Figure 1 shows that the DWCF group had a lower slope, meaning they made the bigger improvement of the two, but it is also true that the ME group finished with a lower mean of errors, below one error per composition.
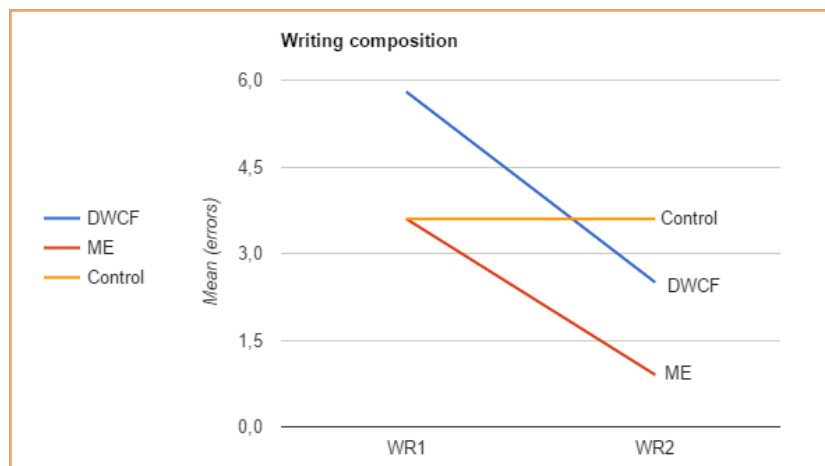
Figure 1: Mean of errors in the written compositions.

As distinctively shown in Figure 1, both treatment groups had similar behaviours and a similar reduction of errors. The participants in the control group showed no improvement whatsoever across tasks, hence not having benefitted from having engaged in task repetition. in contrast, the treatment groups did benefit from having been provided with WCF, which might indicate that the important variable at hand is having access to some form of feedback rather than TR alone. This issue will be further analyzed in the Discussion section below.

II.3.2. Error Correction Test

In terms of the 6 sentences in the ECT including the target structure (see Table 2), the results in the pretest were very low, as a whole. All participants had either the same amount or more correct sentences in the post test, going from a mean of 0.5 correct uses to a mean of 1.9. As indicated in Table 1, the DWCF group had the lowest mean in the pretest and went from 0.1 to 1.6 correct uses of the article. That improvement of 1.5 correct uses was similar to the one the ME group experimented, going from 0.5 to 2.1 correct uses. Finally, the control group obtained the highest mean in the pretest. Nevertheless, this result should be taken with caution since 4 of the 6 members of the group got 0 sentences right, whereas one of the two remaining members obtained the top score, reflected in the high SD. Therefore, the progress of the DWCF and ME groups was similar and their improvement was bigger than that of the control group.

Table 2: Results of the ECT in sentences containing the target structure.

ECT (Articles)	N	Mean (Correct / 6	Standard Deviation	Range
PRE: All participants	22	0.5	1.4	0 (17) - 6 (1)
POST: All participants	22	1.9	1.8	0 (7) - 6 (1)
PRE: DWCF group	8	0.1	0.4	0 (7) - 1 (1)
POST: DWCF group	8	1.6	1.6	0 (3) - 3 (2)
PRE: ME group	8	0.5	1.1	0 (6) - 3 (1)
POST: ME group	8	2.1	1.9	0 (2) - 5 (1)
PRE: Control group	6	1.2	2.4	0 (4) - 6 (1)
POST: Control group	6	1.8	2.3	0 (2) - 6 (1)

(Higher is better)

As for the other half of the ECT (see Table 3), which included 6 distractors, the results were varied and the mean decreased from 2.8 correct sentences to 2.7. As a whole, nevertheless, the results in this half of the ECT were better than those in the other half in both the pre test and the post test. On the one hand, the ME group was the only one to improve, going from 3.1 to 3.4 correct sentences on average. On the other hand, the DWCF group went from 2.1 to 2 correct sentences whereas the control group also went from 3.3 to 2.8 and one of the 2 students who had obtained the highest score in the pretest obtained a lower score in the post test.

Table 3: Results of the ECT in sentences containing distractors

ECT (Distractors)	N	Mean (Correct / 6)	Standard Deviation	Range
PRE: All participants	22	2.8	2	0 (4) - 6 (4)
POST: All participants	22	2.7	1.9	0 (3) - 6 (3)
PRE: DWCF group	8	2.1	1.5	0 (1) - 4 (2)
POST: DWCF group	8	2	1.5	0 (1) - 5 (1)
PRE: ME group	8	3.1	2.4	0 (2) - 6 (2)
POST: ME group	8	3.4	2	0 (1) - 6 (2)
PRE: Control group	6	3.3	2.4	0 (1) - 6 (2)
POST: Control group	6	2.8	2.3	0 (1) - 6 (1)

(Higher is better)

Figures 2 and 3 graphically show the evolution of the three groups. As can be seen in the first one, the ME group made a bigger improvement, as indicated by the higher slope as well as the higher scores overall, while the control group improved the least but obtained a higher mean that the DWCG group anyway in the half of the ECT containing the target structure. As can be seen in Figure 3, the DWCF had the lowest scores in both the pretest and the post test but the control group had the biggest decrease in scores. On the contrary, only the ME group improved in this half of the ECT.

Figure 2: Results of the ECT (target structure)   **Figure 3:** Results of the ECT (distractors)

II. 4 Discussion

It was expected that all participants would improve in their second writing, since two of the groups had received feedback and there was no reason for them to do worse than in the pre test. On the one hand, the DWCF group could infer some rules and improved in the post test. While they only had to remember their particular errors, which were always fewer than nine, they failed to correct all of them. The ME group, however, inferred the rules from their mistakes and most participants made 0 errors in their WR2, suggesting the benefits of ME over DWCF. Similarly, Shintani and Ellis (2013) also reported that the ME group outperformed the DWCF in the writing composition although the differences they observed were not statistically significant. On the other hand, the control group obtained the same results as in WR1, as expected. Nevertheless, the reason behind the control group obtaining the same mean of errors was not caused by participants making the exact same errors as in WR1 in all cases. One of the students took risks and corrected one of their mistakes, making a different error in the process and consequently keeping the mean of errors of their group the same. Therefore, both the DCWF and ME groups improved in the WR2 and the control group seemed not to benefit from engaging in a form of task repetition.

As for the half of the ECT containing the target structure, the results in the pretest were discouraging, which probably can be explained, but not justified, by the fact that students may have been focusing on correcting content words rather than articles in the sentences. In the post test, nevertheless, all three groups improved, since they already knew what to pay attention to and two of them had received feedback earlier in that session. The control group may have improved as a result of test practice effect and engaging in a form of task repetition, as was the case in Sheen (2007). As he reported in his study, the two WCF groups outperformed the control group anyway in this study as well, thus suggesting that WCF had an impact on the improvement of those two groups on top of test practice effect. The ME group obtained the best results, which was expected since they had studied the rules in the handout they were provided, rather than correcting isolated cases like the DWCF group, and were prone to have more explicit knowledge of the use of articles. In this regard, Shintani and Ellis (2013) also reported that the ME group in their study outperformed the other groups in the second administration of the ECT but they did not consider the difference statistically significant. Conversely, the DWCF had to infer the rules themselves and still managed to improve in the post test. As Shintani and Ellis explained after doing a stimulated recall interview with the participants, which was not included in the present study, many of the participants from the DWCF group struggled to infer the rule from the errors corrected in their compositions and some of them did not even thought about trying to infer such rule, which is likely to be the case in this study as well. As a whole, both WCF groups showed greater improvements than the control group (see Figure 2) and, while their improvements were similar, the ME group obtained higher scores in the end.

While other studies do not pay attention to the results obtained in the part of the ECT containing distractors (Shintani & Ellis, 2013), I considered important to see if the participants' scores changed in the post test and how the provision of feedback may have affected them. The results in this half were varied, with the DWCF and control groups obtaining worse scores than in the pre test. These results were surprising since, given the lack of any sort of feedback regarding the contents of those sentences, students were expected to obtain the same results. This may be explained by test practice effect, the fact that they took more risks in the post tests or even cases or hypercorrection. Since they already knew in the post test that the focus of the study was on English articles and they did not know that there were distractors included in the ECT, participants may have double guessed articles that were already correct when they failed to find the real error in the rest of the sentence. The ME group was probably affected by test practice effect and being able to do the same ECT again as well, and was the only group that improved. Since they had studied the rules earlier, they had more explicit knowledge than the rest of participants and were able to notice the real errors with articles and spot the correct uses as well, thus improving their scores slightly in this part. As a result, the ME group was the only one to improve in this half of the ECT and both the DWCF and the control group obtained lower scores, which was unexpected since they were expected to obtain a similar or better score because of task repetition and test practice effect.

Therefore, in response to the first research question of the present study, we can conclude that WCF, both DWCF and ME, can lead to improved accuracy in the use of articles. These result align with the ones reported by Guo (2015), in whose study both WCF groups also outperformed the control group and groups who had received less explicit forms of feedback such as underlining or error codes. Nevertheless, the results of the present study contradict those obtained by Shintani and Ellis (2013), who reported no significant improvements for the DWCF group in either the written composition or the ECT and their ME group only improved slightly in their compositions but with no statistical significance. Similarly, Shintani et al. (2014) reported no improvements in the use of articles for any of the study groups. They explain, however, that they studied both articles and the use of the hypothetical conditional and suggest that participants may have paid too

much attention to the conditional, while neglecting the use of articles. As for the type of knowledge that was fostered in the present study, it is difficult to judge whether or not the implicit knowledge of the participants improved, since only a week had passed between S1 and S2. Nonetheless, considering how the scores of the compositions improved and, more specifically, the scores in the ECT did, it seems that the participants explicit knowledge of the target structure did improve, especially in the ME group. This result is similar to what was reported by Shintani and Ellis (2013), who explained that ME were more likely to improve the participants' explicit knowledge. Thus, the provision of WCF, both in the form of DWCF and ME, improved the participants accuracy in the use of English articles and their explicit knowledge.

In terms of the second research question, it seems safe to conclude that ME can help improve accuracy the most, out of the two types of WCF included in this study. Similarly, Sheen (2007) also reported that focused ME helps improve accuracy. While Bitchener & Knoch (2009) reported no advantages for ME, they suggested that the difference may reside in the ME   provided and wondered how types, amount, frequency and delivery of the ME can affect the results. Nevertheless, one year later, Bitchener & Knoch (2010) reported that the ME group and ME group plus oral form focused review had outperformed the rest of groups in their study. They acknowledge that their participants had an already high level of proficiency but argue that the improvements were indeed significant. Finally, Shintani and Ellis (2013) also reported improvements for the ME group in the immediate post test but argued that in the long term, after a third session, there was no statistical difference between that one and the WR1 in the first session. This result aligns well to those obtained in this study, since only a second session was studied and the effects of time were not put to the test. Consequently, since the ME group improved the most in the written composition and both halves of the ECT, we can conclude that it was more beneficial than DWCF for the participants.

II. 5 Conclusions and Pedagogical Implications

Previous research on the matter and this study seem to empirically back up the belief that WCF can be very beneficial to language learners. As for the effects of TR, it seems that L2 learners with a low proficiency level and little experience or training in L2 writing, like the participants of this study, do not benefit from the expected advantages of TR discussed in the literature (Manchón, 2014). However, while TR alone did not seem to be useful, TR and WCF combined did result in improvements for both groups.

In terms of future pedagogically oriented research, this study is not without shortcomings and, had I had more time and resources, it would have been interesting to extend the number of sessions to study how the passing of time affected the different groups. After all, as mentioned in the previous section, Shintani and Ellis (2013) conducted a similar study and reported different results in the immediate second session and the third session. Consequently, future pedagogically oriented research on WCF should focus on studying the effects of ME, the type of WCF that is usually left out of studies in favor of direct and indirect WCF, and studies should include more sessions and time between each of them, in order to monitor the progression of each group of participants longitudinally.

As for pedagogical conclusions, while ME seems to be more helpful than DWCF even at low proficiency levels, there are implications to consider from the perspective of English teaching. One of the key factors in the success of this type of feedback is that this study was focused. Because of this, the ME handout was concise and detailed to explain a particular target structure, English articles in this case. Nevertheless, the feedback provided by teachers in real classrooms is often unfocused in order the cover as many errors as possible, thus reducing the viability of using ME as the preferred form of feedback. In cases of focused feedback, ME may be the best type of WCF since it seems to benefit learners the most, plenty of work has been put into preparing the handout and, once the handout is prepared, it is easy and quick to provide feedback. Providing unfocused ME, however, would be more difficult, with teachers having to provide an explanation for each error, which would not be as accurate as in focused feedback, and would be extremely time consuming. For this reason, DWCF remains the most common type of WCF in English classrooms. Taking into account how ME seems to be superior to DWCF, perhaps it would be interesting to foster the use of focused feedback in the classroom. This way, for instance, teachers could focus on a different target structure for each unit and consequently be able to provide a detailed handout of ME rather than having to use DWCF. The structures targeted could include conditional tenses, verb tenses or the comparative and superlative form of adjectives, for example. Then again, unfocused feedback has the advantage of covering all types of errors and future research should also study the advantages and disadvantages of focused and unfocused feedback, since it is an issue closely related to that of the different types of feedback that teachers can use in the classroom.

APPENDIX A (Metalinguistic Explanation handout)

El uso de los artículos en inglés (*a / an / the*):

Por una parte, hay que tener en cuenta si hemos mencionado el objeto antes o lo estamos mencionando por primera vez.

1. Si lo mencionamos por primera vez y es singular, usaremos *a / an* en función de si la siguiente palabra empieza por vocal o consonante. Por ejemplo:
   - "I ate <u>an</u> apple for breakfast".
   - "I bought <u>a</u> guitar for my friend".

2. Si hablamos de objetos en plural y es la primera vez que los mencionamos, no usamos *ningún artículo*. Por ejemplo:
   - "I ate __ apples for breakfast".
   - "He bought some __ albums the other day".

3. Cuando ya hemos mencionado uno o más objetos antes y nos volvemos a referir a ellos, usamos *the*. Por ejemplo:
   - "I saw <u>a</u> movie last night. <u>The</u> movie was very interesting".
   - "I bought <u>a</u> guitar for my friend. He loved <u>the</u> guitar".

	Singular	Plural
Primera vez:	**a / an**	**No article**
Ya mencionado:	**the**	

Por otro lado, tenemos que tener en cuenta si nos estamos refiriendo a un objeto en concreto o no.

4. Cuando hablamos de un objeto en general, sin referirnos a ninguno en concreto, usamos el artículo indefinido *ningún artículo* o *a / an*, en función de si el objeto es plural o empieza por consonante o vocal, si es singular. Por ejemplo:
   - "<u>A</u> chair has four legs". "<u>__</u> Chairs have four legs".
   - "<u>An</u> elephant is heavier than <u>an</u> ant".
   - "<u>__</u> Apples are good for your health".

5. Sin embargo, si hablamos de un referente concreto y está claro a cuál nos referimos, usamos *the*. Por ejemplo:
   - "<u>The</u> chair in my room is broken".
   - "<u>The</u> elephant in my city's zoo is enormous".

APPENDIX B (Error Correction Test)

Rewrite these sentences correcting the errors they contain. There are no punctuation or spelling errors. You can follow this example:

> Example: Last Sunday, we goed to the beach.
>
> Answer: Last Sunday, we <u>went</u> to the beach.

1. I was thirsty so I went to buy the bottle of water.
_____

2. You are not allowed to smoked in class.
_____

3. If I won lottery, I would buy a new car for my parents.
_____

4. Joe likes Sally and wants to invite she to dinner.
_____

5. I really want to eat sandwich right now.
_____

6. Last week I ordered a set of new wheels. I was very excited when wheels arrived.
_____

7. If I pass my exams, I will invite mine friends over for dinner.
_____

8. I forgot to do my homework tomorrow.
_____

9. James ate three cake the other day.
_____

10. We was sure that you were from Murcia.
_____

11. My brother gifted me the guitar for my birthday.
_____

12. The mice like to eat cheese.

Bibliografía

- Bitchener, J. & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System*, *37*, 322-329.

- Bitchener, J. & Knoch, U. (2010) Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback*. Journal of Second Language Writing, 19*, 207-2017.

- Coombe, C., Folse, K., Hubley, N. (2007). *A Practical Guide to Assessing English Language Learners*. Ann Arbor MI: University of Michigan Press.

- Manchón, R. M. (2014). The distinctive nature of task repetition in writing. Implications for theory, research, and pedagogy. *ELIA, 14.* 13–42.

- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*, 255-283.

- Shintani. N. & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*, 286-306.

- Shintani, N., Ellis, R. & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning, 64,* 103-131.