

THIS IS AN EARLIER VERSION OF THE MANUSCRIPT. FOR THE FINAL VERSION, PLEASE CHECK THE JOURNAL WEBSITE: Serrano, R. (2011). From metalinguistic instruction to metalinguistic knowledge, and from metalinguistic knowledge to performance in error correction and oral production tasks. *Language Awareness, 20*, 1-16.

From metalinguistic instruction to metalinguistic knowledge, and from metalinguistic knowledge to performance in error correction and oral production tasks

The purpose of this study is to analyze the effect of metalinguistic instruction on students' metalinguistic knowledge on the one hand, and on students' performance in metalinguistic and oral production tasks on the other. Two groups of primary school students learning English as a foreign language were chosen. One of them (Rule group) was provided with metalinguistic instruction on English possessive determiners (PDs) during six weeks (N = 21), and the Comparison group (N = 22) did not receive such instruction. These students' progress was analyzed through a pre-test/post-test design by means of a written error correction task, a 'free production' oral task and a metalinguistic judgment task. The results of the statistical analyses indicate that, although the learners in the Rule group were more advanced in their knowledge and use of the English PDs than their peers in the Comparison group, the differences between groups were not statistically significant in all the tests. Additional analyses revealed that there were correlations between students' knowledge and performance in the Rule group, indicating that the learners who made the most gains from pre- to post-test were the ones who demonstrated a more advanced knowledge of the rule.

Metalinguistic knowledge and second language acquisition

The relevance of metalinguistic instruction in second language (L2) classes has been a controversial issue in the second language acquisition (SLA) field. Traditionally, L2 teaching methods included mostly metalinguistic explanations of L2 structures and translations from the students' L1 (first language) to the L2, and there was hardly any opportunity for communicative practice in the classroom. The typical outcome of these programmes was great levels of accuracy in grammar tests on the part of the learners, but lack of skills to actively use the language in communicative or real-life situations. This type of instruction has been referred to as 'focus on forms' (Long, 1991; Long & Robinson, 1998), and includes methods such as the Grammar Translation Method, the Audiolingual Method, the Silent Way or Total Physical Response. The general failure of these methods to promote fluent language use encouraged the development of other approaches that have communicative competence as a central goal. In these approaches (e.g., the Natural Approach, Procedural Syllabus, or immersion programmes) the focus is on the meaningful use of the L2 and explicit metalinguistic explanations are generally discouraged. The students following L2 programmes that focus on meaning have been characterized as being highly fluent; however, their performance has been reported to be far from native-like due to a lack of grammatical and also pragmatic accuracy. That is why a certain attention to language forms has been claimed to be positive in programmes

that generally focus on meaning (Genesee, 1987; Harley & Swain, 1984; Spada & Lightbown, 1999; Swain, 1998; White & Ranta, 2002)

Even though most researchers of instructed SLA nowadays would agree that approaches that provide metalinguistic instruction exclusively are likely to fail to promote L2 acquisition, there is not much agreement as to the effectiveness of metalinguistic explanations, or the degree to which such type of instruction should be included in L2 classes. Ellis (2003) suggests that explicit form-focused instruction contributes to L2 acquisition. Moreover, considering his review of the research, Ellis concludes that extensive explicit instruction on 'simple' target forms can lead to the acquisition of implicit L2 knowledge. DeKeyser (2003; 2007; 2009) maintains that declarative knowledge of grammar rules helps proceduralization and automatization of the L2. He also argues that explicit learning is especially beneficial for adults (DeKeyser, 2000), and for simple categorical L2 rules (DeKeyser, 1995). Other authors suggest that form-focused instruction should be provided for structures that are problematic for the students, but L2 classes should mainly focus on communicative language use. This approach has been referred to as 'focus on form' (Long, 1991; Long & Robinson, 1998). Although instruction on language forms under this approach can be provided implicitly or explicitly, the former type is encouraged because it is less disruptive (Doughty, 2003).

On the other hand, some claims have been made that explicit attention to language forms might lead to more remarkable L2 gains than implicit instruction in communicative classes. Spada and Lightbown (1999) examined the effect of implicit instruction on the development of question formation in English by Canadian primary school students. Some of the learners included in the study moved up one stage after the treatment;

however, most learners remained in the same question formation stage at which they were before receiving any implicit instruction on question formation. The authors suggest that the learners in their study did not make as much progress as other learners in comparable studies because their treatment included implicit instruction (vs. explicit instruction examined by Pienemann, 1984; 1989).

Research on the acquisition of English possessive determiners (PDs) by francophone learners in Montreal also suggests that explicit instruction (study by White & Ranta, 2002) might be more beneficial in classes that focus on meaning than implicit instruction (White, 1998). These two studies were performed in grade 6 classes in Quebec whose 11-12-year-old francophone students were following an intensive English as a Foreign Language (EFL) programme. This intensive programme offered the students the opportunity to receive EFL instruction over five months in a school year (approximately 400 hours of instruction; 20 hours/week), while the other five months were devoted to their regular curriculum in French.

White (1998) examined the effect of implicit instruction in the form of input enhancement on learners' use of PDs in a passage correction and a picture description task. Although some minor differences existed in the post-test, White (1998) found that there was no difference in the performance of the students that had received the input enhancement treatment and those who had not at a delayed post-test.

White and Ranta (2002) investigated the effect of explicit form-focused instruction on oral production and metalinguistic tasks, and the relationship between both types of performance. Even though the EFL classes under analysis focus on meaning, for the purpose of the study that these authors performed, one of the groups (experimental or

Rule group) received explicit form-focused instruction on the English possessive determiners, while the other (Comparison group) continued with the communicative programme. White and Ranta (2002) found that the students who received metalinguistic instruction on *his* and *her* had higher levels of performance in oral production and in metalinguistic tasks than those who did not receive any explicit instruction on these forms.

The study by White, Muñoz, and Collins (2007) followed the same treatment as White and Ranta's but included teenage participants (grade 8) in Canada and in Spain, who were receiving 'regular' EFL instruction (2-3 hours a week). Their results are in line with those reported by White and Ranta (2002), confirming the positive effect of metalinguistic instruction.

Ammar and Spada (2006) also examined the acquisition of English PDs by primary school students in Canada, but focused on the effect of different types of corrective feedback (prompts, recasts, or no feedback) for groups receiving explicit instruction. Their results indicate that the high-proficiency learners in this study performed similarly regardless of corrective feedback type; however, the low-proficiency learners benefitted more from the most explicit type of corrective feedback, i.e., prompts.

The purpose of the present study is to examine the effect of the treatment developed by White and Ranta (2002) with a population of a similar age range, but in a different context and following a different approach to L2 instruction. The research questions are slightly different to the ones proposed by the above-mentioned authors, since the main interest of the present study is to examine the effect of metalinguistic instruction on

learners' knowledge on the one hand, and on learners' performance on the other. More specifically, the following research questions guide the present investigation:

- 1) After a period of six weeks of instruction on English PDs involving metalinguistic explanations and practice, do learners demonstrate *knowledge* of these forms in a metalinguistic judgment task?
- 2) Do learners receiving metalinguistic instruction on English PDs improve their *performance in an error correction task* that targets these forms?
- 3) Do learners receiving metalinguistic instruction on English PDs improve their *performance in a 'free production' oral task* that targets these forms?
- 4) Is children's metalinguistic knowledge (as expressed in the metalinguistic judgment task) related to their gains in performance in the two tasks under consideration?

In view of the studies that have reported positive results for form-focused instruction in L2 classes (Norris & Ortega, 2000; White, Muñoz, & Collins, 2007; White & Ranta, 2002), it can be hypothesised that metalinguistic instruction on English PDs should have a positive effect on students' knowledge and use of these forms.

Considering the nature of the tasks used in this particular study, it could be expected that learners find it less challenging to demonstrate their knowledge of the target rule in a metalinguistic judgment task that is completed with the help of an investigator (see details in the Procedure section), than to use such knowledge in practical tasks. Both the rule for the use of PDs in English and its formulation are quite simple; therefore, the knowledge of the rule that the students demonstrate in this task is assumed to be a close reflection of their metalinguistic knowledge.

Regarding performance, one would expect it would be easier for learners to demonstrate their metalinguistic knowledge in controlled metalinguistic tasks involving error correction than in spontaneous oral production. Indeed, instructed L2 learners have been claimed to move from a controlled, effortful, conscious performance to a more automatic performance, which requires less voluntary control and attention on the part of the learner (Bialystok, 1994; McLaughlin, 1987; Segalowitz, 2000).

Finally, following theories that maintain that explicit knowledge can be used in spontaneous performance (DeKeyser, 2009; Ellis, 2003), it is expected that the learners' knowledge of the rule is reflected in their improvement in performance. Consequently, even if not all the students should necessarily assimilate metalinguistic explanations—there are individual differences in this respect, as Ranta (2002) suggests—the hypothesis is that those who do will experience more significant improvement in their performance.

Method

Programme and participants

The programme that was chosen for this study is included within the CLIL (Content and Language Integrated Learning) framework. According to Marsh (2002: 15), CLIL refers to 'any dual focused educational context in which an additional language, thus not usually the first language of the learners involved, is used as a medium in the teaching and learning of non-language content.' As explained by Langé (2002), some of the typical characteristics of these programmes are:

- The foreign language is used as a means of instruction from the early grades.
- The foreign language is introduced orally.
- The students learn the course subject and the foreign language simultaneously.
- There is more focus on form in CLIL programmes than in traditional immersion programmes.

The school considered for this investigation is a Catholic semi-private school (partly funded by the State). The students start being exposed to English as soon as they join this school when they are 3 years old. During the first four years they receive two hours a week of English instruction. As early as grade 2, the students start receiving content instruction in English, in addition to their English language class. The two subjects that are taught in this language from grade 2 to 4 are Social Sciences, and Arts and Crafts. In grade 5 and 6, apart from these subjects, the students are taught Science (laboratory) in English. A CLIL programme was chosen because it would have been difficult to find students in grade 6 (same age as the participants in White & Ranta, 2002) with an adequate command of the English language to be able to follow the treatment. Thanks to the increased number of hours of instruction in English—as compared to ‘regular’ EFL programmes—the students enrolled in CLIL programmes are more ‘comparable’ to the students examined by White and Ranta (2002).

Two groups were chosen for this study, group ‘E’ (school classification), which was the experimental group (or Rule group) and group ‘D’, which was the Comparison group. The former consisted of 21 students, all of them in grade 6. The Comparison group had 22 students; four of them were in grade 6, and the rest in grade 5. It was the school that organized the students in such groups for English class. What is important, however, is

that the two groups were comparable according to their teacher. Nevertheless, a test was performed before the treatment to confirm the teachers' impression (see Measures). All these English learners were Spanish/Catalan bilingual, since they lived in a region in Spain where these two languages are official and widely spoken.

The participants included in this study were the same age as the students in White and Ranta (2002) (11-12 years old). Furthermore, they had a proficiency level in English that was advanced enough to do the tasks and follow the treatment used for the Canadian students. Apart from these facts, the participants in this study and in White and Ranta's are quite different.

To begin with, the learners in the present study were following content-based instruction, apart from their English class, while the learners in Canada did not learn the L2 through content, but in intensive classes that focused on the language itself. Second, the intensive EFL programme analyzed by White and Ranta (2002) was an optional programme at grade 6, which offered the students the possibility to receive approximately 400 hours of English over one semester, more or less 20 hours a week, after which, they continued with their regular EFL classes, which were 'non-intensive'. Considering the duration of the treatment (6 weeks), the students in Canada had approximately 120 hours of English. On the other hand, the students in the Spanish groups had fewer hours of instruction overall and a more distributed exposure to the language. During the time of data collection the Spanish students were exposed to English for 4.5 hours a week: English language class, Social Sciences, Arts and Crafts, and Science (laboratory), which makes a total of 187 hours a year, or 27 while the treatment lasted.

Apart from the difference in hours and intensity of instruction, another aspect which is not similar considering the groups in the present study and the groups examined by White and Ranta (2002) is the teaching approach in the English class. The Canadian EFL intensive programmes follow a communicative approach (focus on meaning), whereas in Spain the students' EFL class focuses on forms. According to the English teacher, her classes focus on a grammar point that is chosen before hand (following the textbook). First, she presents it on the blackboard and asks the students what they know about the topic; therefore, they work on the grammar rules together. After the explanation, the students do different exercises, such as writings, readings and listening comprehension activities which include the target feature/s. The students will then do homework on the topic, and some time during the class the students will also do some oral practice, which is sometimes a whole-group discussion, and others a pair-work exercise. Therefore, even if the class is not completely teacher-centred (the students participate even during metalinguistic explanations), it is mainly focused on forms, since it is the teacher, and not the communicative situation, who decides which form should be on focus.

The students under analysis, then, can be said to receive two different types of instruction in English, depending on whether they are learning content (in which case the focus is mostly on meaning), or language (in which case the focus is on forms). In other words, they are learning English mainly implicitly through content in Social Science, Arts and Crafts and Science, while they are learning the language explicitly in their English class, through metalinguistic explanations.

For this particular study, the experimental group received explicit instruction on English PDs during their English class. Additionally, they were probably exposed to

these forms in meaningful contexts while they were being taught content in this language. The adverb ‘probably’ must be used here because the English possessive determiners are relatively frequent and are expected to appear in discussions on Arts, Science, etc. (‘Maria is doing her own experiment’, etc.); however, a controlled study on the input containing these forms in content classes was not performed. The Comparison group, on the other hand, had the chance of receiving input that included these forms in content and language classes, but no explicit instruction was provided for these learners.

It must be mentioned that, in previous years, all the students had been given a basic rule for the English PDs, but no specific practice or instruction had been given in the same year this study was performed. Even if the PDs were included in the syllabus for the two groups under analysis, the teaching of this form was postponed in the Comparison group until this study finished.

Target forms: English possessive determiners ‘his’ and ‘her’

The choice of these forms (*his* and *her*) is appropriate for the context under analysis because, as for the francophone learners examined by White and Ranta (2002), these features present some difficulty for Spanish/Catalan-speaking students (Muñoz, 1994; 2005). As in French, the PDs in Spanish and in Catalan agree with the possessed object (‘Maria estima *la seva mare i el seu pare*’) as opposed to English (‘Mary loves *her* mother and *her* father’), in which language the possessive determiner agrees with the possessor. Since the students’ L1 and L2 behave differently in this respect, PDs tend to cause some problems to Spanish/Catalan-speaking students. The other PDs in English

(*my, your, our, their*) do not seem to be as problematic because English makes no gender distinction and thus learners only have one form to learn. In the case of *his* and *her*, English learners have to learn two forms and make decisions based on gender, and thus it is expected that they initially rely on their L1 knowledge, believing that agreement takes place between the PD and the possessed object. Especially complicated are the cases in which the possessed object is a person of different gender from the possessor ('kin different'), as 'The girl is playing with her dad', as opposed to cases in which the possessed object is a thing ('The girl is playing with her toy') or a person of the same gender as the possessor ('The girl is playing with her mum'). Similarly, possessive determiners with body parts ('The girl is brushing her teeth') is especially problematic for Spanish/Catalan-speaking learners of English, because both Spanish and Catalan would use the indefinite article in such cases ('La niña se lava los dientes' / 'La nena es renta els dents').

Another important reason to consider the PDs *his* and *her* for this study is that the rule of thumb that explains how these forms are used in English is quite simple and easily formulated, which is why children as young as 11 years old are expected to be able to understand metalinguistic explanations referring to these forms and also verbalize them.

Procedure

Data for the present study were collected between the months of October and December. The instructional treatment as well as the measures used to examine the students' knowledge of English PDs were developed by White and Ranta (2002) and also

used by White, Muñoz, and Collins (2007). The treatment took place once a week over a period of six weeks for one of the groups (Rule group). During the first lesson, which lasted for about 40 minutes, the students were taught the rule of thumb (ask yourself ‘Whose ____ is it?’), after which the teacher established comparisons between the rule in English and in Spanish/Catalan. The materials used for the treatment consisted of rational cloze passages with some pictures, which the students had to complete using the determiners *his* and *her*. The text included in this activity was a description of the picture, which showed children with their parents in amusing situations. The students, who were organized in groups, worked through two cloze passages in the first lesson, and also in the subsequent sessions (which lasted for about 30 minutes). They first filled in the blanks individually, and then drew arrows from each PD to its referent. After they finished, they talked to the members of their groups and reached an agreement. Then, they would present their answers to the teacher, who would give them feedback. During each of the sessions the students were reminded of the rule of thumb before they were given the cloze activities to complete.

The Comparison group did not receive any instruction on English PDs. They continued with the established syllabus, and also with their content classes in English, in which they probably received input containing the forms under study, but for which they were not given any explicit explanation.

Measures

The students' progress was examined through a pre-test/post-test design for both groups. The pre-test, before the instructional treatment, consisted of the following tasks:

1. General grammar knowledge task: cloze test
2. Error correction task
3. Oral production task

In the first task the students were asked to do a cloze test consisting of 10 items, so that their general knowledge of English could be established. The reason for this task was mainly to ensure the two groups were comparable.

The error correction activity, with the title 'The Birthday Party', included 16 errors regarding PDs, and other distracter errors. The passage told a story about David's birthday party with illustrations on each page.

The oral production task consisted of picture descriptions. The students were shown six different pictures, which included children with their parents. The students were asked to describe what they saw. It was assumed that the students would consider the children in the picture as the protagonists and would narrate the stories focusing on the children, and thus producing sentences like 'This is a little boy and his mother. His mother is angry because he is dirty', etc. Consequently, the target PD in each picture corresponded to the child's natural gender, even though there was always the possibility of switching perspectives. The idea was to get a balanced number of *his* and *her*, and that is why the pictures included three boys and three girls. The students were shown the pictures one by one on an individual basis. The interviews were recorded and then transcribed.

After the instructional treatment a post-test was performed, which included the following activities:

1. Error correction task
2. Oral production task
3. Metalinguistic judgment task: metacomments on their performance on the error correction passage

The first task was the same as in the pre-test, involving an error correction activity of a passage describing David's birthday party. The second task was similar to the one the students did in the pre-test (describing cartoons with children and their parents); however, six different pictures were included. The same task was included in the case of the error correction activity because this task is more controlled and the students' performance is more easily compared if the same passage is included in the pre-test and in the post-test. Since the oral production task was open, even if the same pictures were kept at both test times, the students' performance would never be exactly the same for the comparison to be established in the same way as in the error correction task. That is why it was considered more appropriate to include different pictures, which would be more entertaining (the pictures depicted new funny situations) and challenging for the students.

Finally, after the picture description, an oral metalinguistic judgment task was performed. The students were asked to explain why they had corrected certain forms in their error correction task and not others. Random sentences were picked for all the students (the same for all of them), which included different contexts for the production of *his* and *her*. The aim of this task was to elicit students' knowledge of the rule of thumb. Since the formulation of such rule is simple (*His* for boys, *her* for girls), the reported knowledge of the target rule in this task can be considered a reflection of the students' metalinguistic knowledge.

Analysis

The cloze activity used in the pre-test was analyzed in terms of accuracy to provide the correct word for a given context. There was always one word which was the most appropriate, but other words were also accepted provided that they were grammatically and semantically correct for the corresponding sentence.

In order to examine the students' performance in the error correction activity ('The Birthday Party'), apart from counting the right corrections on the part of the students out of the 16 incorrect instances of PDs, the coding criteria used by White and Ranta (2002) were also adopted.

(Table 1)

In order to examine the students' progress in the oral production task, the stage analysis elaborated by White (1998), following previous work by Zobl (1985) and Lightbown and Spada (1990) was used.

(Table 2)

Finally, for the metalinguistic judgment task, the coding that was adopted was the same as the coding used by White and Ranta (2002) for this task. The codes are the following:

(Table 3)

Results

The results of the analysis of the scores obtained by the learners in the cloze test indicate that the two groups are comparable, since there were no significant differences in the scores obtained by the Rule group (5.35/10) and the Comparison group (5.19/10): $t(39) = -.243, p = .809$.

Research question 1: Metalinguistic instruction and metalinguistic knowledge (as reflected in a metalinguistic judgment task)

According to the descriptive statistics, in the Rule group 14.3% of the students were in level 1, 28.6% in level 2, 14.3% in level 3, and 42% in level 4 (see Table 3 for details on the levels). In the Comparison group the percentages were 36.4%, 22.7%, 22.7%, and 18.2% respectively. These results are more clearly represented in Figure 1.

(Figure 1)

In order to examine whether there were statistically significant differences between the two groups, a Mann-Whitney U test was performed. According to this test, the difference between the two groups was not significant, although it was leading towards significance in favour of the Rule group ($U 160.5, Z -1.77, p = .076$).

Research question 2: Metalinguistic instruction and performance in an error correction task

This issue was first investigated considering the total scores of the task (/16), with which parametric tests were performed in order to find out if differences existed between groups (same time different groups) and within groups (examining the progress of each group from pre- to post-test). Additionally, more analyses were conducted considering the developmental metalinguistic levels reported in Table 1, for which non-parametric tests were performed, due to the fact that the variables were ordinal and not interval.

Error correction considering scores (parametric tests)

The descriptive statistics including the means and standard deviations (in parenthesis) are reported in Table 4.

(Table 4)

A two-way mixed ANOVA was performed in order to examine differences in terms of time (pre-test vs. post-test), group (Rule vs. Comparison) and interaction between time and group. The results of the tests indicated that there were no significant differences for time ($F(39) = 3.44, p = .327, \text{partial } \eta^2 = .025$), group ($F(39) = .985, p = .071, \text{partial } \eta^2 = .081$) or time*group ($F(39) = .613, p = .438, \text{partial } \eta^2 = .015$). Moreover, when comparing the two groups in the pre-test and in the post-test, it was found that there were

no significant differences (pre-test: $F(39) = 1.62, p = .210, \eta^2 = .040$; post-test: $F(39) = .372, p = .545, \eta^2 = .009$).

The within-group parametric tests indicate that neither the Rule nor the Comparison group experienced significant progress from pre- to post-test ($F(39) = .562, p = .458, \eta^2 = .084$; and $F(39) = 3.57, p = .066, \eta^2 = .084$, respectively)

Error correction task considering stages (non-parametric test)

Figure 2 presents the distribution of the students in the Rule and Comparison group in terms of metalinguistic levels.

(Figure 1)

The results of the between-groups statistical analysis indicate that there were no significant differences in the pre-test ($U 197.5, Z -.381, p = .703$) or in the post-test ($U 189, Z -.605, p = .545$). However, within-groups analyses indicated that, whereas no significant improvement was registered for the Comparison group from pre- to post-test according to the the Wilcoxon Signed Rank test ($Z -1.0, p = .317$), significant progress did happen in the case of the students in the Rule group ($Z -2.12, p = .034$).

Research question 3: Metalinguistic instruction and oral performance

Non-parametric tests were performed in all cases, since oral performance was analyzed using scales, and thus ordinal variables (see Table 2). Figure 3 shows the percentage of students in each group.

(Figure 3)

According to the Mann-Whitney U Test, there were no significant differences between the groups in the pre-test (U 160, Z -1.80, $p = .071$), although the p value tended towards significance in favour of the Rule group. In the post-test, the performance of the learners in the Rule group was significantly more advanced than that of their peers in the Comparison group (U 133.5, Z -2.46, $p = .014$).

When analyzing differences within groups, Wilcoxon Signed Rank Tests indicated that none of the groups under study progressed significantly from pre- to post-test (Rule: Z -.998, $p = .318$; Comparison Z -1.02, $p = .309$)

Research question 4: Relationship between metalinguistic knowledge and gains in performance.

Correlations were made between the gains experienced from pre- to post-test (this value corresponds to the residuals of regressing post-test on pre-test scores) and the levels obtained in the metalinguistic judgment task (which are assumed to be a reflection of the students' metalinguistic knowledge). Parametric correlations were performed for the error correction task and non-parametric correlations for the oral production task, as the

variable was interval in the former case and ordinal in the latter. These correlations were significant for the Rule group considering gains in both, the error correction task ($r = .690$, $p = .001$) and the oral production task ($\rho = .646$, $p = .002$), but not for the Comparison group ($r = .260$, $p = .255$; and $\rho = .154$, $p = .493$ respectively).

Discussion and conclusion

According to the results of this study, it can be claimed that metalinguistic instruction has a slightly positive effect on metalinguistic knowledge on the one hand, and performance in metalinguistic and oral production tasks on the other. However, metalinguistic instruction cannot be considered to have a significant impact on students' knowledge and performance.

First of all, the results of the metalinguistic judgment task demonstrated that there were more students in the Rule group that were in stage 4 according to the codes developed by White and colleagues (see Figure 1) than in the Comparison group (43% vs. 18.2%); nevertheless, the difference between the two groups in terms of metalinguistic knowledge was not significant, although it approached significance ($p = .076$). Second, the analyses of the performance in the error correction task only suggested a significant advantage for the Rule group in terms of progress from pre- to post-test considering metalinguistic stages (low-mid-high). The rest of the analyses did not show any significantly different results between or within groups. Finally, a significant advantage of the Rule group was reported in the post-test in oral production. Nevertheless, such results should be taken cautiously, because the difference between the two groups in the

pre-test was approaching significance ($p = .071$). Additionally, when analyzing the progress experienced from pre- to post-test, no significant differences were found for the students in the Rule group. The correlations between performance gains and knowledge of the rule (as demonstrated in the metalinguistic judgment task) indicate that the learners in the Rule group who benefitted from metalinguistic instruction and assimilated the rule that they were taught were also capable of using such knowledge in their performance in the error correction and the oral production task. In contrast, those learners who did not show a good knowledge of the rule in the metalinguistic judgment task were less capable of making significant gains in their performance.

Considering these results, it can be said that the treatment that included metalinguistic instruction on the English PDs for six weeks had a modest impact. The reasons behind this outcome may have been diverse. The lack of highly significant differences between the knowledge and performance of the Rule group as compared to the Comparison group may be related to the fact that all the students had already been taught the rule in previous years. Consequently, the treatment that the Rule group received could have been useful just for reactivating previously acquired knowledge. This could be one reason why no such clear advantages were found for this group. Additionally, just by doing the pre-test, the students in the Comparison group were practising with the target feature, and this practice might have contributed to the noticing of the possessive determiners and thus a progress in performance, or an activation of knowledge was facilitated.

On the other hand, the lack of significant progress from pre- to post-test reported in some analyses that include the Rule group can be explained by different factors. First of all, the treatment might have been too short (30 minutes a week, over six weeks), or not

intensive enough. Another reason for the present results can be that these learners were not cognitively mature enough yet for them to be able to benefit more fully from metalinguistic instruction (DeKeyser, 2000; 2003; DeKeyser & Larson-Hall, 2005). In fact, White, Muñoz, and Collins (2007) report more robust advantages in the case of teenage students following the same treatment used in the present study. In the research by White and colleagues, both Canadian and Spanish learners in grade 8 receiving metalinguistic instruction on English PDs demonstrated a significant improvement from pre- to post-test and significantly outperformed their peers who were not instructed on these forms.

Nevertheless, the study by White and Ranta (2002) reported significant gains in students' knowledge and performance in metalinguistic and oral tasks in the case of 11-12-year-old children. One important difference between the participants in White and Ranta (2002) and the participants included in the present study concerns the number of hours of instruction. Whereas the students in the present investigation received a maximum of 4.5 hours of English instruction a week, the participants in White and Ranta (2002) received over 4.5 hours a day. It is quite likely that the English PDs were used by these learners, or included in the input they were exposed to, more frequently than in the case of the Spanish students. It has been suggested that for metalinguistic instruction to become internalized and 'proceduralized' massive amounts of practice are necessary (DeKeyser, 2007). While the learners in the Canadian context had a chance to continue practicing what they had learnt in the instructional treatment through approximately 20 hours of weekly exposure to the L2, such practice was not so clearly facilitated for the

students in the Spanish context. White (2008) also suggests that changes in students' performance are not very likely when the hours of exposure to the L2 are limited.

Another difference, which is probably quite remarkable and might have had an effect on the results obtained in the two studies, is the novelty of the treatment. For the students in the Spanish context, instruction that focuses on grammar forms or metalinguistic explanations on target L2 features does not constitute an innovative treatment, because most of these students' EFL classes follow this approach. For the Canadian students, on the other hand, who were following communicative-based instruction, metalinguistic explanations were rare, and the novelty of the treatment might have contributed to raising their interest and motivation. Moreover, notwithstanding this treatment, the students enrolled in intensive English courses in Canada can be said to be highly motivated to learn the L2, since these courses are optional and require a high degree of commitment on the part of the students (White, 2008).

Finally, an important fact that can explain the results reported in this study is that, as has already been suggested, 'metalinguistic instruction is not for everybody'. Metalinguistic knowledge has been associated with language analytic ability both in the case of children and adults, and such skill (also considered a part of language aptitude) is known to be subject to individual variability (Ranta, 2002; Roehr, 2007). It could be the case that for some children in the Rule group metalinguistic instruction was helpful because their analytic skills were quite developed, but for others whose language-analytic ability was not as high, this type of instruction did not promote L2 acquisition. In fact, the correlations between metalinguistic knowledge and performance within the Rule group indicates that it was only those students whose knowledge of the rule was advanced that

made the most gains in performance. Indeed, even though all the students received the same type of instruction, only some benefitted from it, and those who did demonstrated performance gains.

Apart from analytic ability, the effect of metalinguistic instruction has also been claimed to depend on learners' L2 proficiency. In her analysis of the different studies examining the acquisition of the English PDs by different groups of children and teenagers (White, 1996; 1998; White, Muñoz, & Collins, 2007; White & Ranta, 2002), White (2008) suggests that learners' performance concerning English PDs in the pre-test usually determines the progress they experience in the post-test. White (2008) claims that the learners that are at an emergence stage in the pre-test are more likely to benefit from instruction than those who are at a pre-emergence stage, especially if instruction is implicit. Proficiency has also been shown to be related to the effect of implicit/explicit feedback. Ammar and Spada (2006) found that, while high-proficiency learners benefitted equally from implicit and explicit feedback, low-proficiency learners improved significantly more after receiving explicit feedback than when such feedback was implicit. Although it was not one of the research questions in the present study, correlations were performed between gains in the error correction and in the oral production task and proficiency (as defined by performance in the cloze test and performance in the pre-test in the two tasks under analysis). Significant correlations were found between these three different measures of initial proficiency and gains in the oral production task for the Rule group but not for the Comparison group. The evidence from the Canadian studies and from the present study thus suggests that different variables such as L2 proficiency or analytic ability might predict the degree to which metalinguistic instruction would be

advantageous for L2 learners. As White (2008) claims, learners have to be 'ready' in order to benefit from metalinguistic instruction.

The participants included in the present study were bilingual and were learning English as a third language, as was also the case for two of the groups examined by White, Muñoz, and Collins (2007), but different from the groups analysed by White and Ranta (2002). Whether bilingualism had an effect on the learning of PDs is hard to determine, as this variable is confounded with many others, such as age or programme type. What can be said is that the bilinguals in White, Muñoz, and Collins (2007) and those included in the present study behaved differently: the former seemed to benefit more from metalinguistic instruction than the latter, but they were also older and they were enrolled in a different programme type.

In conclusion, and addressing the issue raised by the title of this article, the results of this study suggest that metalinguistic instruction does not necessarily translate into metalinguistic knowledge on the part of the students, but metalinguistic knowledge can certainly affect performance in a positive way, not only in controlled tasks (error correction) but also in free production oral tasks. Moreover, in light of the findings of the present study and other studies that have dealt with this topic (especially White & Ranta, 2002, and White, Muñoz, & Collins, 2007), it can be concluded that the positive effect of metalinguistic instruction is subject to individual variability, with factors such as students' analytic skills, age or motivation being highly influential. Moreover, programme type can also be considered a variable that can potentially have an impact on the effect of a particular treatment in L2 classes: explicit explanations of grammar rules might be more effective within an approach that essentially focuses on meaning. Additionally, the

amount of hours of practice and input that the students are allowed significantly affects the results that L2 classroom instruction can have, and these variables have already been claimed to be significant in L2 acquisition (Serrano, in press). More studies should be performed in order to closely examine how these factors affect L2 acquisition so as to better cater to L2 classroom learners' needs.

References

- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28, 543-74.
- Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition*, 16, 157-168.
- DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17, 379-410.
- DeKeyser, R. M. (2000). The robustness of critical period effects in language acquisition. *Studies in Second Language Acquisition*, 22, 4, 499-533.
- DeKeyser, R. M. (2003). Implicit and explicit learning. In: C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Oxford: Blackwell Publishing.
- DeKeyser, R. M. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97-113). Mahwah, NJ: Lawrence Erlbaum Associates.

- DeKeyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. Long, & C. Doughty (Eds.), *Handbook of second language teaching* (pp. 119-138). Oxford: Blackwell.
- DeKeyser, R., & Larson-Hall, J. (2005). What does the critical period really mean? In J. Kroll & A. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 88-108). Malden, MA: Blackwell.
- Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 256-310). Malden, MA: Blackwell.
- Ellis, R. (2003). Does form-focused instruction affect the acquisition of implicit knowledge? A review of the research. *Studies in Second Language Acquisition*, 24, 223-236.
- Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education*. Rowley, MA: Newbury House.
- Harley, B., & Swain, M. (1984). The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Cripser & A. Howatt. (Eds.), *Interlanguage* (pp. 291-311) Edinburgh: Edinburgh University Press.
- Langé, G. (2002) (Ed.). *TIE-CLIL professional development course*. Milan: TIE-CLIL.
- Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition*, 12, 429-448.

- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramersch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty and J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- Marsh, D. (2002). *CLIL/EMILE—The European dimension: Actions, trends and foresight potential*. Bruxelles: The European Union.
- McLaughlin, B. (1987). *Theories of Second Language Learning*. London: Edward Arnold.
- Muñoz, C. (1994). A case of frequency-based markedness. *Atlantis*, 41, 165-177.
- Muñoz, C. (2005). The development of the personal pronoun system in learners of English. Paper presented at EUROSLA 2005 Conference, Dubrovnik, 14th-17th September.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528.
- Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 23-75). Clevedon: Multilingual Matters.
- Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics*, 10, 217-244.
- Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (Ed.) *Individual differences and instructed second language acquisition*. Amsterdam: John Benjamins.

- Roehr, K. (2007). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics*, 29, 173-199.
- Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggensbach (Ed.), *Perspectives on Fluency* (pp. 200-219). Ann Arbor: University of Michigan Press.
- Serrano, R. (in press). The time factor in EFL classroom practice. *Language Learning*, 61.
- Spada, N., & Lightbown, P. M. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *The Modern Language Journal*, 83, 1-22.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge University Press.
- White, J. (1996). An Input Enhancement Study with ESL Children: Effects on the Acquisition of Possessive Determiners. PhD Dissertation, McGill University, Montreal.
- White, J. (1998). Getting learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language Acquisition* (pp. 85-113). Cambridge: Cambridge University Press.
- White, J. (2008). Speeding up the acquisition of *his/her*: Explicit L1/L2 contrasts help. In J. Philp, R. Oliver, and A. Mackey (eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 193-228). Amsterdam: John Benjamins.

- White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness, 11*, 259-290.
- White, J., Muñoz, C., & Collins, L. (2007). The *his/her* challenge: Making progress in a 'regular' L2 programme. *Language Awareness, 16*, 278-299.
- Zobl, H. (1985). Grammars in search of input and intake. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 329-344). Rowley, MA: Newbury House.

Table 1. Codes for the error correction passage

Metalinguistic Level	Criteria
Low	Below 50% accuracy in correcting errors for <i>his</i> and <i>her</i>
Mid	50-70% accuracy in correcting errors for <i>his</i> and <i>her</i>
High	Above 75% in correcting errors for <i>his</i> and <i>her</i> Above 50% accuracy in kin-different contexts

Table 2. Codes for the Oral Production Task

Pre-emergence	
Stage 1	Pre-emergence: avoidance of <i>his</i> and <i>her</i> (0-1 correct uses, 1-2 incorrect uses) and/or use of definite article.
Stage 2	Pre-emergence: use of <i>your</i> (minimum of 2 times) for all persons, genders and numbers; 0-1 correct uses of <i>his</i> and <i>her</i> .
Emergence	
Stage 3	Emergence of either or both <i>his/her</i> : 2-6 combined total correct uses of <i>his</i> and <i>her</i> , neither to criterion (4 correct uses)
Stage 4	<p>Preference for <i>his</i> or <i>her</i></p> <p>Preference for <i>his</i>: use of <i>his</i> to criterion (4 correct uses); probably accompanied by overgeneralization of <i>his</i> to contexts for <i>her</i>; 0-3 instances of <i>her</i></p> <p>Preference for <i>her</i>: use of <i>her</i> to criterion (4 correct uses); probably accompanied by overgeneralization of <i>her</i> to contexts for <i>his</i>; 0-3 instances of <i>his</i></p>
Post-emergence	
Stage 5	Differentiated use of BOTH <i>his</i> and <i>her</i> without agreement rule: differentiated use of both <i>his</i> and <i>her</i> to criterion (4 correct uses); below criterion (0-1 correct uses) with kin different gender for <i>his</i> and <i>her</i>
Stage 6	Agreement rule applied to <i>his</i> or <i>her</i> (kin different gender): differentiated use of both <i>his</i> and <i>her</i> to criterion (4 correct uses);

	agreement rule applied to kin different gender to criterion (2 correct uses) for either <i>his</i> or <i>her</i>
Stage 7	Agreement rule applied to <i>his</i> and <i>her</i> (kin different gender): differentiated use of both <i>his</i> and <i>her</i> to criterion (4 correct uses); agreement rule applied to kin different gender to criterion (2 correct uses) for both <i>his</i> and <i>her</i> ; errors with body parts may continue
Stage 8	Error-free application of agreement rule: rule applied to <i>his</i> and <i>her</i> (all domains, including body parts)

Table 3. Codes for the metalinguistic judgment task

Level 1	<p>Irrelevant information that focuses on some other feature (e.g. student says there is something wrong with the noun).</p> <p>Wrong information about the PD (e.g. <i>his</i> = singular, <i>her</i> = plural), or completely backwards (e.g. <i>his</i> = feminine, <i>her</i> = masculine)</p> <p>Student says nothing about possession, gender</p>
Level 2	<p>Student appears to be operating with the Catalan rule</p> <p>Or explanation indicates confusion, but may have the idea of possession (e.g. at someone, 's ____)</p> <p>Some right and some wrong referents</p>
Level 3	<p>Information in explanation is mainly correct (1 incorrect referent allowed)</p> <p>Explicitly refers to gender distinction</p> <p>No attempt at a rule of thumb or possession</p>
Level 4	<p>Fluent, little prompting for explanation</p> <p>All information in explanation is clear, correct</p> <p>Refers to gender distinction</p> <p>Refers to rule of thumb and/or possession</p> <p>All referents referred to are correct</p>

Table 4. Descriptive statistics error correction /16

	Pre-test (/16)	Post-test (/16)
Rule Group	7.10 (4.70)	7.70 (6.02)
Comparison Group	5.24 (4.65)	6.71 (4.20)

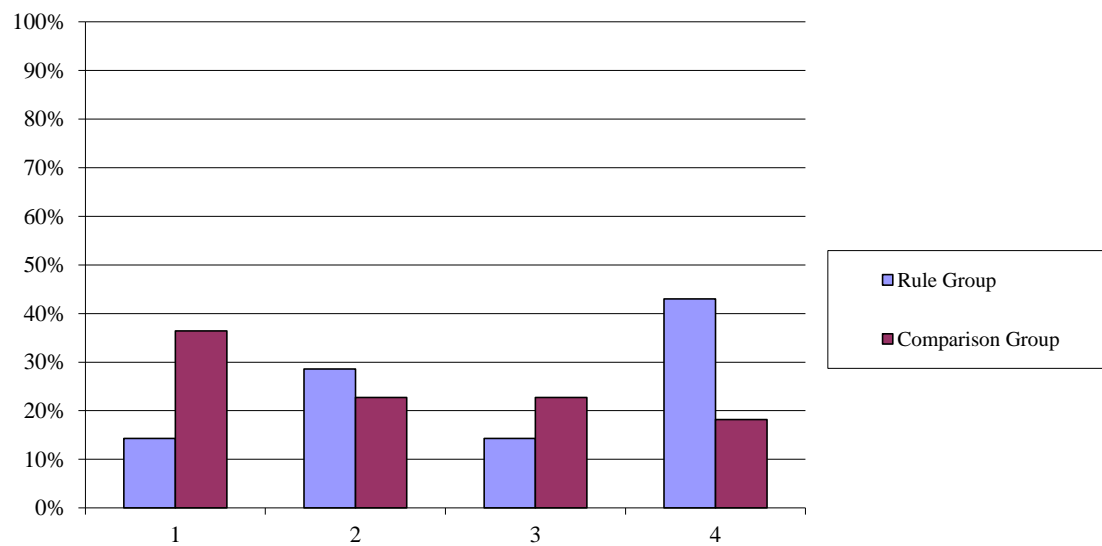


Figure 2. Metalinguistic judgment task

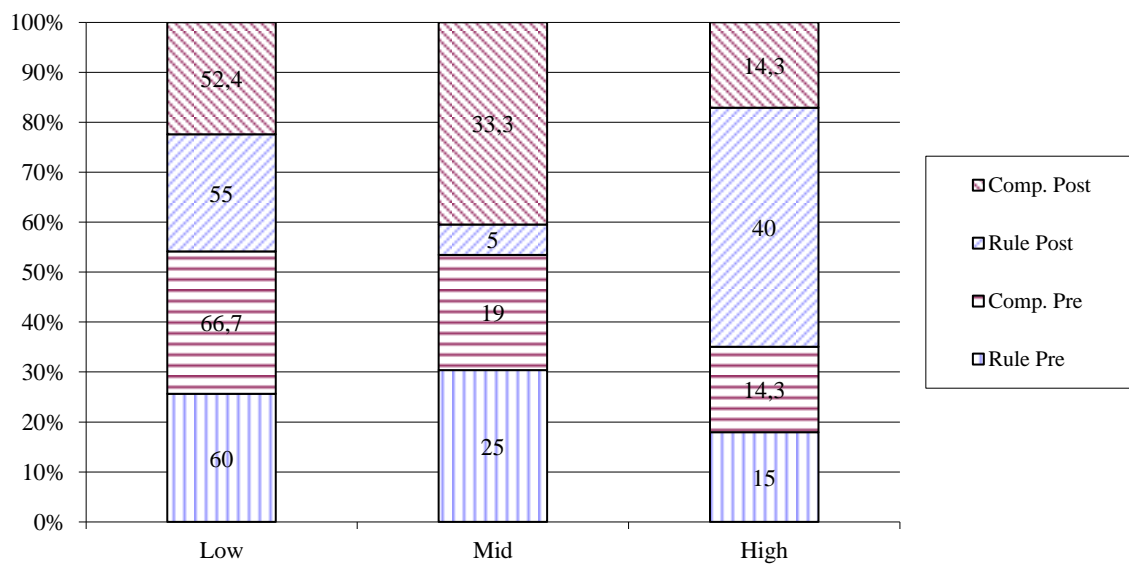


Figure 3. Students according to metalinguistic level (%)

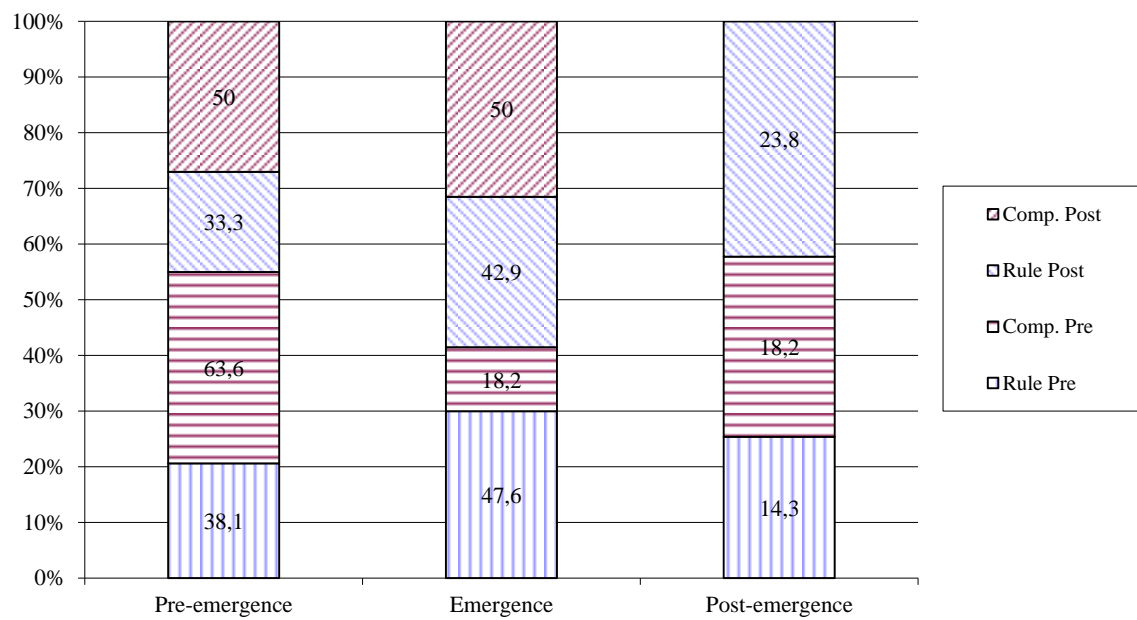


Figure 4. Oral performance (%)