
COMPARING THE EFFECTS OF TWO TYPES OF OUTPUT-PUSHING FEEDBACK ON ADULT EFL STUDENTS' ORAL ACCURACY

UNA COMPARACIÓN DE LOS EFECTOS DE DOS TIPOS DE RETROALIMENTACIÓN CORRECTIVA EN LA PRODUCCIÓN ORAL DE LOS ALUMNOS ADULTOS DE INGLÉS.

Hanne Roothoof

Universidad Pública de Navarra, Spain

hanne.roothoof@unavarra.es

DOI: <http://dx.doi.org/10.12795/elia.2016.i16.03>

Despite growing evidence that providing immediate oral corrective feedback can help second language students speak more accurately, the way in which feedback should be given remains controversial. While a great number of studies support the use of recasts, authors such as Lyster and Saito (2010) claim that output-pushing feedback could be more beneficial for classroom learners. However, few studies have attempted to tease apart the effects of separate output-pushing feedback-types or prompts, which is why the present small-scale study compares two

kinds of prompts; metalinguistic feedback and elicitation. Two groups of intermediate university students ($n=31$) received either elicitation or metalinguistic feedback while carrying out communicative speaking activities. A statistical analysis of students' gain scores from pre- to post-test shows that metalinguistic feedback had greater effects on students' accuracy, although these were not significant. Metalinguistic feedback was also much more successful at helping students correct their errors during the classroom intervention.

Key words: oral corrective feedback, EFL, speaking, metalinguistic feedback, elicitation.

Aunque numerosos trabajos han demostrado que la retroalimentación oral inmediata puede ayudar a los alumnos de una segunda lengua a hablar más correctamente, no hay acuerdo sobre qué tipo de corrección es más adecuado. Si bien un gran número de estudios apoyan el uso de reformulaciones, autores como Lyster y Saito (2010) afirman que, en el contexto escolar, puede resultar más beneficioso proporcionar a los estudiantes los denominados "prompts", que constituyen un tipo de retroalimentación que obliga a los alumnos a producir "output". A día de hoy, apenas se conocen los efectos de diferentes tipos de "prompts", por ello, este estudio a pequeña escala compara dos tipos: la retroalimentación metalingüística y la elicitación. Dos grupos de alumnos universitarios de nivel intermedio ($n=31$) recibieron o bien elicitación o bien retroalimentación metalingüística mientras efectuaron actividades orales comunicativas. Los resultados muestran que los alumnos que recibieron retroalimentación metalingüística mejoraron más entre el test previo y posterior, aunque esta mejora no fue significativa. Asimismo, estos alumnos fueron más capaces de corregir sus errores durante la intervención en el aula.

Palabras clave: retroalimentación oral, enseñanza del inglés como lengua extranjera, destreza oral, retroalimentación metalingüística, elicitación.

1. Introduction

The question of how to provide feedback to second language learners on their spoken performance is an issue of immediate importance to classroom teachers that has given rise to considerable controversy among researchers. A growing body of research carried out within the interaction approach to second language acquisition (SLA) supports the use of immediate oral corrective feedback (CF) during communicative speaking tasks, as several meta-analyses have found positive and durable effects for CF on second language acquisition (e.g., Li, 2010; Mackey & Goo, 2007; and Russell & Spada, 2006). The interaction approach to SLA (Gass & Mackey, 2015) is based on a hypothetical link between interaction and language development, since interaction provides learners with modified input and feedback on their output. Despite these positive findings on the usefulness of interaction and feedback, researchers continue to debate the best method of providing CF, either implicitly in the form of recasts, or by using more explicit techniques. Most studies showing positive effects for recasts have been carried out in a laboratory setting, where CF is provided during one-on-one interaction between a native speaker and a learner (e.g., Ishida, 2004; Leeman, 2003; Loewen & Nabei, 2007; Mackey & Philp, 1998; McDonough & Mackey, 2006). In classroom settings, observational studies almost invariably conclude that teachers predominantly use recasts to deal with their students' spoken language errors (e.g., Sheen, 2004). However, evidence from classroom studies indicates that students may benefit more from being pushed to self-correct by means of prompts or output-pushing CF (Lyster & Saito, 2010). While several studies have compared the impact of recasts to that of a group of prompts (e.g., Ammar & Spada, 2006; Dilans, 2010; Lyster & Izquierdo, 2009; Lyster & Saito, 2010; Nassaji, 2007), very few studies have attempted to compare separate prompts to each other, even though prompts comprise a range of CF-types which may well affect language development differently. To fill this gap in the literature, the present study aims to compare the effects of two types of prompts or output-pushing CF, elicitation and metalinguistic feedback, on EFL students' ability to use the English past simple tense in oral production.

2. Literature Review

2.1. Research on Different Feedback-types

When trying to account for the effects of various oral CF-types, researchers have classified CF-moves based on either their degree of explicitness, or on whether they contain the correct form. Accordingly, recasts are generally classified as implicit (even though they can be more explicit if they are reduced or stressed, see for instance Sheen, 2006) and input-providing, whereas prompts are all output-pushing techniques varying in explicitness. The most explicit type of prompt is metalinguistic feedback, while elicitation can be classified as more implicit. Table 1 provides examples of implicit and explicit input- and output-providing CF in response to an example of a learner error in the use of the past tense. It needs to be noted that degree of explicitness is a continuum and that elicitation is classified as explicit by some authors (Lyster, Saito & Sato, 2013).

Student error: <i>Yesterday I go to the cinema.</i>	I n p u t - p r o v i d i n g (reformulations)	O u t p u t - p u s h i n g (prompts)
Implicit feedback	Recast: <i>Oh, you went to the cinema.</i>	Elicitation: <i>Last weekend, I...</i>
Explicit feedback	Explicit correction: <i>Not go, went.</i>	Metalinguistic feedback: <i>You need past tense.</i>

Table 1: Classification of CF-types based on Lyster et al., 2013

Input-providing CF, particularly in the form of recasts, is hypothesized to foster language acquisition because it provides a brief opportunity for the student to focus on form while remaining focused on the meaning of the message (Long, 1991). Since recasts immediately follow the learner's erroneous utterance, and express the learner's original meaning, researchers (e.g., Goo & Mackey, 2013) have suggested that this type of feedback allows learners to compare the correct form with the error they have just made.

On the other hand, instructor-provided prompts can be expected to promote language acquisition based on Swain's (1995, 2005) Output Hypothesis, which posits that pushing learners to express themselves more precisely and more accurately may lead to interlanguage development. According to de Bot (1996), actively looking for the right form, rather than receiving it from the teacher or interlocutor, strengthens form-meaning connections in learners' minds, and thus helps these learners proceduralize their declarative knowledge.

Many researchers reporting substantial support for the effectiveness of recasts do so by comparing the group receiving this type of CF to a control group receiving no CF, and by conducting their studies in a laboratory—rather than in a classroom—setting (e.g., Han, 2002; Iwashita, 2001; Leeman, 2003; Mackey & Philp, 1998). As mentioned in the introduction, Lyster and Saito's (2010) meta-analysis of 15 classroom studies found larger effects for prompts than recasts. However, the effects of CF are not independent of factors such as the learners' proficiency or the nature of the target structure. For instance, Yang and Lyster (2010) observed that recasts were equally effective for the acquisition of the regular and irregular past tense, but prompts had greater effects on students' accurate production of the regular past. Regarding proficiency, Ammar and Spada (2006) found that recasts benefited high-proficiency learners and low-proficiency learners equally, but prompts had greater effects on the acquisition of English possessive pronouns by low-proficiency learners.

Rather than opposing recasts to prompts, some authors have compared recasts to metalinguistic feedback or explicit correction in order to address the question whether implicit or explicit CF is more effective. It has been found that recasts, due to their implicit nature, are not always interpreted as corrective by learners (Lyster, 1998), especially in the case of morphosyntactic errors (Kim & Han, 2007; Mackey, Gass & McDonough, 2000). Ellis, Loewen and Erlam (2006) compared the effects of recasts to metalinguistic feedback on the acquisition of the past simple tense and found larger effects for the more explicit feedback. Another study by Sheen (2007) showed greater benefits for explicit feedback than for recasts on the acquisition of English articles.

There are thus some indications that prompts are useful for helping learners acquire certain grammatical structures in a classroom context. At the same time, there is evidence that more explicit CF has a greater impact on learners' accuracy than implicit correction techniques. Although prompts encompass a wide range of techniques, both implicit and very explicit, hardly any research has compared prompts with each other. While some studies have focused on metalinguistic feedback (e.g., Ellis, 2007; Ellis et al., 2006; Loewen & Nabei, 2007), to the best of our knowledge no studies have dealt with elicitation as a separate treatment condition. Elicitation is nonetheless used by teachers, and appears to be very successful at pushing learners to reformulate their ungrammatical utterances (Lyster & Ranta, 1997; Sheen, 2004). To contribute to the scarce body of literature on this issue, the present study compares the effects of two types of prompts, metalinguistic feedback and elicitation, on students' oral production of the English simple past tense.

Several studies have shown that the past tense can be influenced positively by CF (Doughty & Varela, 1998; Ellis et al., 2006; Han, 2002; McDonough, 2007; Takashima & Ellis, 1999; Yang & Lyster, 2010). Moreover, the past tense lends itself well to the use of prompts, since students at the intermediate or higher proficiency level should already have some knowledge of the structure and, therefore, may be able to retrieve the correct form when pushed for output by their interlocutor.

2.2. The Target Structure: the Regular and Irregular Past Simple Tense

Even though English text books such as *New English File* (Oxford), *Total English* (Longman) or *Face to Face* (Cambridge) introduce the past simple tense at the elementary or A1 level, it continues to be a problematic area for learners as they progress towards more advanced levels of English, especially in oral production.

Morpheme studies such as Dulay and Burt (1974) or Larsen-Freeman (1975) show that the past tense is acquired late in comparison with other morphemes such as progressive *-ing* or articles.

To explain why certain morphemes appear to pose greater difficulty to learners than others, Goldschneider and DeKeyser (2001) note that perceptual salience is one of the factors which may influence the order of acquisition of these morphemes. In the case of the regular past tense, the regular *-ed* ending is perceptually non-salient, because it is usually realized as an alveolar stop (either voiceless [t] or voiced [d]) and as such does not constitute a syllable. Apart from its lack of saliency, the past tense ending is also often redundant, as both contextual information and adverbial phrases such as “yesterday” or “last week” can be used to interpret a sentence as past (VanPatten, 1996).

This means we can expect learners to make errors when using the past tense. As stated by Ellis et al. (2006: p.251), “the typical error made by learners is the use of the simple or present form of the verb in place of *V-ed*: *Yesterday I visit my sister.” Moreover, in case of irregular verbs, learners may overgeneralize the *-ed* ending to irregular past tense forms.

2.3. Research Questions

Two main research questions underlie the present study. The first question pertains to the immediate impact of these types of oral CF, which was studied by looking at the rates of uptake and repair following the teacher’s CF. Uptake refers to the learner’s response to a teacher CF move and indicates that the learner has noticed the CF. Repair refers to the student’s successful correction of his or her initial error, following the teacher’s CF. The second question is related to the effects of the two types of CF on learners’ subsequent language production on an immediate post-test.

1. Do elicitation and metalinguistic feedback on learners’ past tense errors during classroom communicative activities differ in terms of rates of learner uptake and repair?
2. Are there pre- to post-test gains in learners’ accurate oral production of past tense forms and is there a statistically significant difference between the elicitation group and the metalinguistic feedback group?

3. Methodology

3.1. Participants

The participants were 31 undergraduate students, 5 male and 26 female, from two intact classes (defined by Porte, 2010, p.41 as “classes to which the students have been assigned prior to the study itself”) in the first year of a Primary Teaching Degree at a Spanish university, who had been placed in an intermediate English group based on in-house placement tests. Participants were between 19 and 23 years old, with a mean age of 19.8. They reported having learned English for between 8 and 18 years (13.7 years on average). All of them were native speakers of Spanish. Their English classes took place twice a week for one hour and a half per lesson and they used the text book *Ready for PET*. The lessons consisted of a combination of skills-work and grammar and vocabulary exercises, but relatively little time was devoted to speaking. They did not receive any instruction on the past simple tense during the time of the study, but there is no doubt they did receive explicit instruction on how to form the past tense in English in previous English classes, since all of them had received several years of grammar-oriented instruction in the past. The participants had very limited exposure to English outside of the classroom and most of them had not spent time in an English-speaking country, as they indicated on a background questionnaire they completed during the first session of the experiment. Before the experiment, the participants were informed that they would take part in a study which would allow them to practice their speaking skills. They all signed a consent form in which they agreed to being audio-recorded. Even though 31 students took part in the study, only 16 students attended all four sessions, one of whom was excluded because she had scored zero on the grammaticality judgment pre-test. This means the data related to pre- and post-test gains are only based on 5 students for the metalinguistic feedback group and 10 students for the elicitation group.

3.2. Design

The study took place during the students’ normal class time and the same researcher carried out the tasks with both groups. Four sessions of about 1

hour each, spread over two consecutive weeks, were dedicated to the study. A pre-test/post-test design was used and the classes were randomly assigned to a treatment condition, resulting in a metalinguistic feedback group and an elicitation group. As intact classes were used, there were slightly more students in the elicitation group ($n=18$) than in the metalinguistic feedback group ($n=13$). Figure 1 illustrates the design.

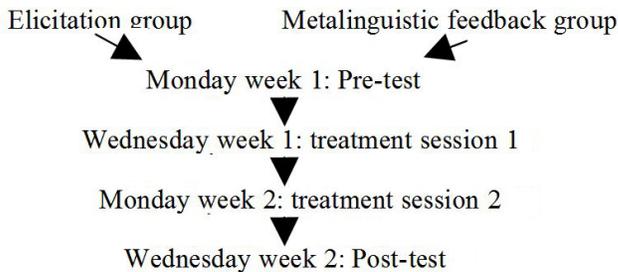


Figure 1: Design of the study

3.3. Procedure and Materials

Table 2 provides an overview of the different materials and tasks used during the study. To test the students' explicit knowledge of the past tense, along with their ability to produce the structure in speaking, they were given a grammaticality judgment test and an oral story-telling test prior to the treatment. As table 2 shows, a similar procedure was followed for the posttest.

Week 1 session 1	Pretest	Background questionnaire Grammaticality judgment test Storytelling test: Story 1: 'The Birthday Party' Story 2: 'The Banana'
Week 1 session 2	Treatment 1	Telling a fairytale in groups, listeners put the pictures in the right order 10 minutes: read the story 5 minutes: prepare for telling part of the story based on keywords and pictures Stories: 'Rumpelstiltskin' 'Cinderella' 'Rapunzel'
Week 2 session 3	Treatment 2	Telling a fairytale in groups, finding the differences between two versions of the same story 10 minutes: read the story 5 minutes: prepare for telling part of the story based on the keywords and pictures Stories: 'Puss in Boots' 'Little Red Riding Hood'
Week 2 session 4	Posttest	Grammaticality judgment test Storytelling test: Story 1: 'The Birthday Party' Story 2: 'The Banana' Story 3: 'The Little Mouse' Exit questionnaire

Table 2: Overview of procedure and materials

The grammaticality judgment test consisted of 15 sentences containing a past tense error and 25 distractor sentences. The following is an example of a test item (see figure 2).

On my last holiday I take a lot of photos.	
Correct	→ Did you use a <i>rule</i> or <i>feel</i> ? _____
Incorrect	→ correction: _____

Figure 2: Example of grammaticality judgment test item

Five of the 15 sentences targeting the past tense contained regular verbs, while the 10 other sentences focused on irregular verbs. These irregular verbs were selected because they are among the most frequent ones and they occur several times in the treatment materials. The following verbs were used: come, tell, say, get, see, take, can, have, give and go.

For the oral test, the same two picture stories were given to the students on the pretest and the posttest. These were adapted from the book *Do and Understand* by Gerngross and Puchta (1996). Each story consisted of 12 pictures that together formed an easily recognizable storyline. The students recorded themselves telling these stories at a computer lab using the free software 'Audacity'. They received some vocabulary they might need and were given the first sentence of the story in order to create a past tense context, for example 'Yesterday Joe went to the market'. On the posttest, as can be seen in table 2, a third story was added in order to investigate whether any improvement found was not simply due to the fact that the students had told the same stories before, in other words, to avoid so-called practice effects.

For both treatment sessions, the students had to tell parts of fairytales to the rest of the class in groups of four or five, after having read a written version of the story which was replaced by pictures and keywords, as in the example in figure 3. First, each group received the written version of their fairytale. They were given 10 minutes to read it and discuss it with their group mates, to make sure they understood the contents. Then they were told to each choose one paragraph of the story,

which they would have to tell to the rest of the class. They were given 5 minutes to practice telling their part of the story. Then the texts were taken away and replaced by pictures and keywords. Each student thus had to tell part of their fairy tale, so that all the students had a chance to be corrected by the teacher.



a long time ago

miller to king: “my daughter can spin straw into gold”

Figure 3: Example of treatment material

Two different communicative tasks based on fairytales were designed, in order to make sure the students were attending to meaning, which is a prerequisite of form-focused instruction (see for instance Williams, 2005). As table 2 indicates, the first task asked students to tell a fairy tale to their classmates, who had to put a set of pictures in the right order. For the second task, two versions of the same fairytale were told by different groups, after which all students had to come up with five differences between those versions. At no point were the students informed about the focus of the study. The students carried out these tasks in the classroom in the presence of their classmates. The same procedure was followed for the two tasks, as can be seen in table 2.

During the treatment sessions, the teacher-researcher provided the students with the corresponding type of CF each time they made a mistake using the regular or irregular past tense.

3.4. Data Collection and Analysis

Both treatment sessions were audio-recorded and faithfully transcribed into English. The recordings from the oral tests were also transcribed. The data from the tests and treatment sessions were then analysed for accurate use of the past tense. Rates of accurate use were calculated following Bardovi-Harlig (2000), who proposes counting identical verb forms only once in order to prevent the results from being biased by the use of very frequent forms such as “was” or “went”. Instances of over-use, such as the use of the past tense instead of the infinitive in “didn’t went” or overgeneralization of regular morphology, as in “goed”, were counted as errors. Irregular verbs that have the same form in the past tense (e.g., “put-put”) were excluded from the analysis. Verb forms which were not pronounced correctly but which could still be recognized as past tense forms were counted as correct, for example when the *-ed* ending was pronounced as /əd/ in forms such as /ɑ:skət/, or in the case of “came” pronounced as /caem/.

To answer research question 1, the treatment sessions were analysed for instances of uptake and repair. This will be explained in more detail in the results section (see 4.1).

4. Results

4.1. Research Question 1: uptake and repair after the two types of prompts

To study students’ immediate reactions to oral CF, Lyster and Ranta (1997) proposed the concepts of “uptake” and “repair”. Uptake refers to the student’s utterance which follows the teachers’ CF and which contains some form of reaction to that CF. If there is no uptake, the student ignores the CF altogether and simply continues speaking about the topic, as in the following example from our study:

- (1) S: She didn’t find the name, then on the third day she... she arrive(d) a message of the little man, which said that he told tell hi- her...

T: He...?

S: ... a strange history

In case there is uptake, it is possible that the student's utterance is still in need of repair, as in (2):

(2) S: After this the... Cinderella hear the clock

T: In the past

S: heard, *no sí*, heard the clock

Uptake with repair, then, refers to the student's successful uptake of the CF. In the case of prompts, this means the student is able to self-correct his initial error, as is the case in example 3:

(3) S: Suddenly a little man go to the...

T: In the past

S: Went to the room

As mentioned in 2.1, for students to be able to repair their errors they already need to have a certain degree of explicit knowledge of the structure in question. The grammaticality judgment test administered at the pretest shows that this condition is satisfied in the case of the past tense, as can be seen in table 3. Recalling that the test in question contained 15 sentences focusing on the target structure, we can see that students were able to successfully correct between 6 and 15 past tense errors, with the elicitation group obtaining a slightly higher average score than the metalinguistic feedback group. However, an unpaired t-test ($\alpha=0.05$, $p=0.27$) showed that this difference was not statistically significant, which means the two groups' explicit past tense knowledge at the beginning of the experiment can be considered comparable. Even though more than the 15 students listed in table 3 carried out the pre-test, we took into account only the test scores of those students who took part in all four sessions of the study.

Student	Elicitation	Metalinguistic feedback
1	6/15	6/15
2	12/15	12/15
3	13/15	14/15
4	9/15	8/15
5	13/15	6/15
6	10/15	-
7	15/15	-
8	14/15	-
9	7/15	-
10	14/15	-
Mean	11.3/15 (SD 3.13)	9.2/15 (SD 3.63)

Table 3: Results of the grammaticality judgment test

Table 4 gives an overview of the analysis of the treatment sessions for both groups, containing the number of correct past tense forms, past tense errors, feedback moves and the instances of uptake and repair. As noted from table 4, the metalinguistic feedback group produced a much larger number of correct past tense forms and a considerably lower number of errors. Nonetheless, pretest scores for both groups did not differ significantly and the mean scores on the in-house placement tests were also similar in both groups, so their overall proficiency and their knowledge of the past tense were comparable at the beginning of the experiment. A possible explanation for this striking difference will be provided in the discussion. Turning to the instances of uptake and repair, we can see in table 4 that both types of prompts resulted in a reaction from the students in almost all cases. However, with regard to the instances of uptake with repair, metalinguistic feedback resulted in a repair rate of 83.1 per cent, while elicitation only led to repair about half of the time.

	Correct past tense forms	Past tense errors	CF	Uptake	Repair
Elicitation (n=18)	116	102	80	77 (96.3%)	38 (47.5%)
Metalinguistic feedback (n=13)	179	62	59	56 (95%)	49 (83.1%)

Table 4: Uptake and repair during the treatment sessions

The two types of CF are thus equally effective at prompting students to modify their initial utterance, but there seems to be an important difference in their ability to help students correct their past tense errors.

4.2. Research Question 2: Pre- to Post-test Effects of the Two Types of Prompts

While the analysis of uptake and repair provides useful information about students' reactions to CF in the context of a communicative task, researchers have warned against interpreting uptake as a sign of language development (e.g., Goo & Mackey, 2013). To find out if the two prompts had an impact on the students' ability to use the past tense accurately in oral tasks, we need to analyse their performance on the oral pre- and post-test.

Tables 5 and 6 contain the results of the story-telling tests of the elicitation and the metalinguistic feedback group, respectively. The results of the posttest are divided into the accuracy scores for the same two stories that were used in the pretest and the scores on the new story, which the students told for the first time in the posttest.

Student	Pretest		Posttest Same stories		Posttest Different story	
E1	1/15	7%	3/15	20%	4/8	50%
E2	13/22	59%	9/18	50%	6/14	42.9%
E3	7/18	38.9%	6/14	42.8%	4/8	50%

E4	7/11	63.6%	12/17	70.6%	6/9	66.7%
E5	7/14	50%	12/22	54.5%	10/13	76.9%
E6	6/23	26.1%	6/17	35.3%	5/8	62.5%
E7	7/20	35%	11/20	55%	7/11	63.6%
E8	10/18	55.6%	14/20	70%	6/11	54.5%
E9	3/15	20%	10/21	47.6%	7/11	63.6%
E10	1/16	6.3%	5/16	31.3%	5/10	50%
Mean score	36.15	(SD 20.9)	47.71%	(SD 16.1)	58.07%	(SD 10.3)

Table 5: Results of the storytelling tests for the elicitation group

Student	Pre-test		Post-test Same stories		Post-test Different story	
M1	4/10	40%	11/17	64.7%	5/9	55.6%
M2	1/14	7.1%	7/12	58.3%	7/13	53.8%
M3	10/16	62.5%	13/17	76.5%	9/11	81.8%
M4	7/16	43.8%	11/22	50%	5/12	41.7%
M5	1/16	6.3%	11/22	50%	6/11	54.5%
Mean score	31.94%	(SD 24.6)	59.9%	(SD 11.1)	57.48%	(SD 14.7)

Table 6: Results of the storytelling tests for the metalinguistic feedback group

An unpaired t-test ($\alpha=0.05$, $p=0.73$) comparing the pretest scores of the elicitation group to those of the metalinguistic feedback group showed no statistically significant difference between the two groups on the pretest. As seen before, the two groups were also found to be comparable on the grammaticality judgment scores.

Two paired t-tests ($\alpha=0.05$) comparing the pretest and posttest scores (same stories) in both groups showed both groups improved significantly after the treatment ($p=0.009$ for elicitation, $p=0.03$ for metalinguistic feedback). When comparing the pretest stories with the results of the new story in the posttest, the elicitation group also improved significantly ($p=0.01$), whereas the metalinguistic feedback group improved but the difference was not quite significant ($p=0.0572$). This indicates that the improvements made in accurate production of the past tense were not just due to practice effects or task repetition.

To answer the question whether these two types of feedback affect the acquisition of the past tense differently, gain scores were calculated by subtracting the pretest scores from the posttest scores (only taking into account the scores for same stories). These gain scores are displayed in table 7. Student E1, for instance, made gains of 13 per cent from the pretest to the posttest for the same stories, whereas student E2's score went down, which results in a negative score of -9 per cent. The mean gain score of the elicitation group was found to be 11.56, while the mean gain score of the metalinguistic feedback turned out to be somewhat higher, i.e. 27.96. An unpaired t-test comparing the gain scores of both groups resulted in a *p*-value of 0.0518, which means the difference is not quite statistically significant, but there are some indications that metalinguistic feedback might be more effective than elicitation for the accurate production of the past tense.

Student	Elicitation	Metalinguistic feedback
1	13%	24.7%
2	-9%	51.2%
3	3.9%	14 %
4	7%	6.2%
5	4.5%	43.7%
6	9.2%	-
7	20%	-
8	14.4%	-
9	27.6%	-
10	25%	-
Mean	11.560% (SD 10.9)	27.960% (SD 19.1)

Table 7: Comparison of gain scores in the elicitation and metalinguistic feedback group.

When interpreting these results, it is important to take into account that only 5 students could be studied for the metalinguistic feedback group, compared to 10 in the elicitation group. Any differences between the gains derived from these techniques could then be attributed to this disparity, and not to the nature of the technique themselves. Moreover, major differences

exist between the improvement of individual students, and some of them only improved slightly (for example students E4 and M4).

5. Discussion

The results of this small-scale investigation into the effects of CF on students' oral accuracy seem to be in line with previous research showing that CF provided in the context of communicative tasks helps students improve their accurate production of grammatical structures (e.g., Ammar & Spada, 2006; Ellis, 2007; McDonough, 2007; Yang & Lyster, 2010). As we have seen, both the students receiving elicitation and those receiving metalinguistic feedback significantly improved their accurate production of the English past tense during storytelling activities.

However, as discussed in section 2.1, output-pushing CF can take many forms, which is why we compared the more implicit elicitation with the more explicit metalinguistic feedback. Indeed, we found that these two types of prompts appeared to affect students' past tense usage differently, as both the immediate and delayed effects of metalinguistic feedback turned out to be greater than those of elicitation, although no statistically significant difference was found for the gains made from pre- to post-test. Since both CF-techniques push learners to self-correct in a similar way, it can be assumed that the observed differences are related to the degree of explicitness of each CF-type.

Looking back to the analysis of uptake and repair, we note that students attempted to respond to almost all of the elicitation- and metalinguistic feedback-moves, which indicates that they noticed the feedback. Uptake following recasts, which takes the form of a repetition of the recast, has been found to be no indicator of subsequent acquisition (Mackey & Philp, 1998). However, uptake after output-pushing feedback usually results in modified output, which can play an important role in language development according to Swain's (1995, 2005) output hypothesis. In support of this hypothesis, McDonough (2005) concluded that the opportunity to modify one's speech was the best predictor of

acquisition in her study on the impact of CF on the acquisition of question forms.

Given that no differences were observed between both groups' past tense knowledge or oral production on the pre-test, it is striking that metalinguistic feedback resulted in a high rate of repair, whereas elicitation only resulted in the use of the correct form about half of the time. To understand this finding, we need to take a closer look at specific instances of learner uptake in both groups. In the case of elicitation, it appears that there were 22 cases in which students responded to the CF-move by simply repeating what they had said before, as in the example below.

(4) S: But she go to house

T: But she...?

S: But she go to house to 11 o'clock

The student in this example may have thought the teacher had not heard the utterance clearly, and thus the corrective intent of the elicitation-move seems to have gone unnoticed. Twenty-two of the 77 elicitation-moves leading to uptake simply contained a repetition of the initial form. In seven other cases, students modified their initial utterance, but they seemed to think the problem was related to vocabulary, as in (5):

(5) S: And a little man arrive...

T: A little man...?

S: Go

Here the student is simply using another verb, instead of the past tense of "arrive." Finally, there were two cases in which students were clearly at a loss as to the teacher's intention:

(6) S: When the prince climb up the tower...

T: And when the prince...?

S: What?

In the case of metalinguistic feedback, the teacher is obviously referring to a grammatical problem, as she explicitly states that the students need to use past tense. This may explain why there were only 6 cases of uptake without repair in the metalinguistic feedback group. In three of these cases, students seemed to misunderstand the intention of the teacher, as in example (7):

(7) S: She had one list of names but it's incorrect.

T: In the past?

S: don't?

There was also one case in which a student chose a different verb, as in example 5 in the elicitation group. The two other cases of uptake without repair in the metalinguistic feedback group occurred after irregular verbs and they resulted in overgeneralization of the regular verb rule, as in (9):

(9) S: After this the... Cinderella hear the clock

T: In the past

S: Heared, *no sí*, heard the clock.

Overgeneralization of the *-ed* ending to irregular verbs also happened once in the elicitation-group:

(10) S: want... to run he fall and he's die

T: he...?

S: he falled

Overall, it seems that the low rate of repair after elicitation is not due to a worse command of the past tense in the elicitation group, but rather to a lack of understanding of the teacher's intention.

The greater explicitness of metalinguistic feedback probably also explains the finding that students in the metalinguistic feedback group produced far fewer past tense errors than those in the elicitation group (see 4.1), even though there were no significant differences between the two groups on the pre-test. Since the other students were aware that the teacher was providing feedback on their classmates' past tense usage, it is likely that they started focusing harder on getting the past tense forms right.

Even though metalinguistic feedback was more successful at prompting students to self-correct their past tense errors during the storytelling tasks, this does not mean metalinguistic feedback is more effective at promoting long-term improvement or learning. As mentioned above, several authors have pointed out that uptake and repair do not predict subsequent acquisition, especially in the case of recasts (Goo & Mackey, 2013; Mackey & Philp, 1998). Indeed, notwithstanding the low rate of repair in the elicitation group, both the elicitation and metalinguistic feedback group made significant improvements in their accurate oral production of past tense forms on the post-test. On the whole, however, the students in the metalinguistic feedback group apparently improved more, although the difference was not quite statistically significant and the results are based on a small sample size.

6. Conclusion

The results of this small-scale study on the effects of output-pushing CF on students' accurate production of the past simple tense during storytelling tasks indicate that metalinguistic feedback may be more effective than elicitation. As mentioned in section 5, we attribute this finding to the more explicit nature of the former CF type. One possible reason why explicit prompts could work better than more implicit output-pushing CF is that it is easier for students to notice their corrective intent, since language learning cannot take place without a certain degree of noticing (Schmidt, 2001). Indeed, as shown in section 5, the students in this study did not always interpret elicitation as being directed at past tense usage.

Furthermore, it is possible that metalinguistic feedback was more successful with these students because of their relatively low level. As they were intermediate level students who had not had many opportunities of putting their theoretical knowledge into practice, the more explicit CF was probably more in line with their developmental level. Similarly, Aljaafreh and Lantolf (1994) posit that more explicit forms of CF could be more useful to low-level learners, while higher-level learners may be better able to regulate their performance in the foreign language and thus more implicit CF might be sufficient for them. Ammar and Spada (2006) also found prompts, which are generally more explicit, to be more effective for low-proficiency learners than more implicit recasts. Obviously, further comparative studies on the effects of prompt type on students at different levels of proficiency are warranted to corroborate these claims.

Another possible reason for the observed success of metalinguistic feedback can be related to skill-acquisition theory (e.g., DeKeyser, 2007), since this theory posits that explicit knowledge can be automatized and proceduralized through meaningful practice. The results of the grammaticality judgment pre-test show that the students generally had a fair level of explicit knowledge of past tense forms, which was activated through the storytelling activities and the teachers' prompts.

Even though this study shows explicit prompts to be more effective, Li (2010) and other researchers have suggested that the effects of implicit feedback take longer to manifest themselves. In Li's (2010) meta-analysis of 33 experimental CF-studies, the author found that recasts had greater effects on the acquisition of second language grammar on the delayed post-test, in comparison with more explicit feedback-types. Since this study only lasted two weeks and no delayed post-test could be performed, we were not able to verify whether this is also the case for implicit prompts, such as elicitation. Therefore, further research on the long-term effects of elicitation and other more implicit prompts needs to be carried out.

Another important caveat concerns the individual differences between students. As can be seen in the results section, not all of the students improved their oral production of the past tense in the same way.

Even though the students were of a comparable level and no significant group differences were found between their explicit knowledge of the past tense and their accurate production of the target structure on the pre-test, one of the elicitation-group students in fact decreased her accuracy score by 9%, whereas gain scores in the metalinguistic feedback group ranged from 6.2 to 51.2%. Such an important variation in students' performance cannot be ignored. To account for these differences, future studies need to take into account student factors such as proficiency, working memory and motivation, amongst others.

Finally, we acknowledge that this study has several limitations. First of all, it was a small-scale study, in which only a limited number of the initial participants were present throughout the entire experiment. Second, no control group in which students did not receive feedback was used and therefore it is possible that the observed pre- to post-test gains were merely due to the fact that students had the opportunity to practice the past tense during the treatment. Third, only one post-test was carried out three days after the second treatment session. As mentioned above, a delayed post-test could have given us more information about the long-term effects of the treatment.

Despite its limitations, we believe that this study provides important insights into students' reactions to two under-researched types of oral corrective feedback. Moreover, it proves that it is necessary to study different types of prompts separately, since differences were found between explicit and implicit prompts.

References

- Acklam, R., & Crace, A. (2005). *Total English*. Harlow, UK: Pearson Education.
- Aljaafreh, A., & Lantolf J.P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 87, 465-483. <https://doi.org/10.1111/j.1540-4781.1994.tb02064.x>

- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts and L2 learning. *Studies in Second Language Acquisition*, 28 (4), 543–574. <https://doi.org/10.1017/s0272263106060268>
- Bardovi-Harlig, K. (2000). *Tense and Aspect in Second Language Acquisition: Form, Meaning and Use*. Oxford, UK: Blackwell.
- de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529 – 555. <https://doi.org/10.1111/j.1467-1770.1996.tb01246.x>
- DeKeyser, R. (2007). Introduction. Situating the concept of practice. In R.M. DeKeyser (Ed.), *Practice in a second language. Perspectives from applied linguistics and cognitive psychology* (pp.1-18). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667275.002>
- Dilans, G. (2010). Corrective feedback and L2 vocabulary development: prompts and recasts in the adult ESL classroom. *The Canadian Modern Language Review*, 66 (6), 787-815. <https://doi.org/10.3138/cmlr.66.6.787>
- Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). New York, NY: Cambridge University Press.
- Dulay, H., & Burt, M. (1974) Natural sequences in child second language acquisition. *Language Learning*, 24, 37-53. <https://doi.org/10.1111/j.1467-1770.1974.tb00234.x>
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (ed.) *Conversational Interaction and Second Language Acquisition: A Collection of Empirical Studies*. (pp. 339–360). Oxford, UK: Oxford University Press.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit negative feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339 – 368. <https://doi.org/10.1017/S0272263106060141>
- Gass, S., & Mackey, A. (2015). Input, interaction and output in second language acquisition. In B. VanPatten & J. Williams (eds.). *Theories in Second*

- Language Acquisition. An introduction.* (pp. 180-206). New York, NY; London, UK: Routledge.
- Gerngross, G., & Puchta, H. (1996). *Do and Understand. 50 Action Stories for Young Learners.* Harlow, UK: Longman.
- Goldschneider, J.M., & DeKeyser, R.M. (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, 51 (1), 1-50. <https://doi.org/10.1111/1467-9922.00147>
- Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in second language acquisition*, 35, 127-165. <https://doi.org/10.1017/S0272263112000708>
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36 (4), 543–572. <https://doi.org/10.2307/3588240>
- Housen, A. (2002). The development of tense-aspect in English as a second language and the variable influence of inherent aspect. In R. Salaberry & Y. Shirai (eds.), *The L2 acquisition of tense-aspect morphology* (pp. 154-97). Amsterdam, The Netherlands: Benjamins. <https://doi.org/10.1075/lald.27.09hou>
- Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form of –te i (ru) by learners of Japanese as a foreign language. *Language Learning*, 54 (2), 311–394. <https://doi.org/10.1111/j.1467-9922.2004.00257.x>
- Iwashita, N. (2001). The effect of learner proficiency on interactional moves and modified output in nonnative-nonnative interaction in Japanese as a foreign language. *System*, 29, 267-287. [https://doi.org/10.1016/S0346-251X\(01\)00015-X](https://doi.org/10.1016/S0346-251X(01)00015-X)
- Kenny, N., & Kelly, A. (2007). *Ready for PET.* Oxford, UK: MacMillan Education.
- Kim, J., & Z. Han (2007). Recasts in communicative EFL classes: Do teacher intent and learner interpretation overlap? In A. Mackey (ed.): *Conversational Interaction in Second Language Acquisition: A collection of Empirical Studies.* (pp. 269–297). Oxford, UK: Oxford University Press.
- Larsen-Freeman, D. (1975). The Acquisition of Grammatical Morphemes

- by Adult ESL Students. *TESOL Quarterly*, 9, 409-419. <https://doi.org/10.2307/3585625>
- Leeman, J. (2003). Recasts and L2 development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25 (1), 37–63.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: a meta-analysis. *Language Learning*, 60 (2), 309-365. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Lightbown, P.M., & Spada, N. (2013). *How languages are learned*. 4th edition. Oxford, UK: Oxford University Press.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (ed.), *Conversational interaction and second language acquisition: A collection of empirical studies*. (pp. 361–377). New York NY: Oxford University Press.
- Long, M. (1991). Focus on Form: A Design Feature in Language Teaching Methodology. In K. de Bot, R. Ginsberg, and C. Kramsch (eds), *Foreign Language Research in Cross-Cultural Perspective* (pp. 39–52). Amsterdam, The Netherlands: John Benjamins. <https://doi.org/10.1075/sibil.2.07lon>
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51-81. <https://doi.org/10.1017/S027226319800103X>
- Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59 (2), 453–498. <https://doi.org/10.1111/j.1467-9922.2009.00512.x>
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37-66. <https://doi.org/10.1017/s0272263197001034>
- Lyster, R., & Ranta, L. (2013). Counterpoint piece: the case for variety in corrective feedback research. *Studies in Second Language Acquisition*, 35, 167-184. <https://doi.org/10.1017/S027226311200071X>
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA. A meta-analysis.

- Studies in Second Language Acquisition*, 32, 265-302. <https://doi.org/10.1017/S0272263109990520>
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language teaching*, 46, 1-40. <https://doi.org/10.1017/S0261444812000365>
- Mackey, A., Gass, S.M., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in second language acquisition*, 22, 471-497. <https://doi.org/10.1017/S0272263100004022>
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: a meta-analysis and research synthesis. In A. Mackey (ed.). *Conversational interaction and second language acquisition: a collection of empirical studies* (pp. 407-453). New York: Oxford University Press.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82 (3), 338-356. <https://doi.org/10.1111/j.1540-4781.1998.tb01211.x>
- McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in second language acquisition*, 27 (1), 79-103. <https://doi.org/10.1017/s0272263105050047>
- McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (ed.): *Conversational Interaction in Second Language Acquisition: A Collection of Empirical Studies* (pp. 323-338). Oxford, UK: Oxford University Press.
- McDonough, K., & Mackey, A. (2006). Responses to recasts: repetition, primed production, and linguistic development. *Language Learning*, 56 (4), 693-720. <https://doi.org/10.1111/j.1467-9922.2006.00393.x>
- Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair in dyadic interaction. *Language Learning*, 57 (4), 511-548. <https://doi.org/10.1111/j.1467-9922.2007.00427.x>
- Oxenden, C., & Latham-Koenig, C. (2006). *New English File*. Oxford, UK: Oxford University Press.

- Porte, G.K. (2010). *Appraising research in second language learning: A practical approach to critical analysis of quantitative research* (2nd ed.). Amsterdam, The Netherlands: John Benjamins. <https://doi.org/10.1075/llt.28>
- Redston, C., Cunningham, G., & Bell, J. (2013). *Face2face*. Cambridge, UK: Cambridge University Press.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar. A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). Amsterdam, The Netherlands: Benjamins.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.003>
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8 (3), 263-300. <https://doi.org/10.1191/1362168804lr146oa>
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10 (4), 361-392. <https://doi.org/10.1191/1362168806lr203oa>
- Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (ed.), *Conversational Interaction in Second Language Acquisition: A collection of empirical studies* (pp. 301–322). Oxford, UK: Oxford University Press.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics*. (pp. 125-144). Oxford, UK: Oxford University Press.
- Swain, M. (2005). The Output Hypothesis: theory and research. In Hinkel, E. (Ed.). *Handbook of research in second language teaching and learning*. (pp. 471-483). Mahwah, NJ; London, UK: Lawrence Erlbaum Associates.

- Takashima, H., & Ellis, R. (1999). Output enhancement and acquisition of the past tense. In R. Ellis (Ed.), *Learning a second language through interaction*. (pp. 173- 188). Amsterdam, The Netherlands/Philadelphia, PA: John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.17.12tak>
- VanPatten, B. (1996). *Input processing and grammar instruction. Theory and research*. Norwood: NJ: Ablex.
- Williams, J. (2005). Form-focused instruction. In Hinkel, E. (Ed.). *Handbook of research in second language teaching and learning*. (pp.671-691). Mahwah, NJ; London, UK: Lawrence Erlbaum Associates.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32 (2), 235–263. <https://doi.org/10.1017/S0272263109990519>

First version received: May 2016

Final version accepted: September 2016