The Effect of Oral Corrective Feedback on Article Errors in L3 English: Prompts vs. Recasts

Pawan M. Doski[1] & Filiz Çele[1]

[1] English Language and Literature Department, Istanbul Aydin University, Istanbul, Turkey

Correspondence: Pawan M. Doski, English Language and Literature Department, Istanbul Aydin University, Istanbul, Turkey. E-mail: pawandoski@outlook.com

Abstract

This study examines the effect of prompts and recasts in providing CF for the article errors by Kurdish-Arabic bilinguals who learn English as a third language. 39 lower-intermediate Kurdish-Arabic bilingual learners of English were tested on three tests: pre-, post-, and delayed post-tests. The participants were randomly put into three groups: (1) prompt group (n =15), (2) recast group (n = 14), and (3) no feedback group (n = 10). Each group completed 28 dialogues, which included articles in a Forced Choice Elicitation Task (FCET) as a pre-test. The same test was given to the three groups as post- and delayed post-tests. Between the pre-test and the post-test, the prompt and recast groups took a treatment which involved an interactional activity that aimed the FCET, in which the former took CF in the form of prompts, and the latter took it as recasts for their article errors in L3 English.

Results showed that all groups were the same in the pre-test. In addition, both the prompt and recast groups were similar in post-test but were significantly better than the group which did not receive any feedback. In delayed post-test, the prompt group significantly outperformed the other two groups. These findings suggest that prompts are more effective than recasts in providing oral feedback over the long term. The error analysis, on the other hand, revealed that among the four contexts of articles, all students had the highest error rate in the [-def, +spec] context in both pre- and post-tests. These were substitution errors rather than omission errors, which shows that the students fluctuated between definiteness and specificity settings. In delayed post-test, the prompt group significantly made fewer errors than the other two groups.

Keywords: oral corrective feedback, prompts, recasts, articles, TLA, Kurdish, Arabic, English

1. Introduction

The role of Corrective Feedback (CF) in Second Language (L2) learning has attracted considerable attention from L2 researchers who have been interested in the question of how learners' errors are useful in improving L2 learning. The current theories of the L2 acquisition assume different assumptions about the role of CF. the Generativist framework takes the view that positive evidence is sufficient for triggering parameter resetting in the L2 acquisition, therefore, CF has limited theoretical interest and usefulness to learners (Schwartz, 1993). However, the Interaction Hypothesis of the cognitivist school suggests that feedback in meaningful L2 interaction provides negative evidence, which, therefore, is useful in L2 learning (Long, 1996). Long claims that feedback-involving reformulation of a learner's own intended utterances can help in the noticing of gaps between the learner's own L2 production and L2 target forms and can prevent first language (L1) influence. According to some cognitive approaches, feedback contributes to the establishment and proceduralisation of declarative knowledge (Ranta & Lyster, 2007).

Empirical studies investigating CF in L2 learning have provided mixed conclusions. Lyster and Ranta (1997) found that recasts are the most common and preferred technique employed by foreign language teachers in spite of the fact that they are the least effective feedback technique. Moreover, they conclude that student-generated corrections are important in language learning since they show effective immersion in the process of a students' language learning, and this effective immersion arises when there is the negotiation of form. Russell and Spada (2006) in their meta-analysis review of 15 studies relating to the effectiveness of CF. compared the effects of implicit CF (i.e. recasts) to no CF. They found medium to large effects for *recast*s. They also calculated an effect size for one study (Carroll & Swain, 1993) that compared implicit with explicit CF types, and they found a large

effect size for explicit CF. Ammar and Spada (2006) compared the effectiveness of recasts and prompts on different proficiency level francophone learners of English which focused on third person possessive determiners. They found that prompts and recasts have an equal effect on high proficiency learners while prompts are significantly more effective than recasts for low-proficiency learners. Ammar (2008) in a quasi-experimental study compared the impact of recasts with prompts and no CF on francophone learners' acquisition of English third person possessive determiners. She found that prompts were more effective than recasts with no CF in facilitating the learners' development in possessive determiner and she further underlined the higher effect of prompts by the reaction time data which proved that prompts made learners to save possessive determiner knowledge faster than recasts. Lyster and Saito (2010) in a meta-analysis review of 15 classroom-based studies examined the effectiveness of several types of CF in the classroom. They found larger effects for prompts than recasts, while similar effects for explicit correction in compare with prompts and recasts.

The present study aims to contribute to the field of L3 research on CF with new empirical data on the comparison of prompts and recasts in providing feedback for article errors made by Kurdish-Arabic bilingual adolescent learners of L3 English. Kurdish is a (+article) language, the choice of article is based on definiteness. Kurdish has definiteness and indefiniteness markers added to the end of NPs, while in English they come before the NPs. Arabic, on the other hand, is a (+article) language which overtly has a definite marker and does not have an indefinite article, and the choice of articles is also based on definiteness. English overtly marks definiteness rather than specificity; it encodes this semantic knowledge in its article system. The difference between Arabic and English article system is the lack of overt indefinite marker in Arabic compared to English. In this study, we focus on whether prompts or recasts are more effective in correcting article errors produced by Kurdish-Arabic learners of L3 English. We expect the findings of the present study to have valuable implications for instructed L3 learning with respect to an effective type of CF.

2. Literature Review

2.1 Prompts and Recasts

Although there is a universal agreement that giving feedback produces important advantages in the performance of learners (Russell & Spada, 2006; Mackey & Goo, 2007; Lyster & Saito., 2010), there is controversy about the efficiency of the various types of oral CF. Many studies have compared diverse types of CF to examine which feedback is the most effective technique in the classroom. Whereas some studies show that there was no significant difference among different feedback types (e.g., Ammar & Spada, 2006; Loewen & Nabei, 2007; McDonough, 2007), others uncovered different effects of diverse types of feedback on L2 development (e.g., Long, Inagaki, & Ortega, 1998; Leeman, 2003; Ellis, Loewen, & Erlam, 2006). Many studies have examined the advantages of prompts in comparison with that of recasts (e.g., Yang & Lyster, 2010; Havranek, 1999; Ellis, Loewen, & Erlam, 2006; Lyster, 2004; Ammar & Spada, 2006). While the research studies in the classroom usually show prompt as more beneficial, laboratory studies suggest that recasts have a facilitating effect on the development of the second language. Lyster's (2004) study of young learners in the French immersion context, for example, found that when (FFI) is combined with prompts, it was more effective in facilitating the acquisition of grammatical gender than when it was combined with recasts. Similarly, Ammar and Spada (2006) found prompts to be more effective than recasts in the acquisition of English third Person Possessive Determiners (PDs) 'his' and 'her' by young French first language (L1) speakers, but that the effect of feedback depended on the learner's proficiency level. That is, whilst high proficiency learners (with pre-test scores above 50%) benefited equally from the two CF techniques, their low proficiency counterparts (with pre-test scores below 50%) benefited much more from prompts than from recasts. A meta-analysis of oral CF with 33 studies in both classroom and laboratory settings showed that explicit feedback (metalinguistic feedback, explicit correction) was more beneficial than implicit feedback (recasts, clarification requests, elicitation, and repetition) in the short term, whilst implicit feedback had more effect on L2 learning in the long term (Li, 2010). Another meta-analysis by Lyster and Saito (2010) investigated oral CF in the classroom; with a slightly different result for the effectiveness of different feedback techniques examined (recasts, explicit correction, and prompts). They revealed that prompts were more effective than recasts. Ellis et al. (2006) with adult ESL learners, it was shown that metalinguistic feedback (i.e. a prompt that consists of repeating the error with a hint of the problem) was more productive than recasts in acquiring the English regular past tense since the learners could recognize the corrective purpose of this CF technique more easily than that of recasts. Yang and Lyster (2010) in the EFL setting analyzed the impacts of recasts, prompts, and no feedback on the learning of the English regular and irregular past tense by grown-up college level learners in China. The findings revealed that on both immediate and delayed post-tests prompts benefitted participants more than recasts on the acquisition of the regular past tense. The studies that compared recasts to other techniques of CF have either produced positive effects for

recasts only (Long, Inagaki, & Ortega, 1998; Mackey & Philip, 1998) or for both prompts and recasts (Lyster & Izquierdo, 2009; McDonough, 2007). For instance, McDonough (2007) in the EFL setting compared the efficacy of recasts with clarification requests on the acquisition of simple past verbs among Thai L1 university undergraduate students and saw that both recasts and clarification requests were useful in producing improvement. Investigation of the acquisition of the French grammatical gender in second language adult learners also found no differences in the efficacy of prompts and recasts (Lyster & Izquierdo, 2009).

Researchers (Goo & Mackey, 2013; Lyster, Saito, & Sato, 2013) express uncertainty about measuring the comparative effectiveness of different feedback approaches, proposing as an alternative a demand for changing the method of CF research, indicating that it may not be useful to try and find just one helpful feedback technique while all feedback types have some helpful effect in L2 development.

With this background, this paper aims to examine the comparison of prompts and recasts in providing feedback for article errors made by Kurdish-Arabic bilingual adolescent learners of L3 English classrooms.

3. The Study

3.1 Research Questions and Predictions

This study attempts to examine the following research questions:

1) Do L3 learners who receive oral CF in the form of recast and prompts perform better than those who do not receive any CF?

In the light of previous research, we predict that the students who get oral CF either in the form of recast or prompt would be more successful than those who do not get any CF in directing L3 articles.

2) Are prompts more effective than recast in providing oral CF for English article errors by Kurdish-Arabic Bilingual learners of L3 English?

This study's prediction was that L3 learners who get prompt in the interactional activities as oral CF would benefit more than students who get recasts as oral CF for the article errors since previous studies suggested that prompts are more effective than recasts in Form-Focused Instruction (FFI) and in classroom learning settings.

3.2 Participants

The study was applied in focused English as a Second Language (ESL) system in Erbil. Focused ESL systems are provided in private schools from kindergarten stage. Students learn English every day for 10 months of the school year. They are taught the standard curriculum subjects (i.e., math, science, etc.) in English.

Three groups of students and a teacher (who was responsible for the application of the treatment) participated in this study. From a population of 87 students in a primary school in Erbil, 39 lower intermediate English level students were selected to participate. All participants were in grade 6 and in the second half of the school year. Using a linguistic background information questionnaire, we found that (the students) had unlimited exposure to English at home, the table below shows their background information.

Table 1. Participants' background information

| Groups | Gender | | Mean age of first exposure to L3 English | Mean age at the time of testing | Age range |
	Male	Female			M (SD)
Prompt (n=15)	3	12	4.3	11	11-13 (0.6)
Recast (n=14)	4	10	4.3	11	10-13 (0.8)
No feedback (n=10)	4	6	4.1	12	11-13 (0.8)

To ensure the homogeneity of the L3 English participants in the experiment, a proficiency test (Quick Oxford Placement Test (OPT)) was given to all the population in three classes each with 28-30 students. According to the results of the proficiency test given to the 87 students, 39 were found to be at the lower intermediate English level. Lower intermediate level was chosen for this study because the students in this level were seen to already know the target structures definite and indefinite articles ('a' and 'the') but not capable of using them properly. Then they were put randomly into three groups, two experimental groups; one for prompts, the other for recasts and a control group with no feedback. The following table shows the OPT results of the participants.

Table 2. Participants' OPT results

Groups	Mean	SD
Prompt (n=15)	34.87	3.18
Recast (n=14)	34.93	3.12
No feedback (n=10)	32.50	2.27

The age range of the participants was from 10 to 13. The students' primary language was Kurdish, but they had some exposure to Arabic outside the home and at school. Most of their prior language of education was English and some L1 Kurdish. The table below shows the participants Kurdish language background.

Table 3. Participants by Kurdish proficiency

	Beginner	Intermediate	Advanced	Near-Native
Reading	0	0	3	36
Writing	0	0	5	34
Speaking	0	0	0	39
Listening	0	0	2	37
Overall Competence	0	0	0	39

The table indicates that participants were native speakers of Kurdish. The general definition of bilingualism is the ability to speak two languages fluently. Bloomfield (Language, 1933) defines bilingualism as the acquisition of two or more languages at the native-like level. According to (Andersson & Boyer, 1978), bilingualism is using two languages for educational teaching in non-language subjects. However, Cummins (1980) states that bilingualism is the skill to continue a simple dialogue, for instance in a shop or in the street which is called-basic interpersonal communicative skills which can be improved with background signals like gestures which regularly occurs in face-to-face conditions. The table below shows participants' level of Arabic.

Table 4. Participants by Arabic proficiency

	Beginner	Intermediate	Advanced	Near-Native
Reading	14	17	8	0
Writing	28	9	2	0
Speaking	26	9	4	0
Listening	23	11	5	0
Overall Competence	22	14	3	0

3.3 Materials and Instruments

3.3.1 Proficiency Test

A proficiency test was given prior to conducting any of the studies' tests on the participants. The test was run in three classes of Grade 6 in the primary school where the study has been conducted. The number of the students in each of the three classes was between 28-30, from a total population of 87 students. They took a quick placement test by (Oxford-University-Press & Cambridge-University-Syndicate, 2001). This test is a standard measure of proficiency in English. It has a standardized scale that verifies learners' levels as a beginner, elementary, lower intermediate, upper-intermediate, advanced, and very advanced including all levels from A1 to C2. It contained 60 questions given to the population and they had 30 minutes to answer all the questions. The results of the test showed different English levels among the population and classified the learners. However, lower-intermediate level students were selected to participate in this study. The table below shows the mean and range of their scores according to each group.

Table 5. Participants by English proficiency

English L3	N	Mean	Range (out of 60)
Lower intermediate (Prompt)	15	34.87	30-39
Lower intermediate (Recast)	14	34.93	31-39
Lower intermediate (No feedback)	10	32.50	31-36

In summary, to ensure the homogeneity of the participants for the experiment, we only focused on the results of lower intermediate L3 English students, excluding other English level learners.

3.3.2 Pretest

The experimental study consisted of a forced-choice elicitation task (Ionin, 2003); the pre-test was applied to measure the participants' current level of accuracy in producing articles in L3 English.

Forced choice elicitation task (FCET):

The only task applied in the study was an article FCET, a task which allowed L3 students to choose the correct article for a certain context. It contained 76 short dialogues from (Ionin, 2003). This task was selected because it was shown to work well and be effective to elicit and test articles in the previous studies. The selection of the test design was made to facilitate the researchers' work on the context categories.

The only task in this study was a *forced choice task* to test the use of articles in different semantic contexts. In each dialogue of the task, the target sentence was missing an article: the learner had to choose between 'a', 'the', and the null article (---), choosing his or her answer according to the preceding context. There were 28 dialogues, 7 in each of the [-definite, -specific], [-definite, +specific], [+definite, +specific], [+definite, -specific] contexts. The target article in each item was underlined. For example:

[+definite, -specific] *At a supermarket*

Sales Clerk: May I help you, sir?

Customer: Yes! I'm very angry. I bought some meat from this store, but it is completely spoiled! I want to talk to **(a, the, --) owner** of this store - I don't know who he is, but I want to see him right now! (Ionin, 2003, p. 277).

3.3.3 Post-test

Immediately after the treatment, the post-test was given and the same FCET, which was used in the pre-test, was used again. The post-test aimed to show the improvement of the students' article use.

3.3.4 Delayed Post-test

18 days later, the delayed post-test was given, to show the learners' processing knowledge and the same FCET which was used in the pre-test and the post-test was used again in this test, but the students did not know.

3.4 Treatment

3.4.1 Observation

Over a six-day period, 45-minute group lessons were observed, and 149 CF occurrences arising from the interactions between the teacher and the students were coded. Unfortunately, the participating teacher did not give any permission for any video recording or audio recording of her lectures during the treatment, but she accepted an observer in her classes on a daily basis. Therefore, with using CF observation scheme adapted from (Ammar & Spada, 2006) 13.5 hours of classroom teaching interaction were observed and coded in real time by the researcher sitting at the back of the classroom. The observation scheme contained the students' errors; the two types of the CF of interest to this study recasts and prompts. All other feedback methods were classified under the group "other" and the errors to which the teacher did not react classified under a group nominated as "ignore".

3.4.2 The Preparation Session

An experienced ESL teacher with master's degree in English Language from Cambridge educational program at a primary school in Erbil was selected to give the treatment involving teaching articles. She was selected after discussions with a range of teachers and her supervisor within the program. In this session, some measures were carried out to make sure that the various requirements of CF methods were applied as designed. First, all the instructional elements were comprehensibly described during the preparation session that preceded the

experimental treatment and the teacher was provided with a booklet which explained the use of CF techniques with examples adapted from (Ammar & Spada, 2006). Second, the researcher remained in a daily contact with the teacher to assure that she was not facing difficulties performing the task. These measures allowed the researcher to find out the frequency with which feedback was given, and they helped to see that the teacher was applying the kind of feedback as required.

3.4.3 The Classroom Interaction Activity

The treatment was divided over a seven-day period with 45 minutes sessions. Over that time there were two major stages: the first stage consisted of one teaching session held on the first day of the treatment. The teacher gave explanations about the articles and the rules for using definite and indefinite articles, this session was arranged to review the student prior knowledge of the the target structure, since prompts cannot be applied to elicit information that learners do not know already (Lyster, 2004), and to free the results from the effects of the experimental differences.

The second stage focused on the FCET, which continued for the next six days in which dialogue build activity was prepared that included sixteen short dialogues. Four dialogues for each of the four (definiteness and specificity) contexts. The dialogues were two-6-line conversations between two people in different semantic contexts. The teacher drew picture of two people on the board, setting the context, for example: *at a supermarket*, then she put the two speakers' pictures on either side of the board. She drew a bubble from the person who speaks first and inserted a prompt, for example *where/going?* Elicited the target sentence, *where are you going?* Shaped and drilled the target language with the whole class and then individually, for each conversation she chose two students to come to the board and act in front of the whole class, after that she drew a reply speech bubble and inserted a prompt, continuing in this way, establishing one line at a time until the conversation was completed (Clark, Young, Redston, & Cunningham, 2005). The sentences of each dialogue were intentionally chosen to elicit [DS, DNONS, IDS, IDNONS] contexts. The students practiced the conversations in pairs, then changed their role and practiced the conversations again during which the teacher was giving oral CF according to each group.

Some examples of the conversation between the students and the teacher, in the recast group:

- S: I am planning to visit oldest armory in Kurdistan. It's in Ruanduz, Soran.

T: Okay, I plan to visit the oldest armory, what else. ➔ *(Recast)*

In the prompt group for example:

- S1: I would like to read most interesting book in the library.

T: MOST interesting book? (Rising intonation) ➔ *(Repetition)*

S1: The most interesting book.

S1: He got the scarf; it is green with big purple stripes.

T: No, it is an unknown scarf until you introduce it to us. ➔ **(Metalinguistic clues)**

- S2: She got **a hat**; she wanted to cover her head from **sun**.

T: Excuse me! I did not understand.➔ **(Clarification requests)**

S2: She got **the hat;** she wanted to cover her head from **sun**.

T: This sentence is incorrect. Who can correct it? ➔ **(Elicitation)**

S3: She got a hat; she wanted to cover her head from **the** sun.

T: Bravo!

3.5 Procedure

As mentioned before, 39 students from a grade 6 primary school participated in this study. First, the OPT was applied to ensure their homogeneity before starting the experiment. After getting the results of the proficiency test the lower-intermediate level students were selected, and then they were put into three groups: two experimental groups (Recasts and Prompts) and a control group (No Feedback). On the following day, a linguistic background information questionnaire was given to check the language background and analyze their errors accordingly. Then, a teacher was selected to give the treatment. One day before the treatment the researcher held a preparation session with the participating teacher, by providing comprehensive information on the CF methods. To show the students' ability in using the target structure (articles), a pre-test was assigned to the three groups, with an FCET, and applied repeatedly in the following tests (post-test, and delayed post-test).

The participants were given appropriate instructions in answering the questions. They were given 45 minutes to complete the task. The next day, the treatment was administered by incorporating the dialogue build activity that intentionally elicited the target structure. Learners were then involved in communicative dialogue build exercises, and their incorrect sentences in *recast* group were corrected. In the prompts group, four methods of prompts namely 'repetition', 'clarification request', 'metalinguistic clues', and 'elicitation' were used independently, or two or more types of prompts were mixed to correct learners' ungrammatical sentences. The treatment continued for a week, and immediately after, the post-test was given. After 18 days, the delayed post-test was given.

4. The Results

4.1 Results from the FCET

The FCET was graded out of 28 points. The descriptive statistics for accuracy scores of the task, containing means and standard deviations for each of the three groups during the pretest, posttest, and delayed posttest, are given in Table 6 below.

Table 6. Overall accuracy means and standard deviations from FCET

	Prompt (n=15)	Recast (n=14)	No Feedback (n=10)
	M(SD)	*M(SD)*	*M(SD)*
Pre-test	18.33(3.94)	18.64(3.03)	20.30(3.30)
Post-test	21.33(3.54)	21.93(3.22)	21.10(3.07)
Delayed post-test	23(3.14)	18.43(4.30)	18.40(4.55)

CF techniques in the L3 classroom during interaction might affect L3 English article use accuracy, but that effect might differ across CF groups. To know whether different semantic contexts effect different accuracy in article use, a one-way analysis of variances tested the impact of three CF techniques on the article use of Kurdish-Arabic bilingual learners of L3 English, in three tests. The results showed no significant differences among the three groups during the pretest: $F (2, 36) = 1.041$, $p = 0.364$; in the post-test, $F (2, 36) = 0.209$, $p = 0.813$. However, in the delayed post-test there was a significant difference between the groups, $F (2, 36) = 6.191$, $p = 0.005$. The prompt group significantly outperformed the recast and no feedback groups, as illustrated in Figure 1.
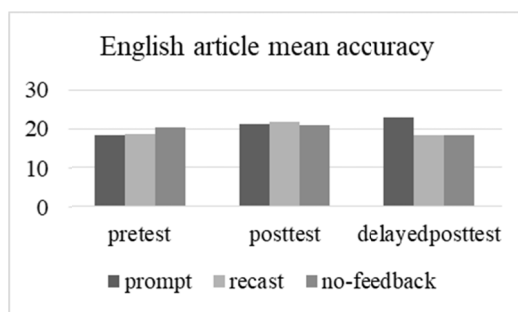

Figure 1. Mean accuracy on FCET over the three test periods

The figure shows that in the pre-test and post-test there is not a significant difference among the three groups but in the delayed post-test the difference among them is very significant.

4.2 Results from the Pre-test

Accuracy analysis:

Accuracy results from three groups through FCET in the pretest are summarized in Table 7.

Table 7. Mean accuracy from four contexts at pre-test

Groups	DS (n=7)	DNONS (n=7)	IDS (n=7)	IDNONS (n=7)
	M(SD)	*M(SD)*	*M(SD)*	*M(SD)*
Prompt (n=15)	6.07(1.10)	4.93(1.58)	2.67(2.16)	4.67(1.45)
Recast (n=14)	5.43(1.34)	5.43(1.45)	3.21(1.48)	4.57(1.55)
No Feedback (n=10)	**6.3(0.67)**	**5.2(1.03)**	**3.5(1.43)**	**5.3(1.34)**

DS (definite – specific contexts), DNONS (definite – non-specific contexts), IDS (indefinite- specific), IDNONS (indefinite – non-specific contexts).

The table above shows that among four contexts all the groups are less accurate in the IDS context followed by DNONS for prompts and IDNONS for recasts. However, all groups were more accurate in the DS context than in the other contexts. There were some differences in article use among the groups in the pre-test. To show the differences, a two-way ANOVA was used with the groups as the between-subject factor and accuracy values in four contexts as a within-subject factor. The results indicated a main effect of contexts $F (3,108) =29.32$, $p = 0, 00$, such that all the groups had higher accuracy in DS context and had the lowest accuracy in IDS context. The interaction effect was not significant $F (6, 108) = 0. 79$, $p = 0.57$. In all cases, the groups showed low accuracy in the IDS context, as shown in Figure 2 below.
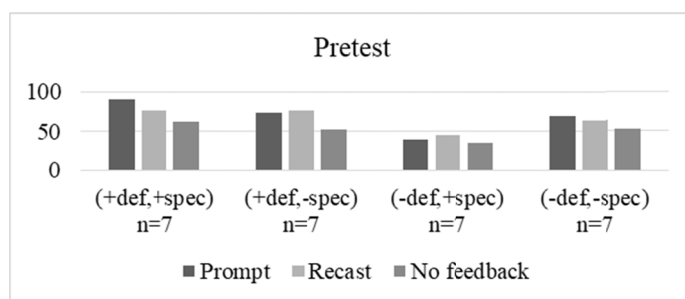

Figure 2. Accuracy from four contexts at pretest

This figure shows that learners' accuracy in the four contexts in which they fluctuate between being more accurate in some contexts and less accurate in others.

Error analysis:

The error results from three groups through FCET in the pretest are summarized in Table 8.

Table 8. Frequency and percentage of article errors in four contexts at pretest

Groups	DS (n=7)		DNONS (n=7)		IDS (n=7)		IDNONS (n=7)	
	S	*O*	*S*	*O*	*S*	*O*	*S*	*O*
Prompt (n=15)	12/105 11%	2/105 2%	18/105 17%	13/105 12%	62/105 59%	3/105 3%	30/105 29%	5/105 5%
Recast (n=14)	20/98 20%	2/98 2%	14/98 14%	8/98 8%	50/98 51%	3/98 3%	31/98 32%	3/98 3%
No Feedback (n=10)	6/70 9%	1/70 1%	9/70 13%	9/70 13%	31/70 44%	4/70 6%	16/70 23%	1/70 1%

S (Substitution), O (Omission), DS (definite – specific contexts), DNONS (definite – non-specific contexts), IDS (indefinite- specific), IDNONS (indefinite – non-specific contexts).

The table shows that among four contexts the highest level of errors was in substituting *a* with *the* in the IDS followed by IDNONS context, substituting *the* with *a* in DNONS context for prompts and in DS context for recasts. However, the highest level of omission was in omitting *the* in the DNONS context, and the lowest level of omission was in the DS context. To present error type differences, a repeated measure ANOVA was used with the groups as a between-subjects factor and article context (DS, DNONS, IDS, IDNONS) and error type (omission and substitution) as the within-subject factors. There was a significant main effect of errors $F (1, 36) = 160.16$, p= 0.00. There was a significant interaction effect between contexts*errors $F (3,108) = 33.50$, p = 0.00. However, this effect was not significant across the groups $F (2, 36) = 1.04$, p = 0.36, which means the groups made similar errors in similar contexts at pre-test, as displayed in Figure 3.
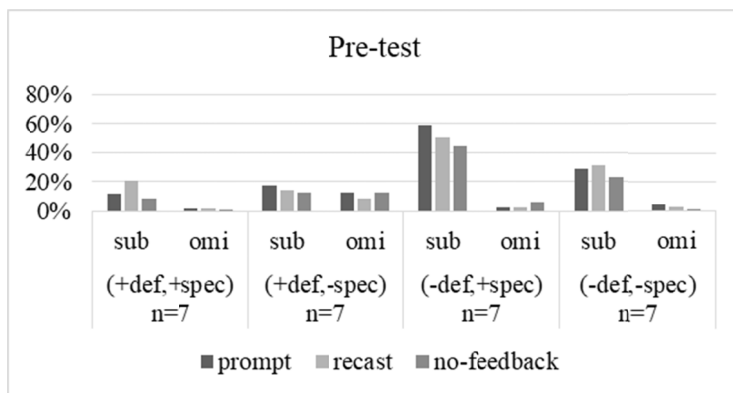

Figure 3. Types of errors in the four contexts at pretest (Sub = substitution, Omi = omission)

This figure shows the students' error types they made in four contexts, and how it varies between omitting and substituting articles across the groups at pre-test.

4.3 Results from Four Contexts in Posttest

Accuracy analysis:

Accuracy results from three groups through FCET in the posttest are summarized in Table 9.

Table 9. Mean accuracy from four contexts at posttest

Groups	DS (n=7) M(SD)	DNONS (n=7) M(SD)	IDS (n=7) M(SD)	IDNONS (n=7) M(SD)
Prompt (n=15)	6.2(0.77)	5.2(1.86)	4.6(1.76)	5.33(1.35)
Recast (n=14)	5.64(1.55)	6.29(0.91)	4.43(1.91)	5.57(1.28)
No Feedback (n=10)	**6.3(0.82)**	**5.9(1.20)**	**4.3(1.77)**	**4.6(1.51)**

DS (definite – specific contexts), DNONS (definite – non-specific contexts), IDS (indefinite- specific), IDNONS (indefinite – non-specific contexts).

This table shows that among the four contexts all the groups were less accurate in the IDS context. However, all groups were more accurate in the DS context followed by DNONS for recast and DNONS for prompt. To present the differences, a two-way ANOVA was used with the groups as the between-subject factor and accuracy values in four contexts as a within-subject factor. There was a significant main effect of contexts $F (3,108) = 10.21$, p = 0.00, that the groups maintained to perform inaccurately in IDS context. There was not a significant interaction effect $F (6, 108) = 1.51$, p = 0, 18, which shows that the groups also performed equally in the post-test, as illustrated in Figure 4.
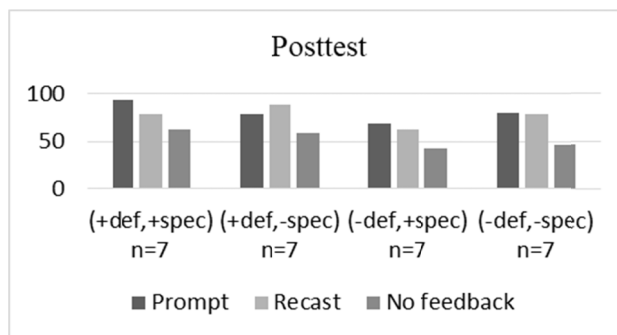
Figure 4. Four contexts accuracy at posttest

This figure shows learners' accuracy in the four contexts at post-test, which shows that the participants were less accurate in the (-def, +spec) context.

Error analysis:

Error results from three groups through FCET in the post-test are summarized in Table 10.

Table 10. Frequency and percentage of article errors in four contexts at posttest

Groups	DS (n=7)		DNONS (n=7)		IDS (n=7)		IDNONS (n=7)	
	S	O	S	O	S	O	S	O
Prompt (n=15)	10/105 10%	2/105 2%	13/105 12%	14/105 13%	30/105 29%	6/105 6%	17/105 16%	8/105 8%
Recast (n=14)	17/98 17%	2/98 2%	3/98 3%	7/98 7%	26/98 27%	10/98 10%	13/98 13%	7/98 7%
No Feedback (n=10)	5/70 7%	2/70 3%	5/70 7%	6/70 9%	19/70 27%	8/70 11%	22/70 31%	2/70 3%

S (Substitution), O (Omission), DS (definite – specific contexts), DNONS (definite – non-specific contexts), IDS (indefinite- specific), IDNONS (indefinite – non-specific contexts).

The table shows that the highest frequency of errors was in substituting *a* with *the* in the IDS, followed by substituting *the* with *a* in DS context for recast group and *a* with *the* in IDNONS for prompts. However, the highest level of omission in all the groups was in omitting *the* in the DNONS context, and the lowest level of omission errors was in the DS context. To display the error type differences, a repeated measure ANOVA was used with the groups as a between-subjects factor and four contexts (DS, DNONS, IDS, IDNONS) and error type (omission and substitution) as the within-subject factors. The results indicated a significant main effect of errors $F_{(1, 36)} = 46, 63$, $p = 0.00$. There was also a significant interaction effect $F_{(3,108)} = 8.21$, $p = 0.00$, which means the errors were different with respect to context. However, between-subject effects revealed that there was no significant difference between the groups in making errors $F_{(2, 36)} = 0.21$, $p = 0.81$. That is groups made the same errors in the same contexts, as shown in Figure 5.
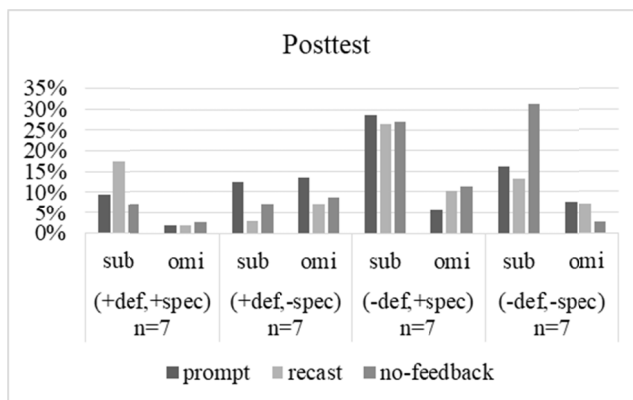
Figure 5. Types of errors in the four contexts in the posttest (Sub = substitution, Omi = omission)

This figure displays students' error types in four contexts at post-test, in which most of their errors were in substitution in the (-def, +spec) setting.

4.4 Results from Four Contexts in Delayed Post-test

Accuracy analysis:

Accuracy results from three groups through FCET in the delayed posttest are summarized in Table 11.

Table 11. Mean accuracy from four contexts at delayed post-test

Groups	DS (n=7)	DNONS (n=7)	IDS (n=7)	IDNONS (n=7)
	M(SD)	*M(SD)*	*M(SD)*	*M(SD)*
Prompt (n=15)	6.53(0.64)	5.8(1.32)	5.73(1.22)	4.93(1.98)
Recast (n=14)	5.43(1.83)	4.21(2.36)	4(1.80)	4.79(1.81)
No Feedback (n=10)	**5.9(1.20)**	**4.3(2.11)**	**3.3(2.11)**	**4.9(1.37)**

DS (definite – specific contexts), DNONS (definite – non-specific contexts), IDS (indefinite- specific), IDNONS (indefinite – non-specific contexts).

This table indicates that all groups were less accurate in the IDS context. However, all groups were more accurate in the DS context followed by IDNONS for *recast*s, DNONS, and IDS for prompt. To present the differences in article accuracy, a two-way ANOVA was used with the groups as the between-subject factor and accuracy values in four contexts as a within-subject factor. There was a significant main effect of contexts F $(3,108) = 6.96$, p = 0.00. There was a significant interaction effect between the three groups with respect to the four contexts F $(6,108) = 1.54$, p = 0.17. The prompt group outperformed the recast and no feedback group, as illustrated in Figure 6.
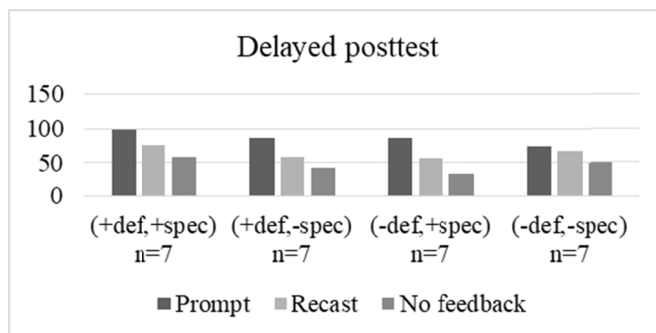
Figure 6. Four contexts accuracy at the delayed posttest

This figure displays students' article accuracy level in four contexts at delayed post-test, where they were more accurate in the (+def, +spec) setting.

Error analysis:

Error results from three groups through FCET in the delayed post-test are summarized in Table 12.

Table 12. Frequency and percentage of article errors in four contexts at delayed post-test

Groups	DS (n=7)		DNONS (n=7)		IDS (n=7)		IDNONS (n=7)	
	S	O	S	O	S	O	S	O
Prompt (n=15)	5/105 5%	2/105 2%	14/105 13%	4/105 4%	16/105 15%	3/105 3%	27/105 26%	4/105 4%
Recast (n=14)	8/98 8%	14/98 14%	16/98 16%	23/98 23%	38/98 39%	4/98 4%	27/98 28%	4/98 4%
No Feedback (n=10)	3/70 4%	8/70 11%	8/70 11%	19/70 27%	32/70 46%	5/70 7%	19/70 27%	2/70 3%

S (Substitution), O (Omission), DS (definite – specific contexts), DNONS (definite – non-specific contexts), IDS (indefinite- specific), IDNONS (indefinite – non-specific contexts).

The table shows that the highest level of errors was in the recast and no feedback groups, in substituting *a* with *the* in the IDS context and omitting *the* in the DNONS context, followed by substituting *a* with *the* in the IDNONS context for all groups. However, the lowest level of substitution in all the groups was in substituting *the* with *a* in the DS context. To display the error type differences, a repeated measure ANOVA was used with the groups as a between-subjects factor and four contexts (DS, DNONS, IDS, IDNONS) and error type (omission and substitution) as the within-subject factors. The results indicated a significant main effect of errors $F(1, 36) = 53.83$, $p = 0.00$. There was also a significant interaction effect $F(3,108) = 26.85$, $p = 0.00$, that is the number of errors differed with respect to the context types, between subject effects showed a significant effect regarding groups $F(2,36) = 6.19$, $p = 0.005$. Tukey post hoc multiple comparisons showed a significant difference between prompt and recast $p = 0.01$, prompt and no feedback group $p = 0.02$, which means that total errors in recast and no feedback groups were significantly higher than in the prompt group, as displayed in Figure .
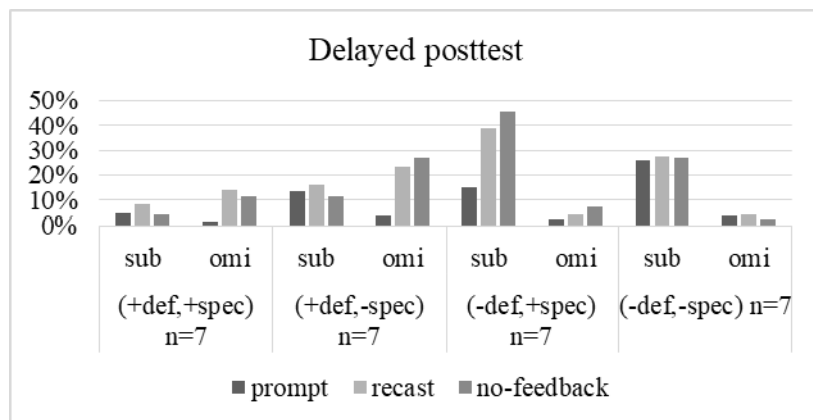
Figure 7. Types of errors in the four contexts at the delayed post-test (Sub = substitution, Omi = omission)

This figure shows error types in four contexts at delayed post-test, in which the level of substitution errors in the prompt group was significantly lower than in the recast and no feedback groups.

To sum up, results obtained from (pre-, post-, and delayed post-) tests indicate that both prompts, and recasts are helpful in L3 English articles use, and prompts are more successful than recasts and no feedback on Grade 6 learners' article use in third language learning classrooms.

5. Discussion

Providing corrective feedback in second and foreign language classes has been frequently examined, however, it is still unclear what types of feedback are more effective in correcting syntactic errors. Some claim that student-generated corrections (prompts) are important and more effective than recasts in language learning classrooms (Lyster & Ranta, 1997). Others claim that recasts are the best CF method because they are positive, unnoticeable, and provide a correct form, at the same time keeping the learners' intended meaning (Doughty & Varela., 1998; Long, 1996).

Following the review of previous CF studies in both second and third language acquisition, this study aimed to examine the effect of two techniques of oral CF; prompts and recast on the article errors in L3 English learning. Important findings have been discovered. Statistics from the FCET showed that both prompt group and recast group developed in the post-test, which means that both prompts, and recasts were useful in acquiring L3 English articles and the prompt group significantly outperformed the recast and no feedback groups at the delayed post-test.

The findings of this study are in line with findings of researchers who argue that recasts are implicit and consequently, maybe not successful enough, in meaning-based teaching classrooms (Lyster & Ranta, 1997; Lyster, 1998; Chaudron, 1977; Fanselow, 1977) and suggest that *prompt* is a more successful method. It shows that the higher effects of prompts, which direct learners' attention to problematic parts of their utterance and might make them self–repair their ungrammatical utterances, give learners opportunities to proceduralize the declarative knowledge of articles. The results of this study do not support the findings of previous studies which defend *recast* as a successful CF technique since it is indirect, comprehensive, unnoticeable, and depending on the learners' intended meaning, which keeps the flow of conversation. In other words, it does not interrupt learners' conversation (Oliver, 1995; Long, 1996; Doughty & Varela., 1998; Doughty, 2001; Leeman, 2003). The smaller effect of recasts might be due to the ambiguity it produces for the students in a way that might make them confused between whether the teacher's feedback was an indication of confirmation or a disapproval feedback on their output (Lyster, 2004). Analyses of the errors showed that Kurdish-Arabic L3 learners of English were least accurate on IDS context, because the source of the errors was substitution rather than omission. Given the Fluctuation Hypothesis, L1 speakers of non-article languages will either fluctuate between definiteness and specificity when learning a language that encodes the features [+definite] or [+specific] or will select the appropriate value for the target language. However, L1 Kurdish and Arabic both are [+article] languages based on definiteness and our participants fluctuated. Error analysis findings of this study suggest that a specificity-based choice of articles can occur even when the L1 is an article-based language which is based on definiteness. The students' errors when fluctuating in their choice of appropriate article might also be due to their

language background differences when learning English. In other words, the effect of their L1 might have a role in their errors due to the differences in the position of the articles in ~~the~~ Kurdish and English, in addition to the lack of the indefinite article in Arabic compared to English. Results from delayed post-test show that prompts are more effective than recasts in correcting substitution errors over the long term.

6. Implications and Limitations

6.1 Implications

Findings obtained from the present study suggest firstly in interactional L3 classrooms that teachers should take into consideration the application of oral CF (in the form of prompts or recasts), to allow learners to work on any kind of knowledge they have developed from teaching. In their development, learners may notice the difference between their language system and the target language system. Secondly, regarding the CF-type that teachers should apply during interaction in the classroom, the study suggests providing oral prompts in the correction of English article errors for Kurdish-Arabic bilingual learners of L3 English rather than recasts. Another crucial point is providing teaching sessions in which instructors describe and compare the students' L1, L2, and L3 article systems and grammar rules.

6.2 Limitations and Recommendations for Further Research

The language learning setting and the instructional tasks are two changeable factors that need to be considered for further research since earlier studies have shown that the effects of a CF method can differ from one setting to the other. Some studies have shown while prompts are very effective in the classroom settings, recasts are more effective in the laboratory setting since this study's aim was to discover the teaching implications of CF by choosing only the classroom setting for conducting this research. Involvement in a laboratory setting would be important for further investigation. Another limitation of this study was implementing one comprehension task (FCET), whereas inserting other tasks such as production tasks might also have different results that have instructional values. Moreover, they can provide absolute evidence as to which CF methods are more successful in language learning development.

References

Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research, 12*(2), 183-210. https://doi.org/10.1177/1362168807086287

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, Prompts, and L2 Learning. *Studies in Second Language Acquisition, 28*, 543-574. https://doi.org/10.1017/S0272263106060268

Andersson, T., & Boyer, M. (1978). *Bilingual schooling in the United States.* Austin, TX: National Educational Laboratory Publishers.

Bloomfield, L. (1933). *Language.* Chicago: University of Chicago Press.

Carroll, S., & Swain, M. (1993). Explicit and Implicit Negative Feedback: An Empirical Study of the Learning of Linguistic Generalizations. *Studies in Second Language Acquisition, 15*(3), 357-386. https://doi.org/10.1017/S0272263100012158

Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning, 27*(1), 29-46. https://doi.org/10.1111/j.1467-1770.1977.tb00290.x

Clark, R., Young, A., Redston, C., & Cunningham, G. (2005). *Face2face pre-Intermediate Teacher's Book with over 100 pages of extra teacher's resources.* Cambridge: Cambridge University Press.

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly, 14*(2), 175-187. https://doi.org/10.2307/3586312

Doughty, C. (2001). *Cognitive underpinnings of focus on form* (Cognition and second language instruction ed.). (P. Robinson, Ed.) New York: Cambridge University Press.

Doughty, C., & Varela., E. (1998). Communicative focus on form. *Focus on form in classroom second language acquisition, 1*, 114-138.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *SSLA, 28*, 339-368. https://doi.org/10.1017/S0272263106060141

Fanselow, J. F. (1977). The treatment of error in oral work. *Foreign Language Annals, 10*(5), 583-593. https://doi.org/10.1111/j.1944-9720.1977.tb03035.x

Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition, 35*(1), 127-165. https://doi.org/10.1017/S0272263112000708

Havranek, G. (1999). The effectiveness of corrective feedback: Preliminary results of an empirical study. *Acquisition et Interaction en langue étrangère* (pp. 189-206). Proceedings of the Eighth EUROSLA Conference.

Ionin, T. R. (2003). Article semantics in second language acquisition. *(Doctoral dissertation, Massachusetts Institute of Technology)*.

Leeman, J. (2003). Recasts and second language development: Beyond Negative Evidence. *Studies in Second Language Acquisition, 25*(1), 37-63. https://doi.org/10.1017/S0272263103000020

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

Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361-377). Oxford: Oxford University Press.

Long, M. H. (1996). *The role of the linguistic environment in second language acquisition* (Handbook of Second Language Acquisition ed.). (W. C. Bhatia, Ed.) New York: Academic Press.

Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal, 82*(3), 357-371. https://doi.org/10.1111/j.1540-4781.1998.tb01213.x

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. (2004). Research on form-focused instruction in immersion classrooms: Implications for theory and practice. *French Language Studies, 14*, 321-341. https://doi.org/10.1017/S0959269504001826

Lyster, R., & Izquierdo, J. s. (2009, June). 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, 20*, 37-66. https://doi.org/10.1017/S0272263197001034

Lyster, R., & Saito., K. (2010). Oral feedback in classroom SLA. *Studies in Second Language Acquisition, 32*(2), 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), 1-40. https://doi.org/10.1017/S0261444812000365

Mackey, A., & Goo, J. (2007). Interaction in SLA: A research synthesis and meta-analysis. In A. Machey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 407-452). Oxford: Oxford University Press.

Mackey, A., & Philip, 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. (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: Oxford University Press.

Oliver, R. (1995). Negative feedback in child NS-NNS conversations. *Studies in Second Language Acquisition, 17*(4), 459-481. https://doi.org/10.1017/S0272263100014418

Oxford-University-Press, & Cambridge-University-Syndicate, o. L. (2001). *Quick Placement Test* (version 2 ed.). U. C. L. E.

Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The awareness-practice-feedback sequence. 141. https://doi.org/10.1017/CBO9780511667275.009

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar. In J. M. Norris, & L. Ortega (Eds.), *Synthesizing Research on Language Learning and Teaching* (pp. 133-164). Amsterdam, Netherlands: John Benjamins Publishing.

Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in second language acquisition, 15*(2), 147-163. https://doi.org/10.1017/S0272263100011931

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*, 235-263. https://doi.org/10.1017/S0272263109990519

Copyrights