



Research Article

Unraveling the Impact of Recasts, Working Memory, and Textual Enhancements on EFL Learning: Intensive vs. Extensive Approaches

Farzaneh Bahadori ¹, Mohammad Bavali ², Mohammad Javad Riasati ³, Samad Mirza Suzani ⁴

1. Ph.D. Candidate, Department of Foreign Languages, Shiraz Branch, Islamic Azad University, Shiraz, Iran

2.3. Department of Foreign Languages, Shiraz Branch, Islamic Azad University, Shiraz, Iran

4. Department of Foreign Languages, Marvdasht Branch, Islamic Azad University, Fars, Iran

* Corresponding author: Mohammad Bavali; Email: mbvl57@gmail.com

ARTICLE INFO

Submission History

Received: 2023-06-01

Accepted: 2023-02-11

Keywords

Intense and extensive educational approaches
Verbal revisions versus textual Improvements; memory function
Language proficiency development

ABSTRACT

One of the challenges in English Language Learning (ELL) research is understanding the impact of different instructional techniques on grammatical and lexical growth. This study addresses this challenge by investigating the effects of extensive versus intensive recasts and textual enhancements (TEs) on Foreign Language (FL) development, while considering the role of working memory capacity and students' perceptions. A quasi-experimental design was implemented, involving four experimental groups and one control condition. The results of one-way ANOVA and mixed between-within group ANOVAs revealed the facilitative effect of recasts on English as a Foreign Language (EFL) learning. Specifically, the intensive recasts group demonstrated significantly higher performance compared to the extensive recasts group, intensive TE group, extensive TE group, as well as the control group, in both tests. The study also examined students' perceptions through interviews, which indicated a notable difference in accurate perception of recasts. Students in the intensive recasts condition exhibited a higher level of accuracy compared to those in the extensive recasts group. Conversely, the extensive recasts group displayed limited error detection. Additionally, a significant correlation between working memory and learners' gain scores was observed in the intensive recasts condition. This suggests that learners with higher storage capacity demonstrated more substantial development in oral tasks, while learners with higher attention domain exhibited greater improvements in written tasks. By shedding light on these findings, this study contributes to our understanding of effective instructional techniques and their impact on EFL development.

Introduction

Recasts are a common form of corrective feedback (CF) used in language classes and researched by Second Language (L2) scholars. Despite their popularity, previous studies have shown that L2 learners may not recognize recasts as CF due to their implicit nature, which weakens their effectiveness (Mackey et al., 2000; Mackey, 2007; Rassaei, 2013). However, combining recasts with other Focus on Form (FOF) methods, such as TE, may increase their saliency and impact on L2 learning (Nguyen et al., 2017). TE is a method used to increase the saliency of target forms in input and is based on the notion that increased saliency of the target forms triggers noticing and facilitates L2 development (Meguro, 2019).

TE and recasts can also be provided extensively on a wide range of errors or intensively on certain preselected errors (Ellis et al., 2001; Hawkes & Nassaji, 2016; Lyster & Ranta, 2013; Nassaji, 2007, 2009). Therefore, it is interesting to know if different enhancements in the saliency of linguistic structures through TE and recasts would raise the saliency and effectiveness of learners' development differently. Nonetheless, little research has compared the effectiveness of extensive versus intensive recasts and TEs concerning the role of WM capacity. The present study is set up to address these issues intertwined. The research questions are:

1. Does intensive or extensive recast and/or TE have an effect on EFL students' improvement in vocabulary and grammar?
2. If yes, which experiment is more effective for EFL students improvement?
3. How do students perceive the effect of intensive or extensive recasts on EFL development?
4. Can WM capacity account for any effects of recasts or TE experiments on EFL development?

Recasts in SLA: Current Perspectives and Emerging Trends

Contemporary second language acquisition (SLA) research focuses on exploring the effectiveness of recasts as a feedback strategy (Lyster & Ranta, 2013; Mackey, 2012). SLA theories view adult language learning as a complex

cognitive process with distinct stages. Initial stages involve acquiring factual information and L2 rules through explicit instruction or observation of native speakers (Spada, 2015; VanPatten, 2004). Proceduralization then transforms declarative knowledge into readily accessible language chunks, enabling efficient language production (Anderson, 2000; Rebuschat, 2013). Challenges arise in transferring procedural knowledge to different contexts and skills, limiting its generalizability (Ortega, 2005).

Extensive practice in the final stage aims for automaticity in language use, enabling spontaneous and fluent production (Segalowitz, 2010; Segalowitz & Hulstijn, 2005). Recasts, as feedback, aid in consolidating declarative knowledge, facilitating procedural knowledge development, and promoting language skill automatization (Lyster, 2004; Mackey & Goo, 2007).

Recent research highlights the potential of recasts in encoding novel knowledge, but findings vary (Hosseini, 2019; Mackey, 2016; Nassaji & Fotos, 2011). Further investigation is needed to determine effective conditions and underlying mechanisms impacting L2 development (Sheen, 2018; Wong & VanPatten, 2003).

Textual Enhancement (TE)

Textual Enhancement (TE), an implicit method of focus on form (FOF), has gained attention in language learning research (Lee, 2007; Meguro, 2019; Reinders & Ellis, 2009; Simard, 2009). TE involves visually highlighting certain aspects of a text, such as using italics, underlining, boldfacing, or capitalizing linguistic features (Rassaei, 2015).

TE aims to direct learners' attention to salient linguistic features in the text, thereby enhancing L2 development (Lee, 2007). This approach is grounded in Schmidt's noticing hypothesis, which posits that learners must notice linguistic features for acquisition (Schmidt, 2001). By modifying the input visually, TE increases the likelihood of noticing and processing language features, facilitating long-term memory (Lee & Huang, 2008).

Research by Reinders and Ellis (2009) explored TE's impact on grammar learning in adult learners. They found that interventions like boldfacing target grammatical structures improved learners' noticing and comprehension, ultimately enhancing grammatical accuracy in production.

Recent studies have also investigated TE's effects on vocabulary acquisition. Meguro (2019) conducted a study with Japanese learners of English, revealing that TE increased learners' attention to target vocabulary, leading to improved retention and recall.

Working Memory

Working memory can be largely described as the temporary storage and processing of incoming information needed for functioning complex cognitive tasks (Baddeley, 2007), such as learning a foreign language (Hasegawa et al., 2002). Although substantial controversy exists on the precise conceptualization of WM, three types of verbal WM are chiefly pertinent to L2 learning. Single resource prototypes (Just & Carpenter, 2002) highlight a trade-off between processing and storage, whereas multiple-resource prototypes (Waters & Caplan, 2013) consider that processing and storage count on distinct resource pools. Baddeley (2007) split WM into a central executive and three slave systems with independent sizes: two short-term storage systems (phonological loop, visuo-spatial sketchpad) and an episodic buffer that assimilates the flow of information between long-term storage and the two short-term memory systems. Among other tasks, the dominant executive organizes information to and from its slave systems and agrees what information to choose and when it is time to change tasks or retrieval plans. Because the central executive incorporates an immense array of functions, some investigators proposed that it might consist of several constituents. For instance, Waters and Caplan (2013) divided Baddeley's central executive into a syntactic processing source and an attention control source and claimed that WM capacity does not influence syntactic processing efficiency (but L2 studies are mixed: Juffs & Harrington, 2011). There is proof that WM affects SL reading

comprehension, processing of SL morphological arrangement, and explicit SL rehearses (Leeser, 2007; Roehr, 2008; Waters & Caplan, 2013).

In summary, while previous studies cast doubt on the perceptual saliency and the effectiveness of recasts and TE, at least for certain instructional settings, it is desirable to understand whether their combination along with the role of WM would enhance the effectiveness of teaching.

Research Design

The present study was done through a quasi-experimental design including five intact classes nominating as four experimental groups and one control condition. At first, the participants took the pre-test and then attended five treatment sessions. After the last treatment session, they took the post-test, followed by stimulated recall interviews. Later, they completed a Working Memory Questionnaire (WMQ). Before the main study, pilot research was run to check the validity of the instruments.

Participants

In this study, the target population consisted of 100 Iranian EFL learners who were enrolled in five intermediate level classes in an online language school in Iran. The participants included 56 females and 44 males, ranging in age from 15 to 35 years old. A convenient sampling method was used to select individuals who met the research criteria, which included proficiency in English. Participants who had studied standard general English for at least five years with the language school were selected for the study. However, to ensure that all participants had a similar level of English proficiency, a placement test was also administered. All participants' native language was Persian to minimize any systematic bias. An EFL teacher with native-like proficiency was recruited to assist the researcher in conducting the experiment, in order to provide consistent teaching quality and a comprehensive impression of the entire treatment.

Instruments

To obtain a more comprehensive understanding of the learners' L2 knowledge, researchers have suggested using multiple

measures instead of relying on a single measure (Ellis et al., 2006; Norris & Ortega, 2001). Consequently, both oral and written tasks were employed in this study. The inclusion of a writing test allowed for a more extensive evaluation of the learners' language proficiency as it requires time for planning and editing, which are not present in an oral task. A detailed description of each instrument is provided below.

Oxford Test of English:

Prior to the treatment, students were administered two separate English tests to assess their level of homogeneity. These tests included a grammar level test and a vocabulary level test. Each test consisted of 40 multiple-choice questions (Oxford Test of English, 2021). The website provided an overall score and proficiency level for each student in the test results section. For example: "You have reached 28 out of 40 points (70%). Your level is B1 (Intermediate)."

The Picture Description Task (PDT)

This oral task was applied to measure learners' ability to use the target forms in a meaning-focused free production task. Learners were asked to describe sequenced pictures depicting a story within seven minutes based on the results of the pilot study. Before starting the test, the teacher asked some comprehension checking questions (such as how many photos do we have?; where are they?) in combination with learners' L1 to assure that the participants understood the concept and aim of the task. Then each learner individually narrated the story with the help of the pictures. The researcher with the help of three experts attempted to choose materials of similar difficulty level for the pretest and posttest, containing target vocabularies and grammar structures with a simple theme. The testing sessions were recorded and transcribed for further analysis. Obligatory contexts for using target vocabularies and forms were determined. This obligatory aspect was confirmed in the pilot study with students at a similar level of language proficiency. Since descriptions involved spontaneous use of the language with a main focus on meaning and without requiring learners to

analyze language structures, the tasks were assumed to engage implicit knowledge more than explicit knowledge (Ellis et al., 2006; 2009).

Learners' scores were measured by dividing the number of correct target forms by the total number of obligatory contexts plus the non-obligatory contexts in which the inappropriate target forms were used in the learners' production (Rassaei, 2020). A second rater also rated 20 percent of the tests to assure inter-rater reliability.

The Story Writing Task (SWT)

In this study, the SWT was utilized to assess the learners' ability to produce the target forms through a writing test. For this purpose, participants were provided with a short narrative to read during each testing session, which lasted for approximately fifteen minutes. The narratives had an average length of 700 words. Prior to commencing the test, the teacher asked several comprehension questions in the learners' native language (Persian) to ensure their understanding of the stories. Following this, participants proceeded to rewrite the story. The same procedure as the PDT was implemented to confirm the validity of the tests and to score the collected data. The SWT engaged both explicit and implicit knowledge, as it required spontaneous production with a focus on conveying meaning and allowed students sufficient time to monitor their written output.

Stimulated Recall Interviews

Another objective of this research was to examine and compare learners' perceptions of different treatments. To achieve this, stimulated recall interviews were conducted after the posttests to gather participants' insights on the extensive and intensive vocabulary and grammar recasts during the treatment sessions. To facilitate this process, the treatment sessions were recorded. Participants were asked to listen to the recorded segments and provide explanations of what transpired during the interaction sessions. Each student's interaction from the recorded sessions was carefully reviewed to identify the treatment conditions. Participants were then shown the relevant segment and

requested to elaborate on the events and their level of awareness of the conditions.

Prior to the interviews, learners received instructions that included the following guidelines: (1) pay close attention to each video clip containing an interaction episode from your storytelling sessions, (2) recall the specific moment when you were involved in the interaction episode, and (3) provide a detailed explanation of what took place during the interaction (Rassaei, 2020). The interviews were conducted in the learners' native language to facilitate communication and reporting. Participants were also allowed to pause and replay the audio as needed. Each interview lasted approximately 20 minutes.

Working Memory Questionnaire

A revised Persian equivalent of the Working Memory Questionnaire (WMQ) designed by Vallat-Azouvi et al. (2012) was applied after the experiment. This questionnaire included three dimensions: short-term storage, attention, and executive control. Equal number of questions in a fixed, pseudo-random order were presented to avoid any response bias (Vallat-Azouvi et al., 2012). The first part (short-term storage) examined the ability to retain information in short-term memory for a short period of time (e.g., "Do you have problems with remembering sequences of numbers, for example, when you have to note down a telephone number?"). The second domain (attention) addressed distraction and mental processing issues (e.g., "Do you need to make an effort to concentrate in order to follow a conversation in which you are participating with many other people?"). The third domain (executive control) was related to decision making and planning subjects (e.g., "When you are carrying out an activity, if you realize that you are making a mistake, do you find it difficult to change strategy?").

Each question was rated on a five-point Likert-type scale, ranging from 0 ("not at all") to 4 ("Extremely"). Three sub-scores were computed for each of the three domains (maximal score 40 for each), as well as a total score (out of 120). Higher scores were corresponded to more

difficulties/complaints. The study reported the reliability of the WMQ as 0.89 measured by Cronbach's alpha (Vallat-Azouvi et al., 2012).

Target Linguistic Structures

The focus of this study was on the grammatical and lexical functions of English language. Specifically, it focused on the usage of the indefinite article "a" used to introduce someone or something for the first time, and the definite article "the" used to refer to someone or something that has been mentioned before. Therefore, English articles and topic-related vocabulary were chosen as target structures, as they can be conveniently elicited from learners during communicative and meaning-focused tasks.

Reliability of the Instruments

To ascertain the instruments' validity, the confirmation of two TEFL experts were sought. Two native speakers also checked and approved the oral and written tasks for obligatory contexts. The reliability of the Oxford tests (i.e., vocabulary and grammar) was assessed by measuring Cronbach's alpha on the 40 items of each test separately. Estimation of Cronbach's alpha reliability was shown to be 0.82 and 0.86 for the vocabulary and grammar tests, respectively.

The inter-rater reliability for the PDT and SWT was checked by a second rater who independently rated 20 percent of all students' productions in pre-test and post-test. An estimation of 0.92 agreement in PDT and 0.96 agreement in SWT was indicated between the two raters by estimating simple percentage agreement. The correlation between the control group participants' scores in the pre-test and post-test also revealed an estimation of test-retest reliability of 0.86 for the PDT and 0.79 for the SWT, respectively.

The reliability of the working memory questionnaire was assessed by measuring Cronbach's alpha on the 30 items of the scale. The estimation of Cronbach's alpha reliability was 0.87 indicating a good reliability. To further evaluate the internal validity of the scale, the correlations between the three sub-scores and the total score of the scale were calculated. The three sub-scores

were significantly correlated with each other (Spearman's Rho ranging from 0.45 to 0.59, Sig. = .0001) and also significantly correlated with the total score (Spearman's Rho ranging from 0.79 to 0.89, Sig. = .0001).

Data Collection Procedure

In the first place, a thorough explanation of the aim of the research was given, it was also clarified that they can drop out of the experiment at any point. Upon their written consent, the researcher performed a pilot study before the main study in order to remove the potential obstacles and problems with learners at a similar level of language proficiency. Two separate Oxford online tests were applied in the beginning. Participants who were found to be homogeneous participated in the PDT and SWT. The intensive recast group received feedback on recasts only on articles and theme words whereas the extensive recast group had feedback on both article and lexical errors and any other errors that happened incidentally during communication. Textual enhancement (TE) was done by boldfacing and underlining the target structures and words with a different color in the text. The intensive TE group observed TEs only on articles and theme words whereas the extensive TE group had TE on both article and target words in a phrase or sentence. The control group had no feedback. Table 1 demonstrates the number of target structures for each treatment sessions.

Table 1

The Number of Target Structures for Each Treatment Session

	Definite Articles	Indefinite Articles	Vocabularies
Session 1	30	22	18
Session 2	35	20	18
Session 3	40	25	17
Session 4	28	19	18
Session 5	30	22	16

Accordingly, five short stories in three versions were prepared for the present study. One version, containing no textual enhancement, was taught to

the control group and the intensive and extensive recasts groups. The second version contained intensive TEs on target grammatical and lexical forms. The third version included extensive TEs on both article and target words in a phrase or sentence. The procedures for each group are described below:

The Recasts Groups

For the intensive and extensive recast condition, learners were provided with a short story of about 700 words in each treatment session and were asked to read the text for 10 minutes. The story was removed from the screen afterward and they were equally divided into groups and participated in retelling the whole story. The teacher was cognizant enough to give equal exposure chance to all the students. Other students who were present in the class were asked to listen to the group presentations. Each group storytelling lasted about five minutes. While narrating the story, the teacher provided an intensive or extensive recast, based on the group, to each target error produced by the students. The aim was to provide equal input and recast to the whole class. Besides the target form errors, some nontarget errors such as lack of subject-verb agreement, the use of inappropriate tense and some lexical errors were corrected via recasts by the instructor in the extensive group. The type of recasts was mainly declarative with no additional stress, repetition, or verbal prompts (Nassaji, 2009; Loewen & Philp, 2006; Sheen, 2008). Moreover, all recasts' sessions were recorded for stimulated recall interviews.

Tables 2 and 3 demonstrate examples of recasts employed in the extensive and intensive recast groups. Table 2 indicates examples from students in the extensive recast group which covered feedback on the other occurred errors as well as the target linguistic structures (i.e., article and target vocabulary usage) while in Table 3 the instructor offered no feedback on non-target forms.

Table 2

Examples from the Extensive Recast Group

Example 1: Target Grammar Feedback Student: Next to <i>tree</i> , there is a girl. Teacher: Next to the tree, there is a girl.(recast)	Example 2: Target Vocabulary Feedback Student: The man is <i>very very hungry</i> . Teacher: The man is starving. (recast)
Example 3: General Feedback	
Student: The weather is <i>winding</i> . Teacher: The weather is windy. (recast)	

Table 3

Examples from the Intensive Recast Group

Example 1: Target Grammar Feedback Teacher: Why do you think your opponent has chosen Maria? Student: Well, Maria is <i>female</i> who <i>wear</i> a green hat. Teacher: A female. (recast)	Example 2: Target Vocabulary Feedback Student: The other ducklings were grey and <i>very very soft and no weight</i> . Teacher: Grey and fluffy. (recast)
Example 3: No Feedback	
Student: She <i>wear</i> glasses. Teacher: Does she wear a hat, too? (no feedback)	

The TEs Groups

For the intensive and extensive TE condition, learners were asked to read the short story within the same time limit and present the story in the same way without receiving recasts for their errors. TEs were operationalized through coloring, underlining and boldfacing. As some examples in

Table 4 indicate, the intensive TE group received textual enhancement only on the target forms (i.e., target vocabulary and article), however, the extensive TE group received textual enhancement on the phrases or sentences which included those target linguistic structures in the texts that the students were required to read.

Table 4

Examples from the TE group

Example 1: Intensive TE Condition As I rode along the highway between Roche and Carthew, I was surprised to see a light coming along behind me in the dark. As the moon came out from behind a cloud, I could make out two galloping horsemen, one holding a light in his hand. At first I imagined that they had been sent out to bring me back to help some other sick person in one of the villages I had just visited. As the two riders approached,....	Example 2: Extensive TE Condition As I rode along the highway between Roche and Carthew, I was surprised to see a light coming along behind me in the dark. As the moon came out from behind a cloud, I could make out two galloping horsemen , one holding a light in his hand. At first I imagined that they had been sent out to bring me back to help some other sick person in one of the villages I had just visited. As the two riders approached,....
--	---

The Control Group

The materials for the control group were similar to the TE procedure except that the stories contained no textually enhanced target forms. The same teaching sessions were presented without TE and recast conditions.

Results

To ensure the normal distribution of data, the researcher conducted a statistical analysis using the

Kolmogorov-Smirnov test based on the results of the Oxford tests of English. The purpose was to examine whether there were any significant differences between the groups in terms of data distribution. The results of the Kolmogorov-Smirnov test, as shown in Table 5, indicate that there were no significant differences among the groups. Therefore, it can be concluded that all five groups exhibited normal and homogeneous data distribution prior to the experiment (with a significance value higher than 0.05).

Table 5
Tests of Normality

		Kolmogorov-Smirnov ^a	
		Statistics	Sig.
Intensive TE Condition	Grammar Test	.976	.142
	Vocabulary Test	.945	.210
Extensive TE Condition	Grammar Test	.961	.251
	Vocabulary Test	.899	.112
Intensive Recasts Condition	Grammar Test	.985	.423
	Vocabulary Test	.971	.513
Extensive Recasts Condition	Grammar Test	.971	.650
	Vocabulary Test	.891	.328
Control Condition	Grammar Test	.980	.243
	Vocabulary Test	.912	.142

To address the initial research question, which investigates the impact of intensive or extensive recasts and/or textual enhancement (TE) on the improvement of vocabulary and grammar among EFL students, descriptive and inferential statistics were utilized to analyze the learners' scores in both

the PDT and the SWT for each group. The findings of the statistical analysis are outlined below, providing insights into the effects of the different treatments on the participants' performance.

Table 6
Descriptive Statistics for the Students' Performance in the PDT and SWT

	N	Pre-Test		Post-Test	
		Mean	SD	Mean	SD
Intensive TE Condition	20	PDT: 36	6.8	49	6.1
		SWT: 34	7.9	50	5.6
Extensive TE Condition	21	PDT: 35	9.1	44	7.6
		SWT: 33	6.1	41	8.3
Intensive Recasts Condition	20	PDT: 36	7.7	58	7.2
		SWT: 35	8.1	60	4.2
Extensive Recasts Condition	20	PDT: 35	6.8	53	5.9
		SWT: 34	6.9	56	4.9
Control Condition	19	PDT: 34	7.1	35	6.8
		SWT: 36	7.5	38	8.1

According to Tables 6 the experimental groups' mean scores improved from the pre-test to the post-test. Moreover, the mean scores belonging to the intensive vocabulary and grammar recasts condition seems to be higher than the other groups. In this vein, the highest mean scores in other conditions indicate extensive vocabulary and grammar recasts, intensive vocabulary and grammar TE, extensive vocabulary and grammar TE and control group, respectively.

To assess students' improvement over time as a result of the treatment conditions, mixed between-within group ANOVA was performed on participants' pre-test and post-test scores. In

relation to the PDT, the findings revealed main effects for time $F(2, 98) = 151$, $\text{Sig}. < .001$, $\eta^2 = .58$, the treatment groups $F(2, 98) = 22.1$, $\text{Sig}. < .001$, $\eta^2 = .57$, and the interaction between time and the treatment groups $F(2, 98) = 25.3$, $\text{Sig}. < .001$, $\eta^2 = .51$. These findings present proof that the participants' English language knowledge improved over time in effect of the treatment conditions. They also highlight the fact that the four treatment conditions had differential effects on students' improvement over time. Table 7 summarizes the results.

Table 7*Mixed Between-Within ANOVA for Improvement over Time (PDT)*

Measure	Degrees of Freedom	F-value	Significance	η^2
Time		151	.001	.58
Treatment Groups	2.98	22.1	.001	.57
Time \times Treatment Groups		25.3	.001	.51

One-way ANOVA on the post-test scores attained from the PDT indicated a significant effect for the treatment conditions, a summary of the results is presented in the table below.

Table 8*One-way ANOVA Results for PDT Post-test Scores*

Source	df	F-value	Sig.	η^2
Treatment	2.98	46.2	.001	.56

Tukey's post hoc comparisons revealed that the intensive vocabulary and grammar recasts condition and also the extensive vocabulary and grammar recasts condition significantly outperformed the control condition (Sig.<.001). Among the experimental groups, the group which received the intensive vocabulary and grammar recasts condition significantly outscored the other

three experimental groups (Sig.<.003). Moreover, the findings indicated that the extensive vocabulary and grammar recasts condition significantly outperformed the intensive vocabulary and grammar TE condition and extensive vocabulary and grammar TE condition (Sig.<.003). Meanwhile, no significant difference was found between the extensive vocabulary and grammar TE condition and control conditions (Sig.=.14).

Concerning the SWT, the results of mixed between-within group ANOVA revealed main effects for time, $F = 48.1$, Sig. <.001, $\eta^2=.34$, treatment conditions, $F (2.98) = 20.7$, Sig. <.001, $\eta^2=.34$, and also the interaction between time and treatment conditions, $F (2, 98) = 20.6$, Sig. <.001, $\eta^2=.44$. Table 9 further displays the students' enhancement over time based on the SWT.

Table 9*Mixed Between-Within ANOVA Results for SWT*

Measure	Degrees of Freedom	F-value	Significance	η^2
Time		151	.001	.34
Treatment Groups	2.98	22.1	.001	.34
Time \times Treatment Groups		25.3	.001	.44

One-way ANOVA on the post-test scores achieved from the SWT revealed significant differences among the groups, $F (2.98) = 39.3$, Sig.<.001, $\eta^2=.66$. The findings of Tukey's post hoc analysis indicated that the intensive vocabulary and grammar recasts group significantly outperformed the extensive vocabulary and grammar recasts group (Sig.<.003), the intensive vocabulary and grammar TE group (Sig.<.003), the extensive vocabulary and grammar TE group (Sig.<.05) as well as the control group (Sig.<.001). Moreover, the findings revealed that the extensive vocabulary and grammar recasts condition outscored the intensive vocabulary and grammar TE condition (Sig.<.001), the extensive vocabulary

and grammar TE condition (Sig.<.05), and the control group (Sig.<.001). However, no significant difference was discovered between the performance of students in the extensive vocabulary and grammar TE group and the control group (Sig.=.98).

To answer the second research question in relation to the effect of different treatment conditions, the effect sizes were sought to differentiate between groups. The effect size describes the strength of the difference between groups, or the influence of the independent variable. Cohen's D describes the difference between groups in terms of standard deviation unit: Small 0.2; Medium 0.5; and Large 0.8 (Pallant,

2013). Table 10 demonstrates Cohen’s d effect size for comparisons between groups in the post-test.

Table 10
Summary of Effect Sizes (Cohen’s D)

	PDT	SWT
Intensive Recasts vs. Extensive Recasts	1.1	1.2
Intensive Recasts vs. Intensive TE	1.9	1.6
Intensive Recasts vs. Extensive TE	1.8	1.9
Intensive Recasts vs. Control	2.1	2.3
Extensive Recasts vs. Intensive TE	0.7	0.5
Extensive Recasts vs. Extensive TE	0.8	0.6
Intensive TE vs. Extensive TE	0.9	0.8

As Table 10 indicates, the effect size values for the combined language components (i.e., vocabulary and grammar) in the intensive recasts condition compared to the other three treatment conditions, separately, is large to very large.

In order to check students’ perceptions of the different treatment recasts’ conditions, their comments were coded and analyzed. Their comments were divided into three categories and coded a) full comprehension of the corrective feedback (CF), b) partial comprehension of the corrective feedback, and c) absent comprehension of the corrective feedback. If a student was able to recognize the source of the error for which he or she received a recast, that comment was coded as "full comprehension". Alternatively, if a learner was able to recognize the corrective nature of a recast

but failed to express the error for which he/she received the recast, that comment was categorized as "partial comprehension". Finally, if a participant failed to identify a recast as a CF in response to his/her error that case was coded as "absent comprehension". The following examples show how the three categories of students' perceptions were coded in the current research:

a) Full comprehension:

Interviewee: I made a mistake and you corrected me. Interviewer: What was your mistake?

Interviewee: Instead of saying the bird, I said a bird.

b) Partial comprehension:

Interviewee: I think I said something wrong and you corrected me. Interviewer: What was your mistake?

Interviewee: .[silence]

c) Absent comprehension:

Interviewee: You repeated what I said. Interviewer: What for?

Interviewee: I don't know!

In general, a total of 650 stimulated recall comments were found and coded for analysis. In line with the aim of the third question (How do students perceive the effect of intensive or extensive recasts on EFL development?) and concerning the accuracy of the students' insights of the recasts they received within the treatment period, Table 11 indicates the frequencies for the three categories of students' perceptions coded based on the stimulated recall interview data.

Table 11
The Frequency of Students’ Perceptions of Recasts

	Full Comprehension		Partial Comprehension		Absent Comprehension	
	F	%	f	%	F	%
Extensive Recasts	106	32.01	179	54.02	37	30.80
Intensive Recasts	189	56.10	136	38.99	19	7.30

As Table 11 shows, the intensive vocabulary and grammar recasts condition had a more accurate perception of recasts. While the participants of the extensive vocabulary and grammar recasts group could notice 32.01 percent of the errors that they received, the learners of the intensive vocabulary and grammar recasts group

could notice 56.10 percent of the recasts they received.

Moreover, while the participants of the extensive recasts condition were not successful at recognizing 30.80 percent of the recasts they received as CF, the participants of the intensive recasts group failed to recognize the corrective

nature of only 7.30 percent of the recasts they received. The results of chi-square analysis are presented below.

Table 12*Chi-Square Results of Intensive and. Extensive Recasts*

	Chi-Square (χ^2)	Significance (p-value)	Effect Size (Cramer's V)
Intensive vs. Extensive Recasts	35.3	.003	0.23

Accordingly, the data revealed a significant difference between the intensive recasts and

extensive recasts conditions in relation to the accuracy of the students' perceptions of recasts, $\chi^2(2, 761) = 35.3, p < .003$, Cramer's $V = .23$.

The descriptive information for the mean scores on the three working memory measures (i.e., storage domain (SD), attention domain (AD), and executive domain (ED)) are presented in Table 13. Pearson product-moment correlation coefficient analyses demonstrated a significant correlation between students' performance on the SD and AD tests ($r = .35$, Sig. $< .05$), but the scores on neither of these measures were revealed to be correlated with the ED results.

Table 13*Descriptive Statistics for the Working Memory Measures*

WM Measure	Condition	N	Mean	SD
Storage Domain	Intensive TE	20	6.16	0.93
	Extensive TE	21	6.06	0.94
	Intensive Recasts	20	6.17	0.95
	Extensive Recasts	20	6.12	0.79
	Control	19	5.72	0.97
Attention Domain	Intensive TE	20	6.45	0.88
	Extensive TE	21	6.52	0.84
	Intensive Recasts	20	6.09	0.95
	Extensive Recasts	20	5.61	0.92
	Control	19	6.01	1.00
Executive Domain	Intensive TE	20	3.08	0.51
	Extensive TE	21	2.98	0.82
	Intensive Recasts	20	3.01	0.61
	Extensive Recasts	20	3.12	0.43
	Control	19	2.87	0.53

The working memory mean scores for all conditions were in a similar range in SD and AD sets (ranging between 5.61 to 6.52) slightly outperforming the AD section ED (ranging

between 2.87 to 3.12). However, one-way ANOVA found no significant difference between the groups for any of the working memory sections (Sig. = .47).

Table 14*Correlations of Working Memory Test Scores and Posttest Scores*

Condition	WM Questionnaire	PDT	SWT
Intensive Recasts	SD	.71 Sig. = .003	.47 Sig. = .001
	AD	.49 Sig. = .001	.75 Sig. = .000
	ED	.08	-.11
Extensive Recasts	SD	.10	.11
	AD	.21	-.16
	ED	.09	.12
Intensive TE	SD	.30	-.28
	AD	-.30	.25
	ED	.27	.31
Extensive TE	SD	.30	-.28
	AD	.22	.28

Condition	WM Questionnaire	PDT	SWT
Control	ED	.30	.27
	SD	.19	.17
	AD	-.08	.12
	ED	.14	.17

Table 14 provides the correlations between the working memory measures and the posttest gain scores. As shown, in the intensive recast group, the extent of participants' development on the PDT and SWT tests showed medium to large positive correlations with their performance on the SD (PDT: $r = .71$; Sig. < .003; SWT: $r = .47$; Sig. < .001) and AD sections (PDT: $r = .49$; Sig. < .001; SWT: $r = .75$; Sig. < .000) of the test but no significant correlations were detected between the ED section and the PDT and SWT tests. However, none of the other components were shown to be significantly correlated with each other.

In sum, working memory did not appear to be associated with the gain scores of students who were not exposed to intensive recasts, but it was associated with the extent of development achieved by learners who received this type of recasts. Specifically, participants with higher SD showed more substantial development on oral tasks (PDT: $r = .71$; Sig. < .003) with less difficulties, whereas participants with higher AD were observed to exhibit greater gains on the written tasks (SWT: $r = .75$; Sig. < .000) with less complaints.

Discussion and Conclusion

The noteworthy effects revealed for recasts are important in light of the theoretical debates concerning the role of recasts in the literature (Lyster & Ranta, 2013; Goo & Mackey, 2013; Nassaji, 2017) and confirm that recasts are beneficial in helping learners identify their errors and enhance their accuracy in the use of the target form. The ANOVA results highlighted the positive effect of the intensive vocabulary and grammar recasts condition significantly over the other three experimental groups evident in both oral and written tests. These findings suggest that intensive recasts had a more pronounced positive effect than did extensive recasts.

This challenges the idea that extensive recasts are necessarily more effective than intensive recasts

discovered in a counterbalance study done by Nassaji (2017), however, he examined the effect of grammar solely. This challenging concept also requires potential explanations for the advantages behind it as all conditions received the same material. It should be noted that both extensive and intensive recasts were presented in the form of declarative recasts with no additional prompts or stress and the number of partial and full recasts was similar. Therefore, the differential effects cannot be attributed to the frequency of recasts on target errors.

It can be argued that while TE just provides students with positive evidence in addition to facilitating noticing target forms precisely or in a sentence, both types of recasts encourage L2 development in a better way by not only presenting students with positive evidence and facilitating noticing of target structures, but also through other mechanisms such as providing students with negative evidence and also triggering students' modified output in either condition. The results of the present research provide some evidence that positive evidence by itself in its enhanced form (i.e., TE) whether intensive or extensive may not be sufficient for L2 development. However, the intensive and direct form again might be more effective in between the two TE conditions based on the results.

The results of the stimulated recall interviews revealed that students who were exposed to the intensive grammar and vocabulary recasts condition had a significantly more accurate perception of recasts compared to the extensive recasts condition. These findings present further evidence that when recasts are intensive, their saliency to the learners as corrective feedback increases significantly. Moreover, the stimulated recall data suggests that the intensive recasts condition outperformed the extensive recasts group in the present study since the students of the intensive recasts condition could more accurately

notice recasts as corrective feedback compared with the challenge of students in the extensive recasts group in detecting the mismatches between the non-target-like forms and the target-like forms delivered through recasts. While the findings of the present study did not present convincing evidence for the benefits of TE treatments, previous research presented mixed results as to the effectiveness of TE with some studies providing evidence for its effectiveness (e.g., LaBrozzi, 2016; Lee, 2007; Simard, 2009). However, as stated by Simard (2009), such contradictory results can be due to different methodological options and also various measures that previous studies used to evaluate TE. Another factor that can affect the result of TE is the target linguistic structures. As hinted by Han et al. (2008), some linguistic structures in input are fundamentally non-salient. In this regard, articles might be considered as non-salient English targets due to their low communicative value. The concept of vocabulary, however, can be explained in terms of the relative short length of the study as some learners might need more time for exposure, processing, comprehension and use of the target terms.

The fact that working memory did not seem to be related to the gain scores of learners who did not receive intensive recasts and their substantial development on oral tasks with storage domain and on written tasks with attention domain is consistent with some previous studies. Therefore, Lyster and Saito (2010) in their meta-analysis of feedback studies discovered more effects of oral feedback on spontaneous oral production tests than on other tests. Révész (2012) also found more improvements on oral than on written tasks in her comparison of the effect of recasts on these two types of measures. As she explained, the benefit of recasts on oral tests can be interpreted in terms of the similarity of the contexts of the oral test and the oral recast and the transfer of learning from one context to another similar context. This explanation is in line with the Transfer Appropriate Processing (TAP) perspective, which recommends that L2 learning is context-dependent and that L2 knowledge acquired in one context can be better used in a context that engages similar cognitive

processes (Lightbown, 2008). In this vein, the effect of examined tasks seem to have affected the participants' performance in the tests.

In this study, the effects of recasts were observed on both the oral tasks and written tasks. These two measures may tap into different types of knowledge. The writing tasks may tap into explicit knowledge since they require learners to reflect on what they are going to write by paying attention to target linguistic forms to identify and correct errors, and the oral production task may tap more into implicit knowledge, as it involves using the language form with a primary focus on expressing meaning rather than language. However, as stated by Ellis (2005), when it comes to the effects of feedback on different types of knowledge, research findings should be treated cautiously. Although a theoretical distinction can be made between implicit and explicit knowledge, in reality it is difficult to differentiate the two as students may rely on both types of knowledge in different extents during any test or language task (Ellis et al., 2009). Future studies can evaluate these effects with respect to other types of feedback such as elicitation, metalinguistic feedback, or explicit correction.

References

- Anderson, J. R. (2000). *Learning and memory: An integrated approach*. Wiley.
- Baddeley, A. D. (2007). *Working memory, thought, and action*. Oxford University Press.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27(2), 141-172. doi:10.1017/s0272263105050096
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51, 281-318. <https://doi.org/10.1111/0023-8333.00170>
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28(2), 339-368. <https://doi.org/10.1017/S0272263106060141>
- Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching*. Multilingual Matters.

- Han, Z., Park, E., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics*, *29*, 597-618. <https://doi.org/10.1093/applin/amn029>
- Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *Neuroimage*, *15*(3), 647-660. <https://doi.org/10.1006/nimg.2001.1008>
- Hawkes, L., & Nassaji, H. (2016). The role of extensive recasts in error detection and correction by adult ESL students. *Studies in Second Language Learning and Teaching*, *6*, 19-41. <https://doi.org/10.14746/ssllt.2016.6.1.2>
- Hosseini, K. (2019). The effects of recasts on L2 learners' acquisition of English wh-questions: A systematic review. *Journal of Second Language Teaching and Research*, *6*(2), 221-243. doi:10.1080/17501229.2017.1419897
- Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*(2), 137-166. doi:10.1017/S026144481000052X
- Just, M. A., & Carpenter, P. A. (2002). A capacity theory of comprehension: Individual differences in working memory. In T. A. Polk & C. M. Seifert (Eds.), *Cognitive modeling* (pp. 131-177). Boston Review.
- LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research*, *20*(1), 75-91. <https://doi.org/10.1177/1362168814549774>
- Lee, H.-S., & Huang, J. (2008). Textual enhancement of input: Issues and possibilities. *Journal of Second Language Writing*, *17*(2), 79-98. doi:10.1016/j.jslw.2007.11.004
- Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, *57*(1), 87-118. <https://doi.org/10.1111/j.1467-9922.2007.00400.x>
- Leeser, M. J. (2007). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, *11*(3), 321-342. doi:10.1177/1362168807079377
- Lightbown, P. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27-44). Multilingual Matters.
- Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *The Modern Language Journal*, *90*(4), 536-556. <https://doi.org/10.1111/j.1540-4781.2006.00456.x>
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *26*(3), 399-432. doi:10.1017/S0272263104263021
- Lyster, R., & Ranta, L. (2013). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*, 37-66.
- 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>
- Mackey, A. (2007). *Conversational interaction in second language acquisition: A series of empirical studies*. Oxford University Press.
- Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning*. Oxford University Press.
- Mackey, A. (2016). Feedback, noticing and second language development. *Applied Linguistics*, *37*(6), 791-807. doi:10.1093/applin/amu062
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407-452). Oxford University Press.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, *22*, 471-497. <https://doi.org/10.1017/S0272263100004029>
- Meguro, Y. (2019). Textual enhancement, grammar learning, reading comprehension, and tag questions. *Language Teaching Research*, *23*(1), 58-77. doi: 1362168817714277
- Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair in dyadic interaction. *Language Learning*, *57*, 511-548. <https://doi.org/10.1111/j.1467-9922.2007.00429.x>
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, *59*, 411-452. <https://doi.org/10.1111/j.1467-9922.2009.00511.x>
- Nassaji, H. (2017). The effectiveness of extensive versus intensive recasts for learning L2 grammar. *The Modern Language Journal*, *101*(2), 353-368. <https://doi.org/10.1111/modl.12378>

- Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. Routledge.
- Nguyen, M. T. T., Pham, H. T., & Pham, T. M. (2017). The effects of input enhancement and recasts on the development of second language pragmatic competence. *Innovation in Language Learning and Teaching, 11*(1), 45-67. <https://doi.org/10.1080/17501229.2016.1248717>
- Norris, J. M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning, 51*, 157-213.
- Ortega, L. (2005). *Understanding second language acquisition*. Hodder Arnold.
- Oxford test of English. (2021). *English level test*. Oxford.
- Pallant, J. (2013). *SPSS survival manual*. Mc Graw-hill education.
- Rassaei, E. (2013). Corrective feedback, learners' perceptions, and second language development. *System, 41*(2), 472-483. <https://doi.org/10.1016/j.system.2013.04.006>
- Rassaei, E. (2015). Effects of textual enhancement and input enrichment on L2 development. *TESOL Journal, 6*(2), 281-301. <https://doi.org/10.1002/tesj.157>
- Rassaei, E. (2020). The separate and combined effects of recasts and textual enhancement as two focus on form techniques on L2 development. *System, 89*, 102-193. <https://doi.org/10.1016/j.system.2019.102193>
- Rebuschat, P. (2013). *Implicit and explicit learning of languages*. John Benjamins Publishing.
- Reinders, H., & Ellis, R. (2009). The effects of two types of input on intake and the acquisition of implicit and explicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & R. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 282- 302). Multilingual Matters.
- Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning, 62*(1), 93-132. <https://doi.org/10.1111/j.1467-9922.2011.00672.x>
- Roehr, K. (2008). Linguistic and metalinguistic categories in second language learning. *Cognitive Linguistics, 19*, 67-106.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.
- Segalowitz, N., & Hulstijn, J. H. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: psycholinguistic approaches* (pp. 371-388). Oxford University Press.
- Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning, 58*, 835-874. <https://doi.org/10.1111/j.1467-9922.2008.00473.x>
- Sheen, Y. (2018). Recasts in second language learning: A meta-analysis. *Language Learning, 68*(4), 868-907. doi:10.1111/lang.12296
- Simard, D. (2009). Differential effects of textual enhancement formats on intake. *System, 37*(1), 124-135. <https://doi.org/10.1016/j.system.2008.08.008>
- Spada, N. (2015). SLA research and L2 pedagogy: Misapplications and questions of relevance. *Language Teaching, 48*(1), 69-81. <https://doi.org/10.1017/S0261444814000285>
- Vallat-Azouvi, C., Pradat-Diehl, P., & Azouvi, P. (2012). The Working Memory Questionnaire: A scale to assess everyday life problems related to deficits of working memory in brain injured patients. *Neuropsychological Rehabilitation, 22*(4), 634-649. <https://doi.org/10.1080/09602011.2012.667246>
- VanPatten, B. (2004). Input processing in adult second language acquisition. In VanPatten, B. (Ed.), *Processing instruction: theory, research, and commentary* (pp. 5-31). Lawrence Erlbaum Associates.
- Waters, G. S., & Caplan, D. (2013). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology: Section A, 49*(1), 51-79. doi:10.1080/713755607
- Wong, W., & VanPatten, B. (2003). The evidence is IN: Drills are OUT. *Foreign Language Annals, 36*(3), 403-423. doi:10.1111/j.1944-9720.2003.tb01933.