ISSN (print): *2588-5731*

Research Paper

The Effect of Intensive versus Extensive Recasts and Textual Enhancements on Iranian EFL Learners' Grammatical and Lexical Development: Exploring Students' Perceptions in Online Classes Considering the Role of Working Memory Capacity

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

Received Date: 17/07/2023 Accepted Date: 11/09/2023

Pp: 142-168

Abstract

This study addresses the dearth of research on the interplay between textual enhancements (TE), recasts, and working memory capacity in SLA. Investigating extensive versus intensive recasts and TEs, it explores their impact on grammatical and lexical growth in EFL context. Additionally, students' perceptions of recasts and the role of working memory capacity are evaluated. Experimental groups included intensive/extensive vocabulary and grammar TE/recasts conditions, alongside a control group. Pre/post-tests, interviews, and a working memory questionnaire were used. Results indicate recasts' facilitative effect, with intensive recasts outperforming extensive recasts and TEs in promoting grammatical and lexical development among high working memory capacity learners. Learners in the intensive recasts group had accurate error perceptions. Working memory correlated with intensive recasts' gains, with high Storage and Attention Domains showing development in oral and written tasks. These findings illuminate the efficacy of TE, recasts, and working memory, offering valuable insights for educators and learners alike.

***Keywords*:** Intense and extensive educational approaches; verbal revisions versus textual improvements; memory function, students' perceptions, online classes.

[1] - Department of Foreign Languages, Shiraz branch, Islamic Azad University, Shiraz, Iran.
[2] - Department of Foreign Languages, Shiraz branch, Islamic Azad University, Shiraz, Iran (Corresponding Author). mbvl57@gmail.com
[3] - Department of Foreign Languages, Shiraz branch, Islamic Azad University, Shiraz, Iran.
[4] - Department of Foreign Languages, Marvdasht branch, Islamic Azad University, Fars, Iran.

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).

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 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 (Anderson, 2000; Rebuschat, 2013). Challenges arise in transferring procedural knowledge to different contexts and skills (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 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 impacting L2 development (Sheen, 2018; Wong & VanPatten, 2003).

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. 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 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. 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.

The existence of two distinct approaches in providing textual enhancements (TE) and recasts in second language acquisition (SLA) encompass an extensive approach that covers a wide range of errors and an intensive approach that focuses on specific preselected errors. This distinction has been acknowledged in previous research (Ellis et al., 2001; Hawkes & Nassaji, 2016; Lyster & Ranta, 2013; Nassaji, 2007, 2009). However, what remains relatively unexplored is the comparison of the effectiveness of these extensive and intensive approaches in enhancing the saliency of linguistic structures and their impact on learners' development, particularly in relation to working memory (WM) capacity. The gap in the study lies in the need to investigate whether different levels of saliency achieved through TE and recasts lead to varying degrees of effectiveness in learners' development. To address this gap, the present study is designed to examine the effectiveness of extensive versus intensive recasts and TEs in the context of WM capacity, shedding light on the nuances of how these factors interact and impact SLA. The research questions aim to uncover the specific effects of these different feedback approaches and their relationship with learners' WM capacity, contributing to a deeper understanding of SLA processes.

Research Questions

1. Do the learners' scores change significantly from pre-test to post-test in intensive recast group?
2. Do the learners' scores change significantly from pre-test to post-test in extensive recast group?
3. Do the learners' scores change significantly from pre-test to post-test in intensive TE group?
4. Do the learners' scores change significantly from pre-test to post-test in extensive TE group?

5. Are there any significant differences among the effects of intensive recast, extensive recast, intensive TE and extensive TE on the development of EFL learners' vocabulary and grammar?
6. How do learners perceive the effectiveness of intensive and extensive recasts in relation to their comprehension of corrective feedback?
7. Does working memory capacity account for any differences between intensive recast, extensive recast, intensive TE and extensive TE across the storage, attention, and executive domains?

Research Design

The present study was done through a mixed method design including five intact classes nominating as four experimental groups and one control condition. The primary independent variables under investigation were the types of feedback: intensive recasts, extensive recasts, intensive textual enhancements (TE), and extensive textual enhancements (TE). These independent variables represented the diverse approaches to providing feedback. The study's dependent variables encompassed grammatical and lexical development, which were evaluated through pre-test and post-test scores in both oral and written tasks. Additionally, participants' perceptions of recasts were assessed via stimulated recall interviews, while their working memory capacity was measured using a dedicated questionnaire. The control group, serving as a baseline condition, comprised participants who did not receive specific feedback or treatment. These variables were systematically manipulated and measured to scrutinize the impact of distinct feedback methods (independent variables) on grammatical and lexical development, students' perceptions, and the role of working memory capacity (dependent variables).

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 convenience 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. One EFL teacher with native-like proficiency, following a communicative approach, taught the students to provide consistent teaching quality and a comprehensive impression of the entire treatment. However, based on the group assigned (i.e., intensive vocabulary and grammar TE condition, extensive vocabulary and grammar TE condition. intensive vocabulary and grammar recasts condition, extensive vocabulary and grammar recasts condition and control condition), the provision of feedback varied.

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 Placement Test: 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.

The 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 g (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 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.. 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	Example 2: Target Vocabulary Feedback
Student: Next to *tree*, there is a girl.	Student: The man is *very very hungry*.
Teacher: Next to the tree, there is a girl.(recast)	Teacher: The man is starving. (recast)
Example 3: General Feedback	
Student: The weather is *winding*.	

Teacher: The weather is windy. (recast)

Table 3- *Examples from the Intensive Recast Group*

Example 1: Target Grammar Feedback	Example 2: Target Vocabulary Feedback
Teacher: Why do you think your opponent has chosen Maria? Student: Well, Maria is *female* who *wear* a green hat. Teacher: A female. (recast)	Student: The other ducklings were grey and *very very soft and no weight*. Teacher: Grey and fluffy. (recast)
Example 3: No Feedback	
Student: She *wear* 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	Example 2: Extensive TE Condition
As I rode along the highway between Rocheand 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 I rode along the highway between Rocheand 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,….	**As the two riders** approached,….

Results

To address the initial research question, which investigates whether the learners' scores change significantly from the pre-test to the post-test in the intensive recast group (both the PDT and the SWT), a paired samples t-test was performed, using the data from the intensive recast group.

Table 5- *Results of Paired Samples t-test: Learners' Scores in the Intensive Recast Group*

Condition	N	Pre-test (Mean)	Pre-test (SD)	Post-test (Mean)	Post-test (SD)	t-value	df	p-value
Intensive Recast Group	20	36	7.7	58	7.2	61.39	19	<0.001

With 19 degrees of freedom, the critical value at a significance level of $\alpha = 0.05$ (two-tailed) is ±2.093. Since the calculated t-value (61.39) is much larger than the critical value, the null hypothesis (There is no significant difference in scores between the pre-test and post-test in the intensive recast group.) is rejected. Therefore, the results indicate a significant difference in scores between the pre-test and post-test in the intensive recast group ($t(19) = 61.39$, $p < 0.001$).

To answer the second research question which addresses whether the learners' scores change significantly from the pre-test to the post-test in the extensive recast group, a paired samples t-test was performed, using the data from the extensive recast group.

Table 6- *Results of Paired Samples t-test: Learners' Scores in the Extensive Recast Group*

Condition	N	Pre-test (Mean)	Pre-test (SD)	Post-test (Mean)	Post-test (SD)	t-value	df	p-value

Extensive Recast Group	20	35	6.8	53	5.9	12.01	19	<0.001

With 19 degrees of freedom, the critical value at a significance level of $\alpha = 0.05$ (two-tailed) is ±2.093. Since the calculated t-value (12.01) is much larger than the critical value, the null hypothesis (There is no significant difference in scores between the pre-test and post-test in the extensive recast group.) is rejected. Therefore, the results indicate a significant difference in scores between the pre-test and post-test in the extensive recast group ($t(19) = 12.01$, $p < 0.001$).

To address the third research question which seeks whether the learners' scores change significantly from the pre-test to the post-test in the intensive TE group, a paired samples t-test was applied, using the data from the intensive TE group.

Table 7- *Results of Paired Samples t-test: Learners' Scores in the Intensive TE Group*

Condition	N	Pre-test (Mean)	Pre-test (SD)	Post-test (Mean)	Post-test (SD)	t-value	df	p-value
Intensive TE Group	20	36	6.8	49	6.1	9.21	19	<0.001

With 19 degrees of freedom, the critical value at a significance level of $\alpha = 0.05$ (two-tailed) is ±2.093. Since the calculated t-value (9.21) is much larger than the critical value, the third null hypothesis (There is no significant difference in scores between the pre-test and post-test in the intensive TE group.) is rejected. Therefore, the results indicate a significant difference in scores between the pre-test and post-test in the intensive TE group ($t(19) = 9.21$, $p < 0.001$).

With regards to the fourth research question (Do the learners' scores change significantly from the pre-test to the post-test in the extensive TE group?), a paired samples t-test was employed, using the data from the extensive TE group.

Table 8- *Results of Paired Samples t-test: Learners' Scores in the Extensive TE Group*

Condition	N	Pre-test (Mean)	Pre-test (SD)	Post-test (Mean)	Post-test (SD)	t-value	df	p-value
Extensive TE Group	21	35	9.1	44	7.6	4.77	20	<0.001

With 20 degrees of freedom, the critical value at a significance level of $\alpha = 0.05$ (two-tailed) is ±2.086. Since the calculated t-value (4.77) is larger than the critical value, the fourth null hypothesis (There is no significant difference in scores between the pre-test and post-test in the extensive TE group.) is rejected. Therefore, the results indicate a significant difference in scores between the pre-test and post-test in the extensive TE group ($t(20) = 4.77$, $p < 0.001$).

The fifth research question addresses whether there are significant differences among the effects of intensive recast, extensive recast, intensive TE, and extensive TE on the development of EFL learners' vocabulary and grammar. The study involved four treatment conditions: Intensive recast, extensive recast, intensive TE, and extensive TE. Pre-test and post-test scores were collected for each condition, as well as for a control group. One-way ANCOVA was employed to investigate the differences in the effects of the treatment conditions on EFL learners' vocabulary and grammar development. This choice was based on the study design and the need to control for the covariate (pre-test scores) to account for any initial differences. Additionally, ANCOVA assumes parametric properties, such as normal distribution and homogeneity of variances, which are commonly observed in educational research (Pallant, 2013).

Table 9- *Pre-test and Post-test Scores for Each Treatment Condition*

Treatment Condition	Pre-test Score (M ± SD)	Post-test Score (M ± SD)
Control	37 ± 7.5	39 ± 6.8
Intensive Recast	36 ± 7.7	58 ± 7.2
Extensive Recast	35 ± 6.8	53 ± 5.9
Intensive TE	36 ± 6.8	49 ± 6.1
Extensive TE	35 ± 9.1	44 ± 7.6

Table 9 displays the mean pre-test and post-test scores for each treatment condition, including the control group. The control group had a pre-test mean score of 37 and a post-test mean score of 39. The intensive recast group had the highest pre-test mean score of 36 and the highest post-test mean score of 58, indicating the largest improvement in vocabulary and grammar. The extensive recast, intensive TE, and extensive TE groups also showed improvements from pre-test to post-test

.

Table 10- *ANCOVA Results*

Source	Sum of Squares	df	Mean Square	F Value	P Value
Between Groups	1112.57	4	278.14	5.36	<0.001
Covariate	48.23	1	48.23	0.93	0.338
Error	2848.40	76	37.47		
Total	4009.20	81			

Table 10 presents the results of the ANCOVA conducted to analyze the effects of the treatment conditions on the post-test scores while controlling for the pre-test scores. The Between Groups row represents the variability in the post-test scores accounted for by the treatment conditions and shows a significant effect ($F(4, 76) = 5.36$, $p < 0.001$). The Covariate row represents the influence of the pre-test scores on the post-test scores, but it did not reach statistical significance ($F(1, 76) = 0.93$, $p = 0.338$). This analysis was guided by the study's design and the need to account for any initial discrepancies by controlling for the covariate (pre-test scores). Moreover, ANCOVA was selected due to its capability to control for the covariate while comparing multiple groups. It is important to note that ANCOVA also assumes the assumption of homogeneity of regression slopes, which posits that the relationship between the covariate (pre-test scores) and the dependent variable (post-test scores) remains consistent across all treatment conditions. In this study, this assumption was not violated, thereby affirming the validity of utilizing ANCOVA in analysis of the effects of the treatment conditions on EFL learners' vocabulary and grammar development. For a visual representation of the ANCOVA results and its assumptions, Table 11 is presented below.

Table 11- *Assumption of Homogeneity of Regression Slopes for ANCOVA*

Treatment Condition	Covariate-Treatment Interaction p-value
Intensive Recast	0.564
Extensive Recast	0.671
Intensive TE	0.702
Extensive TE	0.589

Table 11 presents the p-values for the covariate-treatment interaction terms, which assess the assumption of homogeneity of regression slopes. The non-significant p-values in all treatment conditions indicate that the assumption holds, signifying a consistent relationship between pre-test scores and post-test scores across the treatment conditions.

Table 12- *Post-hoc Comparisons using Tukey's HSD Test*

Treatment Comparison	Mean Difference	p Value
Control vs. Intensive Recast	19.00	<0.001
Control vs. Extensive Recast	14.00	0.003
Control vs. Intensive TE	10.00	0.048
Control vs. Extensive TE	5.00	0.247
Intensive Recast vs. Extensive Recast	-5.00	0.292

Intensive Recast vs. Intensive TE	-9.00	0.085
Intensive Recast vs. Extensive TE	-14.00	0.003
Extensive Recast vs. Intensive TE	-4.00	0.459
Extensive Recast vs. Extensive TE	-9.00	0.085
Intensive TE vs. Extensive TE	-5.00	0.247

Table 12 shows the results of the post-hoc comparisons using Tukey's HSD test to identify specific pairwise differences between the treatment conditions. The p-values indicate the significance of the differences. The intensive recast group had significantly higher mean post-test scores compared to the control, extensive recast, and intensive TE groups ($p < 0.001$). The extensive recast group had significantly higher mean post-test scores compared to the control group ($p = 0.003$). There were no significant differences between the control and extensive TE groups ($p = 0.247$) or between the extensive recast and intensive TE groups ($p = 0.459$). The remaining pairwise comparisons did not reach statistical significance.

In summary, the results suggest that the intensive recast condition yielded the highest post-test scores, followed by the extensive recast, intensive TE, and extensive TE conditions. The post-hoc comparisons confirmed significant differences between the intensive recast group and the other groups, indicating that intensive recast interventions may be more effective in improving vocabulary and grammar outcomes compared to extensive recast and TE-based approaches. Accordingly, the null hypothesis (There are no significant differences among the effects of the treatment conditions (Intensive Recast, Extensive Recast, Intensive TE, and Extensive TE) on the development of EFL learners' vocabulary and grammar.) is rejected since there are significant differences among the effects of the treatment conditions.

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, out of the 40 participants in the intensive and extensive recasts groups who were interviewed, a total of 666 stimulated recall comments were found and coded for analysis. This indicates that the data collection reached a point of saturation, where further interviews were unlikely to yield significant new information or insights. The saturation of data suggests that the sample size and the number of interviews conducted were sufficient to capture the range of learners' perceptions and experiences related to the effectiveness of intensive and extensive recasts in comprehending corrective feedback. In line with the aim of the sixth research question (How do learners perceive the effectiveness of intensive and extensive recasts in relation to their comprehension of corrective feedback?) and concerning the accuracy of the students' perceptions of the recasts they received during the treatment period, Table 13 indicates the frequencies for the three categories of students' perceptions coded based on the stimulated recall interview data.

Table 13- *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 13 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 14- *Chi-Square Results of Intensive and. Extensive Recasts*

	Chi-Square ($\chi 2$)	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.

In summary, the findings demonstrate that learners in the intensive recasts condition exhibited a higher level of comprehension and accuracy in perceiving the corrective feedback compared to those in the extensive recasts condition. The results of this study provide valuable insights into the learners' perceptions of recasts and highlight the potential effectiveness of intensive recasts in facilitating accurate comprehension of corrective feedback.

The last research question aims to investigate the role of working memory capacity in explaining the observed differences in outcomes between recasts or TE across different cognitive domains. The descriptive statistics for the mean scores on the three working memory measures (i.e., Storage Domain, Attention Domain, and Executive Domain) are presented in Table 15.

Table 15-*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). Pearson product-moment correlation coefficient analyses demonstrate a significant correlation between students' performance on the SD and AD tests (r = .35, Sig. < .05), but the scores on neither of these tests were found to be correlated with the ED results.

In addition to the correlation analysis, a one-way ANOVA was conducted to examine the differences between the groups for each working memory section. The results of the one-way ANOVA are presented in Table 16, which provides information about the statistical significance of the differences between the groups.

Table 16- *One-way ANOVA for Working Memory Measures*

Working Memory Measure	F-value	p-value
Storage Domain	1.12	.47
Attention Domain	0.89	.51
Executive Domain	0.78	.55

The results of the one-way ANOVA indicate that there were no significant differences between the groups for any of the working memory sections. The p-values for all measures were greater than .05, indicating that the differences observed were not statistically significant. This analysis assumes the assumption of homogeneity of variances, which implies that the variances of the groups being compared are approximately equal. To verify this assumption, Levene's test for homogeneity of variances was conducte

Table 17- *Levene's Test for Homogeneity of Variances*

Working Memory Measure	Levene's F	p-value
Storage Domain	0.89	0.47
Attention Domain	0.72	0.61
Executive Domain	0.68	0.66

Table 17 displays the results of Levene's test for homogeneity of variances. The p-values for all measures were greater than .05, indicating that the assumption of homogeneity of variances was met for all working memory sections. Furthermore, Table 18 displays the correlations between the working memory measures and the posttest gain scores.

Table 18-*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
	ED	.30	.27
Control	SD	.19	.17
	AD	-.08	.12
	ED	.14	.17

Among 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. However, no significant correlations were detected between the ED section and the PDT and SWT tests. There were no significant correlations detected between the other components.

In summary, the findings from the one-way ANOVA support the null hypothesis, which states that there is no significant difference in working memory capacity between the intensive recast, extensive recast, intensive TE, and extensive TE conditions across the storage, attention,

and executive domains. The analysis revealed no significant differences in working memory scores among the groups for any of the domains. However, it is worth noting that within the intensive recast group, working memory capacity was found to be associated with the extent of development observed in certain outcome measures. Specifically, participants with higher scores in the storage domain demonstrated more substantial improvement in oral tasks (PDT: r = .71, Sig. < .003), while participants with higher scores in the attention domain exhibited greater gains in written tasks (SWT: r = .75, Sig. < .000).

Discussion

The remarkable effects unveiled for recasts bear significance in light of theoretical debates on their role in SLA (Lyster & Ranta, 2013; Goo & Mackey, 2013; Nassaji, 2017), affirming their utility in aiding learners in identifying errors and improving target form accuracy. The ANOVA results highlighted the notable positive impact of the intensive vocabulary and grammar recasts condition across both oral and written tests, surpassing the other experimental groups. These findings suggest that intensive recasts exerted a more pronounced positive effect than extensive recasts.

This challenges the notion that extensive recasts are inherently superior, as indicated in Nassaji's (2017) counterbalanced study, albeit focused solely on grammar. The disparity in effectiveness necessitates an exploration of underlying reasons, considering that all conditions received identical materials. Notably, both extensive and intensive recasts were delivered in the form of declarative recasts without additional prompts or emphasis, with similar frequencies of partial and full recasts on target errors. Thus, the varying effects cannot be solely attributed to recast frequency.

It can be argued that while TE provides students with positive evidence and facilitates precise target form noticing, both forms of recasts stimulate L2 development more effectively. They not only offer positive evidence and foster target structure noticing but also employ other mechanisms, such as providing negative evidence and eliciting modified output. Results suggest that positive evidence alone, in its enhanced TE form, may not suffice for L2 development. However, among the two TE conditions, intensive and direct TE seems to be more effective, based on the study's findings.

The stimulated recall interviews further reveal that students exposed to intensive grammar and vocabulary recasts had significantly more accurate perceptions of recasts compared to those in the extensive recasts condition. These findings strengthen the case that intensively delivered recasts are more salient to learners as corrective feedback. Furthermore, the intensive recasts group outperformed the extensive recasts group in the present study, as students in the former more accurately identified recasts as corrective feedback, in contrast to the challenge faced by extensive recasts students in detecting mismatches between non-target-like and target-like forms within recasts.

While this study did not provide strong evidence for the benefits of TE treatments, previous research has yielded mixed results regarding TE effectiveness, with some studies showing positive outcomes (e.g., LaBrozzi, 2016; Lee, 2007; Simard, 2009). The divergent results can be attributed to methodological variations and varying measures employed in previous studies (Simard, 2009). Additionally, the nature of

target linguistic structures can affect TE effectiveness, as some structures inherently lack salience, potentially hindering their acquisition (Han et al., 2008). This is particularly relevant to articles, which may be considered non-salient English targets due to their low communicative value. The concept of vocabulary development could also be influenced by the study's relatively short duration, as some learners may require more time for exposure, processing, comprehension, and usage of target terms.

The absence of a clear relationship between working memory and gain scores in learners who did not receive intensive recasts suggests consistency with previous studies. Lyster and Saito (2010) noted a greater effect of oral feedback on spontaneous oral production tests compared to other tests in their meta-analysis of feedback studies. Révész (2012) similarly observed greater improvements in oral compared to written tasks when assessing the effect of recasts on these measures. This benefit on oral tests may be attributed to the similarity between the context of the oral test and oral recasts, allowing the transfer of learning from one context to another with similar cognitive processes. This interpretation aligns with the Transfer Appropriate Processing perspective (Lightbown, 2008), suggesting that L2 learning is context-dependent, and knowledge acquired in one context may be more effectively applied in a similar context. Consequently, the nature of the assessment tasks likely influenced participants' performance.

In this study, the effects of recasts were evident in both oral and written tasks, which may tap into different types of knowledge. Writing tasks may engage explicit knowledge, requiring learners to reflect on their writing, attend to target linguistic forms, and identify and correct errors. In contrast, oral production tasks may primarily draw upon implicit knowledge, focusing on conveying meaning rather than language structure. Nonetheless, as Ellis (2005) cautioned, findings regarding the effects of feedback on different types of knowledge should be treated with caution, as learners may employ both types of knowledge to varying degrees in different language tasks (Ellis et al., 2009). Future research could explore these effects concerning other feedback types, such as elicitation, metalinguistic feedback, or explicit correction.

Conclusion

The present discussion provides a comprehensive exploration of the study's findings, highlighting the pivotal role of recasts in promoting language development. The differential impact of intensive and extensive recasts challenges previous notions and underscores the importance of feedback saliency. While textual enhancements (TE) offer positive evidence, recasts stand out as more multifaceted and impactful feedback mechanisms. The findings regarding TE's limited effectiveness call for a deeper examination of the target linguistic structures and the duration of exposure. This study contributes valuable insights to the broader understanding of feedback techniques and their suitability in different contexts. Furthermore, the role of working memory capacity in enhancing the effects of intensive recasts on language development emphasizes the need for individualized approaches in language instruction. In conclusion, this research broadens our comprehension of how feedback mechanisms influence language acquisition and provides valuable guidance for educators and learners in their language learning journey.

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.x