The Effect of Employing Task Complexity with Recasts on Improving Grammar Acquisition among EFL Egyptian Learners

Dr. Elham Sweilam Ahmad Desouky*

Abstract

The main purpose of the current study was to explore the effects of utilizing task complexity when associated with recasts on improving grammar acquisition among EFL Egyptian learners, as well as its effect on relation between aptitude and EFL learners' grammar acquisition after recasts. To achieve the purpose of the study, 120 elementary FL learners were divided in two experimental groups: one of them finished complex decision-making assignments, while the other group accomplished simple information transmission assignments. Recasts were provided as a reaction to the participants' errors when they used the present third person singular verb forms. To collect the data, the researcher used these tools; elicited imitation tests, oral production, and written production to measure EFL grammar acquisition, in addition to LLAMA D, LLAMA E, and LLAMA F to evaluate learners' aptitude. The results showed that it was more useful to employ tasks with less difficult cognitive demands. In complex task conditions, the grammar acquisition was predicted through the outcomes of the participants in LLAMA D, which measured both oral and written production. Added to this, grammar acquisition was also predicted by the learners' performance on LLAMA E, which was measured over elicited imitation.

Key words: Elementary FL Egyptian learners; Task Complexity with Recasts; English grammar

Introduction

Longe (1996, 2015) inspired other researchers in the field of instructed EFL acquisition to concentrate on form approach to teaching a foreign language. As a result of Long's interaction hypothesis, this approach postulates that facilitating subsequent FL development results from the ability to draw students' attention to linguistic components when involved in meaningful interaction. At the same time, Ellis, Loewen, and Erlam (2006) thought that giving corrective feedback to students is a way to stimulate a focus on form. Some meta-analysis studies (Li, 2010; Lyster & Saito, 2010) agreed with this view by showing that corrective feedback can assist interlanguage development. These views raised the debate among researchers concerning the most useful types of corrective feedback and under what conditions can it help students learn. One of these disputed types was recasts, which has been the target of an excessive amount of studies. Recasts

* Lecturer of Curriculum & Instruction (TEFL) - Faculty of Education, October 6 University
Email: elesweilam@gmail.com

can be defined as a way that the student's utterance can be reformed in one or more altered versions without changing the main ideas in the text. The current study displays how participants' utterances can be reformulated via utilizing recasts so as to modify an error in the present third person singular verb form as shown in example (1).

Example (1)

Student: She eat mango
Teacher: She eats?

Loewen and Sato (2018) claimed that there are different variables affecting recasts which lead to interlanguage development; one of these variables is task complexity. The current study tried to investigate the effect of task complexity concerning ingrained cognitive requests, as well as the individual differences in FL aptitude. The results of these two variables were different in many previous studies. In fact, some previous studies (e.g. Révész, 2009; Baralt, 2013; Révész, Sachs, & Hama, 2014; Kim, Payant, & Pearson, 2015) asserted that task complexity impacts the efficacy of recasts, but at the same time they were not sure whether recasts was more useful for the participants during cognitively complex assignments or during simple ones. Concerning the relation between the effects of recasts and aptitude, the results were diverse to some extent, as some of the studies (Sheen, 2007; Yilmaz & Grañena, 2016) proved that there was no relation between learning through recasts and aptitude, while some other studies (Trofimovich, Ammar, & Gatbonton, 2007; Li, 2013; Yilmaz, 2013) found a link between them.

Suzuki and DeKeyser (2017a) stated that based on indications in the literature, aptitude interacts with different learning conditions. Thus, the reason behind these contradicted findings concerning task complexity may be that these studies were unsuccessful in controlling the possible differences in student aptitude. Robinson (2011) suggested that when dealing with aptitude and task complexity, the cognitive complexity of the task during which learners obtain feedback might temper the degree to which aptitude predicts FL development.

Study Significance

The importance of the current study lies in the fact that it provided new insights into how to adopt recasts, accompanied with task complexity, to enhance grammar acquisition among EFL Egyptian learners. Another contribution is that it helped to improve EFL learners' acquisition of grammar**.**

Study Problem

The Egyptian TEFL elementary learners were lacking basic grammar skills. They were not accurate in using the present third person singular (e.g., give-s, eat-s), which was noticed during practicum days. The current study tried to assist the learners in solving this problem.

Accordingly, the current study sought to address the following questions:

1- Are there any effects of task complexity on developing EFL learners' use of the present third person singular, and what are these effects?

2- Does aptitude predict development of EFL learners' use of the present third person singular? To what extent?

3- Does task complexity influence the relation between EFL aptitude and development in EFL learners' use of the present third person singular? To what extent?

Review of Literature

This part is divided into three sections: Section One discusses the studies on recasts and task complexity, Section Two deals with the studies of aptitude patterns and measures, and Section Three looks into the extent to which EFL grammar acquisition benefits from recasts in terms of task complexity and aptitude.

SECTION (A)

Recasts and Task Complexity

In 2009 and 2014, Skehan presented the limited attentional capacity model, after 10 years; Robinson (2001, 2011) formulated the cognition hypothesis model. Both of them were considered the two cognitive models that were used as a base for previous studies on corrective feedback and task complexity. In Robinson's model (2001, 2011), task complexity was defined as the cognitive demands enjoined on FL students through the ingrained characteristics of the assignment in which they take part. According to Robinson, there are passive ways to affect language production and focus on form when the cognitive demands of assignments are increased along resource-dispersing elements (such as planning time). So the students' memories and attentional resources will be scattered. On the other hand, as Robinson predicted when task complexity is increased along resource-directing dimensions such as reasoning demands, the students' attentional and memory resources will hence be addressed to the functional and linguistic demands of the assignment, leading to positive impacts on production and incorporation of information offered by focus on form interventions (e.g., recasts).

As well as, Levelt (1989) offered his pattern of speech production in which he showed the four phases of speech production that begin with conceptualization. The second step was formulation, which refers to encoding the message grammatically, lexically and phonologically. The third step involved production of speech sounds accompanied by articulation. The last step was to know whether the content produced is suitable and accurate through self-observing associated with evaluation. Skehan (2009-2014) utilized Levelt's model as the basis to present his limited attentional capacity pattern to show the impacts of corrective feedback under both complex and simple conditions. Based on Levelt's model (1989), Skehan (2009, 2014) concluded that considerable cognitive complexity such as greater reasoning demands will make linguistic encoding less important than conceptualizing. This is because tasks with greater

cognitive demands complexify conceptualizer processes, resulting in fewer attentional resources available for linguistic encoding, and thereby leading to the production of less complex language and/or lower accuracy. On the other hand, Révész et al. (2014) used Skehan's model to show that reduced pressure on the conceptualizer may have better effect on helping recasts draw students' attention to linguistic encoding, as presenting recasts may be more beneficial when students perform tasks with lower cognitive requests.

A few studies (Révész, 2009; Baralt, 2013; Révész et al., 2014; Kim et al., 2015; Kourtali & Révész, 2019) investigated the impacts of using recasts and task complexity together. Révész's (2009) for example, tried to apply both recasts and task complexity using the past progressive form to develop FL. In Révész's study, the participants were divided into two groups: one of them obtained recasts while involved in a task in which they were asked to describe a photo from memory without contextual support. The other group had given recasts during the same photo description task, but this time with giving contextual support, as the photo was within reach during the performance of the task. The results of Révész's study showed that giving recasts without contextual support was more useful for the learners. Nonetheless, the results of this study gave clear evidence that the absence of contextual support facilitated the efficacy of feedback, but concerning the role of task complexity, this study did not offer clear conclusions. Some other studies (Robinson, 2003; Skehan, 2014) had different views regarding whether the presence or absence of contextual support creates higher cognitive demands, and this point is still a matter of discussion (Kourtali & Révész, 2019).

Nearly during the second half of this decade, Révész, Sachs, and Hama (2014) tried to employ various task manipulations to investigate to what extent expanding the reasoning demands of duties may have an impact on the effectiveness of recasts in evolving students' knowledge of the past counterfactual construction in a context mediated by a computer. The role of the learners was to use a picture story they had read before to point out the relation between causes and results of events, which were designed to be clearer in the simple condition than in the complex one. The results showed that recasts provided during simple tasks were notably more useful than those presented in complex assignments. The researchers used both Skehan's (2014) limited capacity model and Levelt's (1989) speech production model as a basis to prove their point of view concerning the participants' performance during the complex task conditions, as learners gave less attention to both linguistic targets and recasts, whereas they may assign more attentional resources to task achievement.

Another study conducted by Baralt (2013) used face-to-face and computer-mediated environments as a new method to investigate the impacts of recasts and task complexity on improving FL learners. The Spanish past subjective was the linguistic target of this study, and task complexity was utilized to present the international reasoning requests posed in storytelling assignments. The role of the learners in the complex case was to consider the intentions of the characters during the stories' retelling, but the situation was different in the simple case, as the stories already contained the characters' intentions.

Results revealed that recasts was more useful in the face-to-face method after had given complex assignments with larger cognitive requests, while in using computer-mediated method, recasts was more effective in simple assignments asking for less thinking. These contradictory results were due to many reasons like the number of turns, contingency of recasts throughout the two methods, and differences in discourse length, as interpreted by Baralt.

It can be concluded that the findings of previous studies investigating the combined impacts of corrective feedback and task complexity on FL development which have been reviewed so far indecisive. Also, depending on some previous studies on aptitude–treatment interaction like Suzuki and DeKeyser (2017a), both Kim et al. (2015) and Kourtali and Révész (2019) tried to present one possible interpretation of the inconclusive results through showing that current studies did not control for individual differences in cognitive abilities like FL aptitude and working memory capacity. Therefore, the target of the current study is to reconsider this possibility by investigating how aptitude might moderate the link between the effect of recasts and task complexity.

SECTION (B)

Patterns and Measures of Aptitude

Grañena (2013) stated that language aptitude refers to cognitive and perceptual capabilities that simplify the acquisition of foreign language. One of the pioneers of FL aptitude research was Carroll (1965, 1980), who conceptualized aptitude covering four elements: (1) inductive language-learning capability; (2) grammatical sensitivity; (3) associative memory or rote learning ability; and (4) phonetic coding ability. The four elements were successfully illustrated by Carroll as follows: the role of phonetic coding is to make connections between sounds and their symbols and then hold them over, which means that phonetic coding requires identification of sounds. The second element was inductive language learning, which indicates the ability to derive rules from content. The third element was grammatical sensitivity, which helps students to pinpoint the functions of the words in sentences. The last element was associative memory, which refers to the ability of recognizing connections between sounds and meanings as well as the ability to hold them over. In 1959, Carroll and Sapon designed the Modern Language Aptitude test (MLAT), which was not successful in measuring inductive language-learning ability; however, it measures all subconstructs of aptitude presented by Carroll. Therefore, this test is considered one of the most effective language aptitude trials.

Skehan (1998) tried to build partly on Carroll's (1959) work from an adopted information-processing perspective to pattern aptitude and described FL aptitude as a construct comprising cognitive differences in memory-as-retrieval, phonetic coding, and language analytic ability, with language analytic ability subsuming Carroll's grammatical sensitivity and inductive language awareness. Skehan (2002) added working memory and attentional control to the subconstructs of aptitude as an update to the pattern; at the same time, Skehan (2002, 2016) further suggested that the various aptitude subconstructs are

engaged at different steps of FL acquisition. It is clear that in this pattern, various elements of aptitude can be linked to the cognitive processes engaged in learning grammar via giving feedback and when the feedback is segmented, so working memory and attentional control are likely to be connected in this phase. Participants may be able to realize highlighted errors and the corrective function of feedback through the help presented by working memory and phonetic coding ability. Language analytic ability, phonetic coding ability, and working memory may be able to facilitate recognizing models via exposure to feedback. In addition, students will be able to avoid errors by the expected help presented via working memory and retrieval memory (Kourtali & Révész 2019).

Over the last two decades, various new aptitude tests have been created with different purposes and targeted participants, but at the same time the MLAT Test is still widely used. One of these tests was the CANAL-F test, which was designed by Grigorenko, Sternberg, and Ehrman (2000) to evaluate participants' potential to deal with novel FL learning situations. Another test, called the Hi-LAB battery, aims to pinpoint cognitive abilities that enhance the accomplishment of advanced FL skills, designed by Linck et al., (2013). The Hi-LAB battery test incorporates auditory differentiation, measures of the central executive elements of working memory, phonological short-term memory, associative memory, long-term memory retrieval, processing speed, and implicit learning. In spite of the reservations of Bokander and Bylund (2019) about the validity of LLAMA as a test of aptitude, the current study decided to employ the LLAMA test as a measure of aptitude (Meara, 2005; Kourtali & Révész, 2019). The present study considered the LLAMA test as a tool used to test aptitude in current FL acquisition study and is composed of four subtests with different goals. These four subtests utilized to measure inductive language learning ability (LLAMA F), and after that, the capability to differentiate patterns in spoken language (LLAMA D), then measuring rote, associative memory (LLAMA B), and finally the ability to connect sounds with symbols (LLAMA E). The participants in the present research were asked to recognize sound sequences (LLAMA D), phonetic coding (LLAMA E), draw grammatical inferences (LLAMA F), and process oral feedback targeting a grammatical feature, all of which were deemed pertinent to the development of the participants' learning.

SECTION (C)

Task Complexity, Recasts, Aptitude, and EFL Grammar Acquisition

Some previous studies of aptitude and corrective feedback (e.g., Yilmaz, 2013; Yilmaz & Grañena, 2016; Yilmaz & Koylu, 2016) indicated that the provision of explicit feedback was more useful for the participants with higher aptitudes, and these results were in line with the predictions obtained from Skehan's pattern (2014). At the same time, the current research produced varied results regarding the relation between recasts and aptitude. The results of some studies (Sheen, 2007; Yilmaz & Grañena, 2016) showed that the validation of recasts was not related to participants' aptitudes, while some

other studies (Trofimovich et al., 2007; Li, 2013) revealed that the extent to which recasts leads to EFL grammar acquisition gains is linked to elements of aptitude such as attention control and language analytic ability. One reason behind these paradoxical results relates to the nature of the linguistic feature targeted by recasts. The study of Yilmaz (2013) proved this point of view. As Yilmaz discovered, language analytic ability did not help EFL grammar acquisition learning when recasts targeted a less salient construction. But when the focus was a salient feature, the higher language analytic ability simplified learning by recasts. Sato and Loewen (2018) argued that, in spite of the fact that previous studies examined the impacts of corrective feedback on learning linguistic constructions that vary in salience, there is a deficiency in searching behind the role of aptitude in simplifying the advantages of corrective feedback for various types of linguistic features.

On the other hand, some studies considered that cognitive task complexity could be added to the nature of the linguistic target as another factor that may moderate the interaction between aptitudes and recasts. One of these studies was Robinson's (2011), in which he postulated two assumptions: firstly, the learner elements interact with task complexity in order to locate the benefit from pedagogical interventions that the learner will gain. Secondly, he assumed that in terms of cognitive abilities, individual variance would be increasingly linked with learning as tasks increase in cognitive complexity. This hypothesis would lead to lower differences in gains when participants finish assignments with lower cognitive demands than when they implement complex assignments. Another experimental study by Kim et al. (2015) was in line with the former prediction, as it showed that the learners who benefited most from receiving recasts had high working memories and carried out assignments with greater cognitive demands. In spite of these results, the experimental question remains as to whether task complexity influences the relation between the impacts of recasts in developing EFL grammar acquisition use and other aptitude factors like language analytic ability and phonetic coding.

Study Purposes

The purpose of the current study was to investigate three issues: firstly, was there any effect of task complexity on Egyptian EFL students' when learning the present third person singular (e.g., give-s, eat-s).The second aim was to find out whether the ability of the learners can facilitate EFL grammar acquisition learning by showing the relationship between learners' aptitude and their learning. The third aim was searching for whether task complexity influences the relationship between aptitude and learning after providing corrective feedback to the learners. Feedback was presented in this study in the form of recasts which correctly reformulate the students' errors. Aptitude found to be related to learners' EFL grammar acquisition gains under complex task conditions only, while the participants were benefited from recasts to a greater extent when they were under simple task conditions in comparison to complex ones.

Method

Design

A pretest–posttest design with two experimental groups was utilized in the current study, and the learners were divided into two empirical groups via stratified random sampling, considering participants' length of time studying English, their pretest, grammar acquisition, and aptitude test results. One of the two groups carried out two complex decision-making tasks, which provoked reasoning demands on participants, whilst the other group worked on two simple information transmission tasks with no reasoning demands during the treatment sessions. It was clear (see Example 1) that both groups obtained interrogative recasts in response to errors linked to the target feature, like what was mentioned by Norris (2010) and Révész (2014), every treatment task was followed by a posttask questionnaire, which was included to evaluate the perceived cognitive demands provoked by the treatment tasks and thereby indicate the validity of the task complexity manipulation. As with Meara (2005), the current study measured EFL aptitude by the LLAMA test; at the same time, EFL grammar acquisition was assessed by an elicited imitation (EI) test, an oral production test, and a written production test.

Participants

The learners who participated in this study were 68 males and 52 females. Their ages ranged from 10 to 14 years (M=22.92, SD=1.64). All of them were from the countryside Giza governorate in Othman Ibn Affan governmental school. All of the learners were Arabic speakers who had studied formal English for at least 5 years (M= 9.3, SD=2.34). To face the variables of length of previous English study, performance on the proficiency test, and age, the researcher conducted a group of independent-samples t tests to assure that the two sets were comparable: English study, t = 0.66, p = .75, d = 0.16; grammar acquisition, t = 0.14, p = .95, d = 0.02; age, t = 1.56, p = .44, d = 0.28. The participants' level were A2 (elementary); they did not show inclusive prior knowledge of the target feature on any of the pretests, and they had never lived in an English-speaking country before the research. The researcher followed the same track as Kourtali and Révész (2019) when using the listening element of the Trinity College ISE Foundation test, which targets level A2 in the Common European Framework of Reference (CEFR) to evaluate learners' grammar acquisition. The participants needed to attain an ISE score in the range of 2 to 4 to be eligible for the current study. Additionally, the learners' prior knowledge of the target feature was evaluated via pretests, and to bypass ceiling impacts, the participants who reached higher than 35% on any of the result measures were eliminated from the study. The distribution of scores (the next highest score was 90% on the written production test and 50% on the oral production test) was the baseline to determine this threshold. The idea behind excluding learners who obtained high scores on the pretests was that comprehensive prior knowledge could conceal the role of aptitude and task complexity in EFL grammar acquisition study. In addition, if the study were to

incorporate participants acquiring higher than 90% on one of the pretests, then the study would have faced a ceiling impact on improvement.

Instruments

To collect data, the learners completed a background questionnaire in their first language (Arabic). Then, they worked on the pretests and the grammar acquisition test. The oral production pretest was the first test the learners' started with. Thereafter, they continued with the Elicited Imitation pretest. Finally, the written production pretest was done. The target during the oral tests was to raise the chance of the learners depending on procedural knowledge. The tests of oral and written Production were prepared to evaluate learners' capacity to produce EFL grammar acquisition in oral and written forms. Each test included 12 pictures to measure learners' ability to speak about the habits of imaginary figures and how they were dissimilar from those what appeared in the tasks of the treatment. Additionally, elicited imitation test was used in which the participants were subjected to listen to a group of stimulus items (or phrases, words, sounds) and then they were asked to repeat the sentences verbatim. The elicited imitation test consisted of 72 sentences; 48 of them worked as distractors, while the other 24 targeted the present third-person singular.

Validity and Reliability

The current study tested 120 Egyptian elementary learners of English as a foreign language, their ages ranged from 10 to 14 years (M=22.92, SD=1.64). The Participants started by completing a grammar acquisition test firstly. Then, they carried out three tests that assessed the use of agreement on the verbs which express a third person singular in the present tense (e.g., give–s, eat–s). All the tests consisted of an oral production, written production, and elicited imitation. Every treatment task was followed by a post-task questionnaire, which was included to evaluate the perceived cognitive demands provoked by the treatment tasks and thereby indicated the validity of the task complexity manipulation. In addition, the current study utilized a computer-administered aptitude test called LLAMA that is known as a test of language independent that employs an unfamiliar language and visual stimuli. The forms of the tests were submitted to four TEFL professors and three senior English teachers to evaluate the tests in terms of appropriateness of the tasks to the learners' level, their age, and the Egyptian environment. All the involved professors and teachers confirmed the suitability of the tests to the learners; however, they highlighted several points to be taken into consideration. Firstly, learners should be provided with a detailed explanation of the tests and their requirements before starting the tests. Secondly, it is important to use simple words while demonstrating the tests instructions. Finally, the elements from the aptitude test should be administered and delivered in the learners' native language (Arabic). In order to establish the reliability of the tests included in the current study, it was administered to a randomly selected group of thirty students. Those students had been excluded from the whole experiment. Two weeks later, the tests were administered to the same group once more. The current study used test\ re-test results method. Additionally,

the Person correlation coefficient between the test \re- test results was calculated. The reliability was found to be high in tests of oral and written production, elicited imitation, and aptitude test. Therefore, the tests conducted in the current study were considered reliable.

The Target Structure

The feature of the present third-person singular verb form was regarded as a difficult structure to obtain. So it was chosen as the linguistic target of the current study. There are many reasons behind the difficulty of the form of the present third-person singular verbs; one of them is that it comprises a bound morpheme, which is understood via three allomorphs [əz], [z], and [s]; therefore, the grammatical function maps onto multiple forms, making it hard for learners to discern the form–function relation via exposure to oral information. The second reason is that none of the three allomorphs are stressed; thus, the structure lacks physical salience and the learners may neglect the form. In 1996, Van Patten mentioned the third reason, as the present third-person singular is not a communicatively required characteristic, as its meaning can be relocated successfully via subject noun phrases that supply the use of verbal inflection in a sentence not required. In addition, Han (2013) added that the present third-person singular is prone to fossilization in adulthood, which could be because of these characteristics. Based on the former reasons, it is important to determine which learning conditions may enhance its accurate use among EFL learners; for example, recasts conveyed during complex assignments or simple ones; additionally, it is theoretically important whether different EFL aptitude constructs may be related to the learning of a nonsalient and therefore not required characteristic.

Treatment Tasks

The learners were asked to discuss habits of fictional characters in both the complex (decision-making) and simple (information transmission) task conditions. The role of the learners in the first decision-making task was to act as the director of a building, while the role of the researcher (first author) was to help them. Depending on the task instructions, the participants were able to find many elements left by a transportation company at the entrance of their residence, and the tenants of the apartment block were the owners of these elements. In the next step, the participants (i.e., administrators) were able to use information they had about the tenants' habits in order to decide which elements belonged to whom (e.g., "The saw is Yehia's because he works as a carpenter"). In order to help the researcher to return the tools to the tenants, the learners had to tell them about these decisions. At the same time, the learners did two missions, one of which was taking responsibility for helping the administrator (i.e., researcher) in the corresponding information transmission task, and the other mission was to provide data about what the tenants typically did on weekends so the administrator could make decisions about who were the owners of the tools. It can be explained through the example of the learners when they told the researcher that Yehia works as a carpenter, and this data helped the researcher to understand that the saw was Yehia's. Each learner

in the second decision-making mission was required to visualize that they worked at a hospital and they needed to utilize a table showing the possessors' items and habits, and then explain their decisions. The participants in the task of information transmission were asked to play the role of hospital employee to assist the researcher to become aware of the owners of the lost elements; learners were then asked to use a table with pictures to present data about the habits of the possessors. All instructions for all tasks were given to learners in their first language (Arabic).

The decision-making tasks were considered more cognitively complicated because the learners were asked to be logical. In addition, the studies of Robinson (2001, 2011), Baralt (2013), Kim et al. (2015), and Révész et al. (2014) mentioned that in the transmission tasks there was no thinking needed to fulfill the data. The researcher tested the tasks with a sample of children similar to those who participated in the experiment and showed that the age was suitable; and the tasks were successful in fulfilling the structure's purpose.

Type of Recasts

When the learners were engaged with the treatment tasks, they produced the present third-person singular verb form inexactly. So, they received recasts, and the following example showed this:

Example (2)

Student: Habiba on Fridays paints in the public garden from 1 to 2 p.m. She is at home…….

Teacher: She paints?

The present study followed Lyster's (1998) classifications in presenting recasts separately and interrogatively, as no extra data were given to the participants, just reformulating the incorrect parts of the learners' statements. Additionally, the uses of recasts had been reduced to consist of a verb in the present third-person singular and a personal pronoun. The participants concentrated only on one change, either a substitution when the researcher exchanged a nontarget like statement (using another tense) with the present third-person singular or the addition of the allomorph [s], [z], [ɪz], or [əz], which made the way of presenting the recasts focused and intensive. According to previous studies (Bardovi-Harlig, 1987; Sheen, 2006; Kourtali & Révész, 2019), in addition to the former characteristics and the utterance-final position of the correction, recasts in the current study can be classified as explicit. In this study explicit recasts were operationalized as recasts which were stressed, partial, and with only one change from the erroneous utterance. In doing so, the corrective force of recasts was quite obvious to the participants whose performance was very low at the beginning of the experiment. Therefore, it was easy for them to attend to the correction of their erroneous utterances and at the same time to make cognitive comparison between their erroneous utterance and the researcher's corrective reformulation. The previous studies (Nassaji, 2009;

Yilmaz, 2012) showed that implicit recasts was less effective in enhancing noticing and learning, especially when targeting nonsalient constructions, and that was the main reason for using explicit recasts in the current study.

Posttask Questionnaire

The current study had adapted the posttest questionnaire from Robinson (2001), which was administered to the learners in both groups after each treatment task in order to collect data The learners were asked to give ratings on a 9-point semantic differential scale about 1-the quality of their own task performance, 2-the linguistic demands posed by the task, 3-the amount of mental effort required by the task and 4- the overall task difficulty. According to the nature of the task manipulation, it was important to distinguish mental effort from the difficulty of the task. The researcher presumed that decision-making tasks would demand greater mental effort compared to just data transmission because it was unlikely that the learners would have difficulty in perceiving the decision-making tasks. The results were subjected to further statistical analyses. The elements were represented in the participants' L1 to be easily understood, and through collection of data, the researcher provided clarification to concerned participants when appropriate.

Assessment Tasks

Tools

To evaluate learners' capacity to produce EFL grammar acquisition orally and in writing, the current study used tests of oral and written Production. Additionally, elicited imitation test was used in which the participants were subjected to listen to a group of stimulus items (or phrases, words, sounds) and then they were asked to repeat the sentences verbatim. These tests are clarified as follows:

Tests of Oral and Written Production

In order to evaluate learners' capacity to produce EFL grammar acquisition orally and in writing, the researcher designed the oral and the written production tests. Each test included 12 pictures to measure learners' ability to speak about the habits of imaginary figures and how they were dissimilar from those who appeared in the tasks of the treatment. In addition, two forms were designed and counter balanced them across the pretest and posttest to be used in both tests. All the pictures in the two forms represented the same activities to obtain the same behaviors. In both groups the same procedure was followed, as one form was applied as a pretest for half of the learners while the other half used the same form as a posttest. Each test was formulated to extract actions with three allomorphs, as well as the number of pictures being the same for the two test forms to generate a similar number of mandatory contexts for the target feature. In both oral and written tests pictures demonstrated the same behaviors, and the researcher gave the learners the test instructions in Arabic (L1) to ensure that the actions created were the same in these two ways.

It was important to tap into grammar acquisition; thus, oral and written evaluations were included. The researcher was keen on making both tests incorporate procedural knowledge, as mentioned by (Anderson, 1993; DeKeyser, 2007), but the written production test naturally produced less time pressure than the oral production one. As a result of this, performance on the oral production test was considered a better indicator of procedural knowledge in the process of automatization than the written production test. The learners in the writing production test were able to re-examine their production and deploy their explicit declarative and procedural knowledge due to the absence of time pressure during the process. At the same time, the researcher was able to utilize more meticulous standards for the research after employing tests in different modes that helped to gain a complete view of the learners' improving knowledge of the target structure. In order to ensure that learners' prior knowledge did not confound the findings gained for aptitude and task complexity, the researcher excluded participants who acquired low scores on the oral test while they scored 100% on the writing production test in the premier pool.

Elicited Imitation Test

Microsoft PowerPoint was used to compose the elicited imitation test. The learners were asked to listen to a sentence carefully; then she gave them two pictures named 1 and 2 and learners had to select the picture that was linked to the meaning of the sentence that they had heard, then say loudly either 1 or 2. In order to avoid the production of the sentence engaging in rote repetition (Erlam, 2006), the researcher gave the learners 4-second intervals between the presentation of the stimulus and the extracted response. In the next step, the participants were given 10 seconds to produce the sentence in correct English after the color of the slide changed from blue to pink.

The elicited imitation test consisted of 72 sentences; 48 of them worked as distractors, while the other 24 targeted the present third-person singular. The elicited imitation test started with seven practice items that did not include the target structure; also, half of the sentences were grammatical and half were ungrammatical for both distractors and target items. Concerning the target items, the three allomorphs were doled out equally for each of the three allomorphs [s], [z], and [əz] containing three grammatical and three ungrammatical sentences. In dealing with ungrammatical target sentences, the base form of verb was utilized, and the items were pseudo-randomized so that the same allomorphs were not presented in sequence, as suggested by Keating and Jegerski (2014).

In order to combat item impacts, two versions for each item were designed, as mentioned by Keating and Jegerski (2014). It was clear that the two versions differed as to whether the target verb was given as grammatical or ungrammatical (e.g., "She always eat mango" vs. "She always eats mango"). As Keating & Jegerski (2014) suggested, learners did not face both versions of the same item to avoid recurrence. Therefore, half of the learners in both sets were administered Version (1) of the test, while the other half

took Version (2). In addition, the test versions and the order of the test elements were counterbalanced between the pretest and the posttest.

Current study took the view of Keating and Jegerski (2014) into consideration, as the critical verbs inflected with the present third-person singular were of the same length (i.e., one-syllable verbs) and existed in the same position in each target sentence to force analogous processing requirements on the learners. Referring to Vocabprofiler, previous studies (Heatley, Nation, & Coxhead, 2002; Cobb, 2016) considered the verbs were the 1,000 most recurrent words in the English language (Qa group), and from the second most repeated 1,000 words (Qb group). The main reason behind having words from the Qb group was that they repeatedly appeared in textbooks, as well as the experimental study showing that the words were recognized by Arabic EFL participants who have analogous background; for example, the word algebra. The structure prior to the present third-person singular in each target sentence was composed of a personal pronoun followed by a two-syllable adverb of frequency, for example "She at times reads books with her father". The adverb was included to establish a mandatory context for the present third-person singular, as well as the personal pronoun working as the subject of the critical verb. Both stimuli and distractors have the same length of the target; all 144 sentences consisted of six syllables, and the reliability of the two versions of the Elicited Imitation test was found to be high (Version 1 α = 1.684; Version 2 α = 1.734).

It was clear that both tests (Elicited Imitation and oral production) were given in the oral manner, but the main difference between the two tests was that the Elicited Imitation test engaged the processing of oral input and production of prearranged utterances, whilst the oral production test was less controlled, permitting entrants to give their own works. Some researchers considered the type of knowledge Elicited Imitation tests measure to be a subject of discussion concerning whether they capture implied knowledge (Erlam, 2006) or, as mentioned by (Suzuki & DeKeyser, 2015), automatized straightforward knowledge. At the same time, Suzuki's study (2017) claimed that this type takes a number of years in inundation contexts to utilize implied knowledge more dependably. Drawing on this point, the researcher found that it would be more likely that learner EFL entrants in the current research would depend on automatized explicit knowledge, given their limited exposure to the English language.

Procedures

The collection of the data continued over six months; firstly the learners completed a background questionnaire in their first language (Arabic), and after that they worked on the pretests and the grammar acquisition test. The oral production pretest was the first test the learners began with, and then they continued with the Elicited Imitation pretest and, finally, the written production pretest. The target during the oral tests was to raise the chance of the learners depending on procedural knowledge. Implementing the written production test before the oral test may help learners to be in a better prepared status and into a more explicit mode during the following oral test because the written production test is more probably reliant on declarative, explicit knowledge. For practical reasons, the

current study has preferred to make the oral production test before the Elicited Imitation test; one reason was that the oral production test was shorter; thus, the learners who did not suit the implying standard could be excluded at an earlier stage. In order to apply random tasks that controlled for likely differences in aptitude across the two groups, the learners were administered the LLAMA D, E, and F tests after the pretests.

The learners who met the selection standard were asked to take part in the two treatment sessions, which continued for about ten minutes with a five-minute break between the two sessions. In the time of receiving recasts in response to errors with the target feature, the learners performed either two decision-making assignments or two data transmission tasks. A posttest questionnaire was given after each treatment task with a five-minute break after fulfilling each treatment, after that the learners administered the posttests (oral and written production tests, Elicited Imitation test) in the same order as the pretests. It was important to perform the research in a quiet environment because some of the tests and the treatment engaged processing oral input; hence, the researcher in all sessions met the learners individually.

Analysis of the Data

Transcription

The researcher transcribed all the oral production information from the tests and treatment assignments. Another transcriber was asked to transcribe ten percent of the information that was chosen by applied random sampling to ensure equal representation of the experimental sets and different versions of the treatment and testing assignments to verify the reliability of the transcriptions. The researcher compared the two transcripts with concentration on the target verbs, then two types of discrepancies appeared; one was that some items were introduced in one transcript but deleted from the other, while the other discrepancy was that some items were differently transcribed. The current study calculated Cohen's kappa at .96 in addition to dividing the whole number of items transcribed equally by the total number of items, and it was found to be at .98, which means it is high.

Scoring and Coding

The learners' responses to recasts were considered the first step to show whether recasts led to successful uptake, as was previously mentioned by Lyster and Ranta (1997). The current study adopted the definition of Kourtali and Révész (2019), which defined uptake as "a learner's utterance that follows the instructor's feedback at once and that models a reaction in some way to the instructor's desire to pay attention to some parts of the learner's first pronouncement." example (3) showed that uptake was coded as successful when learners corrected their first error specific to the present third-person singular verb form and unsuccessful when they did not. For more illustration, when learners made the same error again or made a different one, they answered "yes" or continued with the next picture in the assignment (Lyster & Ranta, 1997; Ellis, Basturkmen, & Loewen, 2001).

Example (3)

Student: Sandy on Friday 4 or 5 o'clock at public garden and paint the portrait.

Teacher: She paints

Student: Paints the portrait.

Mandatory contexts for the target feature were determined to be used in the oral and written production tests, as the learner is given one grade if the target feature was accurately produced. Also, errors of orthography, pronunciation, and vocabulary were neglected. The overall grades were utilized in more analysis because the learners generated various numbers of compulsory contexts for each item; hence, it is out of the question to make an item-based analysis. The current study followed previous studies (Erlam, 2006; Kourtali & Révész, 2019) in preparing the scoring standard for the Elicited Imitation test. In order to ensure that learners had treated the meaning of the pronunciation, they were given one grade when selecting the right picture. The researcher included in the study only the learners who achieved at least 90%. In dealing with the Elicited Imitation test, the learners were given one grade when they created the target feature, but when they produced other constructions, they gained zero; in addition, the grades of learners' Elicited Imitation tests were measured for grammatical and ungrammatical items combined and separately. Another coder coded 10% of the information for each outcome measure; also, the values of Cohen's kappa were high: .97 for the written production test, .98 for the Elicited Imitation test, and .96 for the oral production test.

Statistical Analyses

The current study followed previous studies (Faul, Erdfelder, Lang, & Buchner, 2007; Kourtali & Révész, 2019) in using GPower 3.1 to hold out a power analysis for all statistical tests. It was found that the size of the sample was enough to reveal impact sizes in the medium range for all tests. It was necessary in this study to utilize SPSS Version 22 to measure descriptive statistics, rating the reliability of the Elicited Imitation test, number of mandatory contexts elicited, aptitude profiles, and pretest scores. Concerning the other statistical analyses, the study of Bates, Maechler, Bolker, and Walker, 2015) and R Core Team (2018) advised constructing linear mixed-effects models employing the lm and glmer functions from the lme package (Version 1.1-21) in the R statistical environment. Drawing on the lm mission, the multiple regression analyses were applied for the oral and written output test data. The pretest scores worked as a covariate and the foretellers of interest were task complexity (First Research Question), the LLAMA scores (Second Research Question), or their interactions relying on the research question (Third Research Question). The researcher considered the overall score obtained on the oral or written production test as the dependent variable; for the Elicited Imitation test, mixed-influences logistic regression analyses were employed using the function glmer due to the twofold nature of the dependent variable. In this study, learners and items were included in the random effects, and the score of the Elicited Imitation was the dependent variable;

in addition, the LLAMA scores (Second& Third Research Questions), their interactions (Third Research Question) and/or the constant effects were task complexity (First & Third Research Questions). It was found that the data for the Elicited Imitation scores combined, in spite of grammaticality, was found to be a significant foreteller of the whole Elicited Imitation scores in a preliminary analysis using grammaticality as a constant effect, Value = –5.56, SE = .4, z = –27.96, p < .01, and participant (SD = 3.58) and item (SD = .1) as random effects. The main reason behind this decision was that analogous outcomes were found for the patterns with and without grammaticality as a random effect in addition to the split-up analyses held for grammatical and ungrammatical Elicited Imitation items (see Appendix A).

The estimates of R2 were acquired to calculate effect sizes for the multiple regression analyses and to evaluate the size of effects on the mixed-effects logistic regression patterns, and the odds ratios (ORs) were calculated. In order to examine normality assumptions for the patterns, as well as linearity and homoscedasticity, the current study utilized residual plots. Also, d values of 0.8, 1.4, and 2 and R2 values of 0.12, 0.32, and 0.72 were considered small, medium, and large, respectively as mentioned in some previous studies (Plonsky & Oswald, 2014; Kourtali & Révész, 2019).

Results

The researcher divided this section into four parts as follows:

Part A: Initial analyses

Part B: Research question (1) Are there any effects of task complexity on developing E FL learners' use of the present third person singular, and what are these effects?

Part C: Research question (2) Does aptitude predicts development of EFL learners' use of the present third person singular? To what extent?

Part D: Research question (3) Does task complexity influence the relation between FL aptitude and development in E FL learners' use of the present third person singular? To what extent?

Part A

Initial analyses

The current study compared learners' conceptions about the intellectual effort they made while fulfilling the treatment assignments to be able to validate task complexity manipulation. The participants concluded that the information transmission tasks were less cognitively demanding than the decision-making assignments, as shown in Table (1). It was assured through the independent-samples t tests that the reported intellectual effort during both treatment sessions: Treatment 1, t = 8.94, p < .01; while Treatment 2, t = 6.76, p < .01 was significantly lower for the information tasks than for the decision-making assignments, which were remarkably higher. In Treatment one the effect size was

sizable (d= 2.3), while in Treatment two the effect size was average (d = 1.72), which indicates that according to the learners' conceptions, task complexity manipulation was successful.

Table 1. Descriptive statistics for the perceived intellectual effort scale

		Treatment task (1)			Treatment task (2)		
Set name	N	M	SD	95% CI	M	SD	95%CI
Information transmission	60	7.66	3.48	[6.36, 8.96]	6.86	3.42	[5.58, 8.14]
Decision-making	60	11.86	3.82	[10.44, 13.08]	10.46	4.7	[8.7, 12.22]

Note. CI = confidence interval.

The researcher showed the number of recasts that the learners received, the amount of successful uptake they generated, and descriptive statistics for the number of compulsory contexts that the assignments produced for the target feature during the treatment (see Table 2). It was clear that the two sets had an analogous number of opportunities to present the target feature and receive feedback on their use, as the independent-samples t-test indicated no significant difference between the number of mandatory contexts for the two sets, t = 2.38, p = .48, d = 0.62. Moreover, there was no significant difference between the number of recasts the two sets received, t = 3.78, p = 0.78, d = 0.96, as presented by the independent-samples t-test. The independent-samples t-test revealed that after receiving recasts, the two sets generated comparable amounts of successful uptake, which means that there were no differences between the experimental sets in terms of the extent to which they corrected the target feature in response to recasts, t = 1.98, p = .66, d = 0.5. Regarding potential relationships between successful uptake and aptitude in the two sets, a series of Pearson correlations were used and showed that in the complex condition, there was only a weak positive correlation between LLAMA E and successful uptake (r =0 .74, p = 0.1). No significant relationship was identified for LLAMA D (r = –0.24, p = 1.04) or for LLAMA F (r = 0.52, p = 0.32). While in the simple condition, there was no relationship found for LLAMA D (r = –0.68, p = 0.12), for LLAMA E (r = 0.18, p = 1.3), or for LLAMA F (r = –0.04, p = 1.8).

Table 2. Descriptive statistics for compulsory contexts, number of recast, successful uptake (percent) per set.

Set	N	M	SD	95% CI
Compulsory contexts				
Data transmission	60	61.32	10.02	[57.58, 65.06]
Decision-making	60	65.32	15.42	[59.56, 71.08]
Number of recasts				
Data transmission	60	30.8	16.82	[24.52, 37.08]
Decision- making	60	39.46	18.56	[32.54, 46.4]
Successful uptake				
Data transmission	60	100.94	73.3	[73.56, 128.32]
Decision- making	60	82.74	68.74	[56.7, 108.8]

Note. CI = confidence interval.

Part B

Effects of task complexity on developing elementary EFL learners' use of the present third-person singular

Exploring the effects of task complexity on development in the knowledge of the target feature when recasts are given was the main target of the first question. From the pretest to the posttest, the learners showed progress in all tests, as the data transmission set was extremely superior in both tests, the oral and written productions more than the decision-making set. In spite of this, the difference between the two groups' gains on the Elicited Imitation test was small.

Table 3. Descriptive statistics for pretest and posttest scores on the evaluation assignments per set

		Pretest			Posttest		
Set	N	M	SD	95% CI	M	SD	95%CI
Oral production							
Data transmission	60	7.89	16.38	[1.86, 14.12]	67.22	61.48	[44.28, 90.18]
Decision- making	60	7.5	13.82	[2.36, 12.68]	34.62	53.08	[14.8, 54.44]
Written production							
Data transmission	60	8.2	18.44	[1.32, 15.1]	114.1	88.64	[81, 147.2]
Decision- making	60	9.86	18.6	[2.92, 16.82]	53.54	79	[24.04, 83.04]
Elicited Imitation							
Data transmission	60	8.52	8.5	[5.34, 11.7]	13.6	10.6	[9.62, 17.56]
Decision- making	60	8	8.2	[4.92, 11.06]	12	10.4	[8.1, 15.88]

Note. CI = confidence interval. The total score was 48 for the Elicited Imitation test.

Any significant difference revealed at the time of the posttest could not be assigned to differential use of the third-person singular –s at the beginning of the experiment because the independent-samples t-tests that were completed to explore whether the two sets had differential prior knowledge of the target structure at the time of the pretest found no significant difference for any of the three tests: oral production, $t = 0.24$, $p = 1.9$, $d = 0.06$; written production, $t = 0.7$, $p = 1.46$, $d = 0.18$; EI overall, $t = 0.3$, $p = 1.76$, $d = 0.12$. The linear regression analyses were held for the two tests one by one, applying the linked pretest score and task complexity as foreteller variables in order to evaluate whether task complexity influenced the learners' progress in the use of the third-person singular –s on the oral and written production tests. The task complexity was the predictor of interest in both analyses, and the pretest score worked as the covariate. In both the oral and written production tests, the regression analyses yielded a significant effect for task complexity with effect sizes being in the tiny range (see Table 4).

Table 4. Findings for the linear regression patterns exploring the effects of task complexity on the oral and written production tests

Factor	Est	SE	T	P	R
Oral production					
Intercept	88.02	22.34	7.88	<.01	
Pretest	2.78	0.94	5.98	<.01	0.26
Task complexity	-31.94	13.9	-4.6	.03	0.16
Overall R^2					0.4
Oral_post ˜1+Oral_pre+ Task_complexity					
Written production					
Intercept	171.26	34.56	9.9	<.01	
Pretest	1.04	1.2	1.74	.39	<.01
Task complexity	-61.42	21.74	-5.66	<.01	0.24
Overall R^2					0.26
Writtenl_post ˜1+written_pre+ Task_complexity					

A mixed-effects regression analysis was executed to investigate the effects of task complexity on learners' production on the Elicited Imitation test. The random effects were item and learner, the dependent variable was the Elicited Imitation score, and the fixed effects were learners' pretest scores and task complexity. Table 5 showed that the pretest scores worked as the covariate, the foreteller of interest was task complexity, and no significant effects appeared for task complexity.

Table 5. Findings for the linear mixed- effects patterns checking the effects of task complexity on the elicited imitation test

	Fixed effects				Random effects	
Factor	Est	SE	Z	P	Factor	SD
Intercept	-2.92	1.02	-5.66	<.01	Learner	2.2
Pretest	5.06	0.4	25.88	<.01	Items	0.42
Task complexity	-0.5	0.66	-1.56	.44		
EI_post ˜1+EI_pre+ Task_complexity+ (1\ Learner)+ (1\Item)						

The previous results showed that despite the set assignment, the learners had acquired parallel gains on the Elicited Imitation test, but they had significantly greater gains on the oral and written production tests when they finished information transmission compared to decision-making tasks during the treatment sessions.

Part C

Aptitude predicts development of EFL learners' use of the present third-person singular

To be able to answer the second question, which was trying to investigate the extent to which aptitude predicts progress in the use of the target structure when participants

receive recasts, the researcher ran a chain of Pearson correlations to check the relationships between the different LLAMA elements, but the results showed that there was no significant correlations, suggesting that the different LLAMA tests quarried various constructs (see Table 6).

Table6. Correlation between LLAMA scores

	LLAMA D	LLAMA E
LLAMA E	0.164 (1.068)	
LLAMA F	-0.444 (0.178)	0.248 (0.692)

Note. p values in parentheses.

Table (7) below shows the descriptive statistics for the LLAMA aptitude scores by set, and there was no significant difference between the two sets on any of the LLAMA test elements in a chain of independent- samples t-tests: LLAMA D, t = 0.82, p = 1.36, d = 0.22; LLAMA E, t = 0.08, p = 1.94, d = 0.02; and LLAMA F, t < 0.02, p = 2, d = 0 (see Table 7). Based on the previous findings, the two sets mostly had comparable aptitude profiles, except the possibility that differences in aptitude among the sets might account for the differential gains noticed in both sets. Additionally, the reason behind including aptitude as a fixed effect in more analyses was that the criteria deviations were large for both sets, indicating sizeable within-set variation in aptitude outcomes.

Table 7. Descriptive statistics for EFL aptitude scores per set

Component	Set	N	M	SD	95%CI
LLAMA D	Information transmission	60	65.32	22.72	[38.84, 55.82]
	Decision making	60	44.32	32.66	[32.14,56.52]
LLAMA E	Information transmission	60	85.32	52.5	[65.72, 104.94]
	Decision making	60	86	63.7	[62.2, 109.8]
LLAMA F	Information transmission	60	43.32	40.7	[28.14, 58.54]
	Decision making	60	43.32	39.68	[28.52, 58.16]

-*Note*. . CI = confidence interval. The total score was 150 for LLAMA D and 200 for LLAMAs E and F

For investigating whether aptitude accounted for the extent to which learners showed development in the use of the target structure on the oral and written production tests as a result of finished receiving feedback and communicative assignments, the current study carried out multiple regression analyses (See Appendix B), which also shows a correlation matrix summarizing the relationships among the LLAMA and gain scores. The pretest scores and the LLAMA D, LLAMA E, and LLAMA F aptitude scores were assigned as the foreteller in the analyses for both tests; in addition, Table 8 demonstrated that none of the aptitude measures appear to be a significant predictor as the pretest scores worked as the covariate, and the aptitude scores were the predictors of interest.

Table 8. Findings for the multiple regression analyses investigating the relationship among gains on the oral and written production tests and aptitude

Factor	Est	SE	T	P	R^2
Oral production					
Intercept	12.04	19.98	1.2	.55	
Pretest	2.84	0.96	5.88	<.01	0.26
LLAMA D	0.7	0.54	2.6	.20	0.06
LLAMA E	0.28	0.26	2.18	.28	0.04
LLAMA F	<0.01	0.38	<0.01	1.00	<.01
Overall R^2					0.36
Oral_post ˜1+Oral_pre+ LLAMA_D+ LLAMA_E +LLAMA_F					
Written production					
Intercept	35.24	31.3	2.26	.27	
Pretest	0.74	1.28	1.14	.57	0.02
LLAMA D	1.2	0.86	2.78	0.17	0.02
LLAMA E	0.3	0.4	1.42	.48	<.01
LLAMA F	0.24	0.62	0.8	.69	<.01
Overall R^2					0.12
Write_post ˜1+Write_pre+ LLAMA_D+ LLAMA_E +LLAMA_F					

One more mixed-effects regression analysis was conducted to examine the extent to which aptitude interpret learners' progress on the Elicited Imitation test, the fixed effects were learners' pretest scores and the LLAMA D, LLAMA E, and LLAMA F scores while the dependent variable was the Elicited Imitation score. The researcher added item and learner as random effects to each pattern. Table 9 demonstrated that the predictors of interest were the aptitude scores, pretest scores worked as the covariate, and in the analyses the only one that seemed a significant predictor was LLAMA E. It is obvious that the learners with higher LLAMA E scores performed better on the Elicited Imitation test, OR = 2.04, CI.95 = [2, 2.06], while those who accomplished one point higher on the LLAMA E test were 4% more likely to score a point higher on the Elicited Imitation test. In using the linguistic target there was no significant predictors of learners' development found, as for the LLAMA D scores, OR = 2.02, CI.95 = [1.98, 2.08], and for the LLAMA F scores, OR = 2, CI.95 = [1.98, 2.04].

Table 9. Findings for the linear mixed-effects patterns investigating the relationship between gains on the Elicited Imitation test and aptitude

	Fixed effects				Random effects	
Factor	Est	SE	Z	P	Factor	SD
Intercept	-5.7	0.86	-13.22	<.01	Learner	1.98
Pretest	5.04	0.38	25.96	<.01	Items	0.42
LLAMA D	0.02	0.02	2.38	.23		
LLAMA E	0.02	<0.01	5.56	<.01		
LLAMA F	<0.01	0.02	0.72	.72		
EI_ post ˜1+EI_pre+ LLAMA_D+ LLAMA_E+ LLAMA_F (1\ Learner)+ (1\Item)						

Part D

Task complexity influence the relation between EFL aptitude and development in EFL learners' use of the present third-person singular

The current study used separate various regression analyses for the oral and written production tests to answer the third research question, and these analyses showed a significant interaction between LLAMA D and task complexity. Table 10 demonstrated that the pretest scores served as the covariate, and the foretellers of interest were the interactions; also, the predictors in the patterns were the pretest scores, the interactions between task complexity and the LLAMA scores, the three LLAMA scores, and task complexity.

Table 10. The effect of task complexity on the relationship between development in the oral and written production tests and aptitude

Factor	Est	SE	T	P	R^2
Oral production					
Intercept	199.22	72.36	5.5	<.01	
Pretest	2.8	0.92	6.12	<.01	0.26
Task complexity	-110.58	40.72	-5.44	<.01	0.16
LLAMA D	-3.64	1.96	-3.72	.07	0.06
LLAMA E	-0.38	0.82	-0.92	.65	0.04
LLAMA F	-0.5	1.16	-0.88	.67	<.01
Task complexity × LLAMA D	2.5	1.12	4.5	.03	0.1
Task complexity × LLAMA E	0.42	0.5	1.68	.41	0.02
Task complexity × LLAMA F	0.18	0.74	0.46	.82	<.01
Overall R^2					0.66
Oral_ prod_ post ˜1+Oral_prod_ pre+ LLAMA_D* Task_ comp+ LLAMA_E * Task_ comp+ LLAMA_F* Task_ comp					
Written production					
Intercept	404.8	111.94	7.24	<.01	
Pretest	1.22	1.18	2.06	.31	<.01
Task complexity	-219.96	63.44	-6.94	<.01	0.24
LLAMA D	-5.92	3.02	-3.9	.06	0.08
LLAMA E	-1.46	1.26	-2.3	.26	0.02
LLAMA F	-1.62	1.82	-1.78	.38	<.01
Task complexity × LLAMA D	4	1.72	4.68	.02	0.14
Task complexity × LLAMA E	1.08	0.78	2.82	.17	0.06
Task complexity × LLAMA F	0.92	1.14	1.6	.43	<.01
Overall R^2					0.56
Written_ prod_ post ˜1+Written_prod_ pre+ LLAMA_D* Task_ comp+ LLAMA_E * Task_ comp+ LLAMA_F* Task_ comp					

The follow-up regression analyses were run for both experimental sets one by one, using the pretest scores and LLAMA D as foretellers to investigate the interaction effects. LLAMA D appeared with both the written and oral production data to be a significant predictor of progress for the decision-making set in the oral production test, value = 1.44, SE = 0.5, t = 5.8, p < .01, R2 = 0.38, and in the written production test, value = 2.2, SE =0.8, t = 5.46, p = .01, R2 = 0.4, but not for the information transmission set in the oral

production test, value = –0.9, SE = 0.96, t = –1.86, p = .36, R2 = 0.06, or the written production test, value = –1.44, SE = 1.48, t = –1.96, p = .34, R2 = 0.06. Former findings means that the learners in both the oral and written production tests in the decision-making set with higher LLAMA D scores showed greater gains, but in the information transmission set, the effect sizes were in the tiny range because LLAMA D made no difference in gains.

A mixed-effects logistic regression analysis was run to examine whether task complexity moderated the extent to which aptitude predicted learners' progress on the Elicited Imitation test. The fixed effects were learners' pretest scores, task complexity, the LLAMA scores (LLAMA D, LLAMA E, or LLAMA F), and the interactions between task complexity and the LLAMA scores, and the dependent variable was the Elicited Imitation score. Also, learner and item worked as random effects in the pattern, the pretest scores served as the covariate, the researcher predictors of interest were the interactions, and patterns did not show a significant interaction effect (See Table 11).

Table 11. The effects of task complexity on the relationship between progress on the Elicited Imitation test and aptitude

	Fixed effects				Random effects	
Factor	Est	SE	Z	P	Factor	SD
Intercept	-1.98	0.94	-4.24	.03	Learner	1.94
Pretest	1.9	0.14	25.98	<.01	Items	0.42
Task complexity	0.36	1.18	0.6	.76		
LLAMA D	-0.54	0.6	-1.84	.36		
LLAMA E	-0.02	1.02	-0.02	.99		
LLAMA F	0.2	1	0.42	.83		
Task complexity × LLAMA D	<0.02	0.68	-0.02	.99		
Task complexity × LLAMA E	0.56	0.64	1.8	.37		
Task complexity × LLAMA F	-0.12	0.64	-0.36	.86		
EI_ post ~1+EI_pre+ Task _ comp* LLAMA_D+ Task_ comp * LLAMA_E+ Task_ comp* LLAMA_F (1\ Learner)+ (1\Item)						

Discussion

Grammar Acquisition and task complexity

Studying the effects of task complexity on participants' development in the acquisition of the target structure when they received recasts was the main target of the first research question. The findings of grammar acquisition tests detected that the participants benefited less when they executed tasks and received recasts in more complex, decision-making tasks imposing greater reasoning demands, while they benefited more in the simple condition engaging only information transmission. The current study found that the effect size for this difference was tiny, harmonious with the

small overall effect size found for accuracy in Jackson and Suethanapornkul's (2013) meta-analysis exploring the influences of task complexity on EFL output.

It was previously discussed Skehan's (2009, 2014) limited attentional capacity model that predicts that greater cognitive requests will put raised pressure on the conceptualizer, leaving fewer attentional resources for linguistic encoding processes and the findings here reflect the predictions of this pattern. At the same line like Skehan, participants likely had less capacity left to pay attention and process feedback focusing on linguistic errors, given the increased effort the task required in terms of conceptualization when they were under the complex condition. This explanation may mean that learners had few attentional resources available to assign to the target structure, a communicatively superfluous element for successfully accomplishing the assignment, and this happened because of the need to devote greater attention to the communicative requests of the assignment. Learners under the simple condition were likely to have more cognitive capacity to observe and process feedback focusing on a nonsalient grammatical structure because it assumed that learners had to give less attention to conceptualization.

It was noticed that the results of the first research question correspond with the study of Révész et al. (2014), as the participants in both studies benefitted more from recasts under simple task conditions than complex ones. Additionally, in both studies the learners received recasts when they produced a non-target structure during simple and complex monologic assignments. In spite of the harmony between the current study and Révész's (2014) study, Baralt (2013) opposed these results because the oral face-to-face recasts were more efficient under complex interactive dialogic tasks than simple conditions, which means that this point needs more research to illuminate possible effects of task design on the efficacy of feedback provided during cognitively simple and complex tasks such as dialogic tasks versus monologic.

In addition to the previous findings, it is important to mention that regardless of the simple condition seeming more useful than the complex condition, both sets showed remarkable FL gains, although the short intervention and these gains may be due to the way of presenting recasts. In the present study, recasts could be categorized as reduced, focused on one change (Lyster, 1998), explicit (Sheen, 2006), provided in an intensive and focused way (Han, 2002), and entailing the correction in utterance-final position (Bardovi-Harlig, 1987), as previously mentioned in the methodology section.

Finally, it is important to discuss the reasons behind the disparity in the learners' improvement, as they showed little development on the Elicited Imitation test and somewhat less progress on the oral production test, while they improved more on the written production test. The only test that permitted production of the target structure under less time pressure, thereby giving more opportunities for the deployment of declarative knowledge, was the written production test. The participants maybe faced no problem accessing the declarative knowledge they had gained through exposure to oral content when involved in assignments in the written condition because declarative knowledge is adjustable to transfer through modalities (Anderson, 1993). Nevertheless,

the treatment did not enable learners to accomplish the stage of automatic production, which would have permitted them to perform better on the oral tests that imposed more time pressure, and this inability was because of the short intervention.

Aptitude as A foreteller of EFL Grammar Acquisition

Investigating the extent to which aptitude predicted EFL learners' use of the target feature when they received recasts was the main target of the second research question. In general, the findings showed that better phonetic coding ability was connected to higher scores on the Elicited Imitation test, as determined by the LLAMA E test; also, the learners who obtained one point higher on the LLAMA E test were 2% more likely to score a point higher on the Elicited Imitation test. It was found that to some extent, the small effect size corresponds to the outcomes of Li's (2016) meta-analysis, which shows a moderate generally association between EFL grammar acquisition and aptitude (r =0.62).

It is clear that the findings of the LLAMA E test are not unexpected, as the LLAMA E test assessed learners' phonetic coding ability, which helped the participants to divide words into phonetic units and analyze how these connect to symbolic units. In similar processes, learners were involved when processing recasts during the treatment and when they received oral content on the Elicited Imitation test, as participants needed to divide words into phonetic units and analyze the meanings linked with these units. As a result, the learners who achieved higher LLAMA E scores were better able to decode recasts, perceive the sound models connected with the present third-person singular morpheme, and support this form with its mission. Another possibility is that the ability also likely assisted the learners with processing the oral content on the Elicited Imitation test more successfully and hence to cope with the requirements of the task, like understanding oral content and giving output content. To sum up, a positive relationship between this aptitude element and EFL grammar acquisition, even if relatively small, would take into consideration the overlaps in the constructs calculated by LLAMA E, the Elicited Imitation test, and the nature of the treatment.

The findings of the current study related to potential relationships among the LLAMA test elements and implicit versus explicit learning seems to correspond with Grañena's (2013) observations, as she mentioned that the LLAMA E test might be more sensitive to tapping aptitude elements engaging in explicit learning (learning via deduction of rules), while the learning conditions generated by the LLAMA D test look like implicit learning environments (learning through exposure) to a greater range. Depending on the previous results, the researcher expects that as the Elicited Imitation test is traditionally regarded as a test of implicit knowledge as mentioned by Ellis (2005) and Suzuki and DeKeyser (2015), so the learners who achieved higher scores on LLAMA D than LLAMA E would obtain greater gains on the Elicited Imitation test. In order to deal with this seemingly paradoxical result, the researcher chose to use the Elicited Imitation test because it may lead at least in part to the use of automatized explicit knowledge, as discussed by Suzuki and DeKeyser (2015). It is not unlikely that learners

with high LLAMA E scores were better able to enhance their explicit, declarative knowledge over exposure to explicit recasts because the learners are from a foreign language school that trained them to analyze language explicitly. Also, the learners were able to analyze the oral content of the Elicited Imitation test and obtain better scores with the assistance of the increased explicit knowledge that high LLAMA E participants had achieved.

It is important to mention that the facilitative role of LLAMA E also agrees with Skehan's (2002) suggestion that phonetic coding is expected to support EFL results in the first stages of EFL grammar acquisition, as the entrants in the current research were low-proficiency learners. LLAMA E in previous experimental studies has been used to measure phonetic coding ability and its role has been proved. Another result discovered by Saito (2017) was that there was a positive relation between morphological accuracy and the ability to associate sounds with symbols; at the same time, Yilmaz and Koylu (2016) demonstrated that phonetic coding ability was linked to the range in which feedback is useful for learners. Both Yilmaz and Koylu's (2016) research and the current study employed explicit forms of feedback; hence, both correspond with Grañena's study (2013), which suggested that LLAMA E might be implicated in explicit learning.

Another important point is to discuss why LLAMA F scores were unsuccessful as foretellers of participant gains as with LLAMA D, either for one of the sets or the two sets combined. The type of feedback provided in the current research, as well as the nature of the target structure might be linked and affect this result. It is clear that high grammatical inferencing ability might not have played a pivotal role in discovering the linguistic model in the current study because the rule of adding a morpheme to the verb connected to the present third-person singular is relatively simple in spite of the morpheme –s is linked to different meanings, which might pose difficulty to the acquisition of the third-person –s. The analytic ability might be more instrumental in facilitating "more diverse, sophisticated, and complex lexicogrammar usage" as said by Saito (2017, p. 670); the study of Suzuki and DeKeyser (2017b) corresponded to this idea and discovering that LLAMA F was a significant indicator of enhancing automatized explicit knowledge of complex textures. The explicit partial recasts may have assisted participants to observe the nonsalient form, thereby neutralizing the role of analytic ability, which might be the reason for the lack of a relationship between LLAMA F and learners' gains. More experimental studies are needed to investigate whether analytic ability is of greater importance if giving more implicit recasts to the learners.

The Connection among Aptitude and EFL Grammar Acquisition: Task Complexity as a Moderating Variable

Exploring the effect of task complexity on the relation between FL aptitude and progress in EFL learners' use of the target structure when they received recasts was the target of the third research question. The findings showed that when participants were under complex task condition, higher LLAMA D scores correlated with better EFL results on the oral and written production tests, while when they were under simple task

condition, there was no connection found between aptitude and EFL grammar acquisition. The effect size was in line with the outcomes of Li's (2016) meta-analysis, as it appeared to be tiny.

The results of the current study from a theoretical perspective was consistent and support the findings of Robinson (2011) that individual differences in cognitive abilities were related to learners' performance and progress when tasks require greater cognitive requirements. These results were also compatible with those of Kim et al. (2015), who found that when participants were involved in cognitively complex tasks and had better working memories, they benefitted more from recasts. The current study established that the range within which EFL aptitude predicts participants' benefit from feedback may be reduced due to implementing tasks with low cognitive requirements. Also, the results correspond with Erlam (2005) and Kourtali and Révész (2019) in demonstrating that learning conditions may affect whether aptitude explains EFL grammar acquisition gains, as explicit directives were found to compensate for low aptitude.

It was not surprised to find the positive relation between the progress in the use of the present third-person singular form and LLAMA D, despite observing the complex task version only. The learners needed to process sounds and keep them in long-term memory in two cases: when taking the LLAMA D test and when receiving recasts, and this result agreed with Meara's (2005) prediction indicating that the ability to recognize models in spoken language, which LLAMA D calculates, help participants in recognizing morphological variations. At the same time, it is important to mention that during oral production, Saito (2017) found no relation between participants' LLAMA D scores and morphological accuracy, which indicates that the extent to which sound recognition ability plays a role in EFL progress may rely on the nature of the grammatical construction. Saito applied accurate morphology as a proportion of morphological errors in a wide range of structures, including more salient ones like aspect and modality features and fewer salients like verb agreement and plural. The ability to recognize sound sequences may facilitate the acquisition of morphemes realized through redundant, nonsalient forms like the present third-person singular, but may be less useful in enhancing knowledge of more salient grammatical structures. The idea of gathering more and less salient features in one accuracy ratio might reveal a relationship between grammatical accuracy ratings and LLAMA D scores.

Drawing on Lightbown's conclusion (2007), the positive role of LLAMA D in complex task performance was only observed on the oral and written production tests. The participants' ability to transfer and remember becomes better if the cognitive processes active during learning are similar to those active during retrieval. The additional learning gains might have been easier to deploy on the oral and written production tests because of the learners' excellent ability to recognize sounds during the treatment.

Limitations and Further Research

The current research encountered many limitations that need to be considered in future research. To gain a better picture of participants' progress, future studies could employ comprehension-based result measures like a grammaticality judgment task. It would have been better if the study investigated the longer-term impacts of the treatment. The researcher could not control for exposure to the target structure between the immediate and the delayed posttest because using a delayed posttest would not have created valid outcomes for the present study. The individual differences in cognitive abilities only captured by an aptitude test were considered another weakness of the current study. As well as indicators of aptitude in the follow-up study, working memory and attentional control could be added to supply a more comprehensive account of the role of cognitive individual differences in FL learners' ability to benefit from task-based interaction and recasts. Another weakness of this study is that the researcher used only one type of recasts (interrogative, partial recasts) and one type of grammatical construction. The current study also proposes that future studies could examine the influence of form interventions and other target structures like more salient grammatical features. Previous studies (see Marsden, Morgan-Short, Thompson, & Abugaber, 2018) justified both direct and conceptual replications that appeared in the present research like using learners from different FL backgrounds or proficiency levels, other types of task manipulation, and adult participants.

Conclusion & Recommendations

The main target of the current study was to examine the extent to which (1) task complexity affects EFL grammar acquisition resulting from task-based interaction, (2) aptitude foretells EFL gains, and (3) task complexity affects the relation among aptitude and EFL outcomes when recasts are given. This study used the third-person singular –s structure, a linguistic feature that includes different challenges in the acquisition of English as a foreign language. It was found that the participants who fulfilled less cognitively demanding tasks developed their use of the present third-person singular to a greater extent than participants involved in tasks with greater cognitive demands. This finding is along the same lines with Skehan's prediction (2009, 2014). Another interesting finding is that the participants of high aptitude, especially those with excellent sound recognition ability, were better able to compensate for the raised requests posed by the complex tasks than their low-aptitude opposites, mirroring Robinson's (2011) prediction concerning the relation between task complexity and aptitude. In other words, the findings of the current study recommended that low-complexity tasks can reduce the degree to which participants' differences in FL aptitude foresee improvement in task-based contexts when giving feedback to participants. One temporary pedagogical implication is that less complex tasks might be more useful for improving the grammatical knowledge of language participants in an intact language class, unless individual participants used FL aptitude profiles to give instruction types.

1398   Asghar Salimi et al. / Procedia - Social and Behavioral Sciences 29 (2011) 1390 – 1399

Brindley, J. (1987). Factors affecting task difficulty. In D. Nunan, (Ed.), Guidelines for the development of

curriculum resource (pp.45-56). Adelaide: National Curriculum Resource Center.

Brown, R., & Bellugi, U. (1964). 'Three processes in the child's acquisition of syntax'. Harvard Educational Review.

34: 133-151.

Brown, G., & Yule, G. (1983). Teaching the spoken language: an approach based on the analysis of conversational

English. Cambridge: Cambridge University Press.

Candlin, C. (1987). Towards task-based language learning. In Candlin, C. & Murphy, D. (Eds.), Language

Learning Tasks (pp.5-22). Englewood Cliffs, NJ: Prentice Hall.

Ellis, R. (2003). Task-based language learning and teaching. Oxford: Oxford   University

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Baralt, M. (2013). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition*, *35*, 689– 725. https://doi.org/10.1017/S0272263113000429

Bardovi-Harlig, K. (1987). Markedness and salience in second-language acquisition. *Language Learning*, *37*, 385– 407. https://doi.org/10.1111/j.1467-1770.1987.tb00577.x

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1– 48. Retrieved from https://www.jstatsoft.org/article/view/v067i01

Bokander, L., & Bylund, E. (2019). Probing the internal validity of the LLAMA language aptitude tests. *Language Learning*, Advance online publication. https://doi.org/10.1111/lang.12368

Carroll, J. B. (1965). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training, research, and education* (pp. 87– 136). New York, NY: Wiley.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83– 118). Rowley, MA: Newbury House.

Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test*. San Antonio, TX: Psychological Corporation.

Cobb, T. (2016). VocabProfiler [Computer software]. Retrieved from http://www.lextutor.ca/vp/eng

DeKeyser, R. (2007). Skill acquisition theory. In B. Van Patten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 97– 113). Mahwah, NJ: Erlbaum.

Ellis, R. (2003). Task-based language learning and teaching. Oxford: Oxford University Press.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*, 141– 172. https://doi.org/10.1017/S0272263105050096

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, *51*, 281– 318. https://doi.org/10.1111/1467-9922.00156

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

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*, 147– 171. https://doi.org/10.1191/1362168805lr161oa

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, *27*, 464– 491. https://doi.org/10.1093/applin/aml001

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175– 191. https://doi.org/10.1037/0033-2909.112.1.155

Grañena, G. (2013). Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In G. Grañena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 105– 130). Amsterdam, Netherlands: John Benjamins.

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The CANAL-F theory and test. *The Modern Language Journal*, *84*, 390– 405. https://doi.org/10.1111/0026-7902.00076

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 543– 572. https://doi.org/10.2307/3588240

Han, Z. (2013). Forty years later: Updating the fossilization hypothesis. *Language Teaching*, *46*, 133– 171. https://doi.org/10.1017/S0261444812000511

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, *19*, 379– 400. https://doi.org/10.1017/S0272263197003045

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). Range and frequency programs [Computer software]. Retrieved from http://www.victoria.ac.nz/lals/resources/range.aspx

Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, *63*, 330– 367. https://doi.org/10.1111/lang.12008

Keating, G. D., & Jegerski, J. (2014). Experimental designs in sentence processing research: A methodological review and user's guide. *Studies in Second Language Acquisition*, *37*, 1– 32. https://doi.org/https://doi.org/10.1017/S0272263114000187

Kim, Y., Payant, C., & Pearson, P. (2015). The intersection of task-based interaction, task complexity, and working memory: L2 question development through recasts in a laboratory setting. *Studies in Second Language Acquisition*, *37*, 549– 581. https://doi.org/10.1017/S0272263114000618

Kourtali, N.-E., & Révész, A. (2019). The roles of recasts, task complexity, and aptitude in child second language development. *Language Learning*, 1– 40. Wiley Online Library https: //doi.org/10.1111/ lang.12374

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*, 309– 365. https://doi.org/10.1111/j.1467-9922.2010.00561.x

Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal*, *97*, 634– 654. https://doi.org/10.1111/j.1540-4781.2013.12030.x

Li, S. (2016). The construct validity of language aptitude: A meta-analysis. *Studies in Second Language Acquisition*, *38*, 801–842. https://doi.org/10.1017/S027226311500042X

Lightbown, P. M. (2007). Transfer appropriate processing as a model for classroom second language acquisition. In Z.-H. Han (Ed.), *Understanding second language process* (pp. 27– 44). Clevedon, UK: Multilingual Matters.

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., … Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63,* 530–566. https://doi.org/10.1111/lang.12011

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, *51,* 285–329. https://doi.org/10.1017/S0261444818000125

Longe, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & B. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413– 468). New York, NY: Academic Press.

Longe, M. H. (2015). *Second language acquisition and task-based language teaching*. Hoboken, NJ: Wiley. Blackwell.

Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, *20*, 51–81. https://doi.org/10.1017/S027226319800103X

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*, 37– 66. https://doi.org/10.1017/S0272263197001034

Lyster, R., & 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

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, *68*, 321–391. https://doi.org/10.1111/lang.12286

Meara, P. (2005). *Llama language aptitude tests: The manual*. Swansea, UK: Lognostics.

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

Norris, J. M. (2010, September). *Understanding instructed SLA: Constructs, contexts, and consequences*. Plenary address delivered at the annual conference of the European Second Language Association (EUROSLA), Reggio Emilia, Italy.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64,* 878– 912. https://doi.org/10.1111/lang.12079

R Core Team. (2018). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 159– 180). Amsterdam, Netherlands: John Benjamins.

Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, *31*, 437– 470. https://doi.org/10.1017/S0272263109090366

Révész, A. (2014). Towards a fuller assessment of cognitive models of task-based learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics*, *35*, 87– 92. https://doi.org/10.1093/applin/amt039

Révész, A., Sachs, R., & Hama, M. (2014). The effects of task complexity and input frequency on the acquisition of the past counterfactual construction through recasts. *Language Learning*, *64*, 615– 650. https://doi.org/10.1111/lang.12061

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*, 27– 57. https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2003). The cognitive hypothesis, task design, and adult task-based language learning. *Studies in Second Language Acquisition*, *21*, 45– 105. Retrieved from https://doi.org/http://hdl.handle.net/10125/40656

Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3– 38). Amsterdam, Netherlands: John Benjamins.

Saito, K. (2017). Effects of sound, vocabulary, and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning*, *67*, 665– 693. https://doi.org/10.1111/lang.12244

Sato, M., & Loewen, S. (2018). Metacognitive instruction enhances the effectiveness of corrective feedback: Variable effects of feedback types and linguistic targets. *Language Learning*, *68*, 507– 545.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, *10*, 361– 392. https://doi.org/10.1191/1362168806lr203oa

Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301–322). Oxford, UK: Oxford University Press.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69– 93). Amsterdam, Netherlands: John Benjamins.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*, 510– 532. https://doi.org/10.1093/applin/amp047

Skehan, P. (2014). Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–260). Amsterdam, Netherlands: John Benjamins.

Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Grañena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 17–40). Amsterdam, Netherlands: John Benjamins.

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, **38**, 1229– 1261. https://doi.org/10.1017/S014271641700011X

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*, 860– 895. https://doi.org/10.1111/lang.12138

Suzuki, Y., & DeKeyser, R. (2017a). Exploratory research on second language practice distribution: An aptitude × treatment interaction. *Applied Psycholinguistics*, *38*, 27– 56. https://doi.org/10.1017/S0142716416000084

Suzuki, Y., & DeKeyser, R. (2017b). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*, 747– 790. https://doi.org/10.1111/lang.12241

Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A.

Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171– 195). Oxford, UK: Oxford University Press.

Van Patten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Westport, CT: Ablex.

Yilmaz, Y. (2012). The relative effects of explicit correction and recasts on two target structures via two communication modes. *Language Learning*, *62*, 1134– 1169. https://doi.org/10.1111/j.1467-9922.2012.00726.x

Yilmaz, Y. (2013). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics*, *34*, 344– 368. https://doi.org/10.1093/applin/ams044

Yilmaz, Y., & Grañena, G. (2016). The role of cognitive aptitudes for explicit language learning in the relative effects of explicit and implicit feedback. *Bilingualism: Language and Cognition*, *19*, 147– 161. https://doi.org/10.1017/S136672891400090X

Yilmaz, Y., & Koylu, Y. (2016). The interaction between feedback exposure condition and phonetic coding ability. In G. Grañena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 303– 326). Amsterdam, Netherlands: John Benjamins. https://doi.org/10.1111/1467-9922.00156

Appendix A

Further findings From Linear Mixed-Effects paradigms

Outcomes for the Linear Mixed-Effects paradigms investigating the Task Complexity on the Elicited Imitation Test with considering grammaticality

Test	Factor	Fixed effects				Random effects	
		Est	SE	Z	P	Factor	SD
	Intercept	-3.06	2.1	-2.92	.14	Learner	3.02
EI	Pretest	3.44	0.44	15.6	<.01	Items	0.68
Overall	Task complexity	-0.66	0.86	-1.54	.44	Gram.	2.28

EI_Post ~ 1 + EI_Pre + Task_comp + (1|Participant) + (1|Item) + (1|Grammaticality)

Test	Factor	Fixed effects				Random effects	
EI	Intercept	-0.84	1.32	-1.26	.53	Learner	2.76
Grammatical	Pretest	3.78	0.54	13.98	<.01	Item	1
	Task complexity	-0.74	0.82	-1.78	.37		

EI_Gram_Post ~ 1 + EI_Gram_Pre + Task_comp + (1|Participant)+ (1|Item)

Test	Factor	Fixed effects				Random effects	
EI	Intercept	-6.84	2.12	-6.46	<.01	Learner	3.84
Ungramm	Pretest	2.92	1.1	5.36	<.01	Item	1.02
	Task complexity	-0.28	1.3	-0.42	.83		

EI_Ungram_Post ~ 1 + EI_Ungram_Pre + Task_comp + (1|Participant) + (1|Item)

Note. Gram. = Grammaticality

Outcomes for the Linear Mixed-Effects paradigms investigating the Aptitude on the Elicited Imitation Test with considering grammaticality

		Fixed effects				Random effects	
Test	Factor	Est	SE	Z	P	Factor	SD
	Intercept	-6.72	1.96	-6.84	<.01	Learner	2.74
EI	Pretest	3.42	0.44	15.62	<.01	Items	0.68
Overall	LLAMA D	0.04	0.02	2.56	.20	Gram.	2.28
	LLAMA E	0.04	<0.02	5.54	<.01		
	LLAMA F	<.02	0.02	0.52	.79		

EI_Post ~ 1 + EI_Pre + LLAMA_D + LLAMA_E + LLAMA_F + (1|Participant) + (1|Item) + (1|Grammaticality)

	Intercept	-4.06	1.02	-7.96	<.01	Learner	2.38
EI	Pretest	3.9	0.5	15.5	<.01		
Gramm	LLAMA D	<.02	0.02	1.12	.58		
	LLAMA E	0.04	<.02	4.72	.02		
	LLAMA F	0.04	0.02	2.1	.29		

EI_Gram_Post ~ 1 + EI_Gram_Pre + LLAMA_D + LLAMA_E + LLAMA_F + (1|Participant)

	Intercept	-10.4	1.9	-10.98	<.01	Learner	3.2
EI	Pretest	3.02	1.08	5.54	<.01	Item	1
Ungramm	LLAMA D	0.08	0.04	3.22	.11		
	LLAMA E	0.06	0.02	4.96	.01		
	LLAMA F	-0.02	0.04	-1.9	.34		

EI_Ungram_Post ~ 1 + EI_Ungram_Pre + LLAMA_D + LLAMA_E + LLAMA_F + (1|Participant) + (1|Item)

Note. Gram. = Grammaticality

Outcomes for the Linear Mixed-Effects paradigms investigating the Effects of Task Complexity on the relation among Results on the Elicited Imitation Test and Aptitude

		Fixed effects				Random effects	
Test	Factor	Est	SE	Z	P	Factor	SD
	Intercept	-2.46	2.02	-2.42	.23	Learner	2.68
	Pretest	1.3	0.16	15.66	<.01	Items	0.68
	Task complexity	0.46	1.56	0.6	.77	Gram	2.28
	LLAMA D	-0.68	0.78	-1.72	.39	.	
EI Overall	LLAMA E	0.08	1.36	0.12	.95		
	LLAMA F	0.32	1.32	0.48	.81		
	Task com* LLAMA D	<0.02	0.88	0.04	.99		
	Task com* LLAMA E	0.68	0.82	1.66	.41		
	Task com* LLAMA F	-0.2	0.84	-0.46	.82		

EI_Post ~ 1 + EI_Pre + Task_comp * LLAMA_D + Task_comp * LLAMA_E + Task_comp * LLAMA_F + (1|Participant) + (1|Item) + (1|Grammaticality)

	Intercept	0.3	1.22	0.48	.81	Learner	2.38
	Pretest	1.72	0.24	14	<.01	Item	
	Task complexity	0.16	1.52	0.22	.92		
EI Gramm	LLAMA D	-0.76	0.76	-1.96	.33		
	LLAMA E	<-0.02	1.34	-0.02	1.00		
	LLAMA F	0.22	1.3	0.36	.86		
	Task com* LLAMA D	0.06	086	0.14	.95		
	Task com* LLAMA E	0.6	0.82	1.44	.47		
	Task com* LLAMA F	0.08	0.82	0.18	.93		

EI_Gram_Post ~ 1 + EI_Pre + Task_comp * LLAMA_D + Task_comp * LLAMA_E + Task_comp * LLAMA_F + (1|Participant) + (1|Item)

	Intercept	-6.54	1.96	-6.68	<.01	Learner	3.16
	Pretest	0.66	0.24	5.52	<.01	Item	1

	Task complexity	2.66	0.36	2.24	.26		
EI	LLAMA D	-0.32	1.2	-0.54	.79		
Ungramm	LLAMA E	-0.16	2.02	-0.16	.94		
	LLAMA F	0.04	2	0.04	.98		
	Task com* LLAMA D	-1	1.34	-1.482	.46		
	Task com* LLAMA E	1.12	1.26	1.78	.38		
	Task com* LLAMA F	-0.46	1.24	-0.74	.71		

EI_Ungram_Post ~ 1 + EI_Pre + Task_comp * LLAMA_D + Task_comp * LLAMA_E + Task_comp * LLAMA_F + (1|Participant) + (1|Item)

Note. Gram. = Grammaticality

Appendix B

Interconnection between aptitude and Grammar acquisition in both simple and complex condition

Simple Condition (N = 60)

Result calculate	LAMA B		LAMA D		LAMA E		LAMA F	
	rho	p	rho	p	rho	p	rho	P
Oral Production	-0.034	1.858	-0.352	0.706	-0.18	1.27	-0.152	1.378
Written Production	-0.094	1.608	-0.332	0.76	-0.266	0.97	-0.288	0.896
Elicited Imitation	0.34	0.736	0.314	0.812	-0.224	1.108	-0.214	1.15

Complex Condition (N = 60)

Result calculate	LAMA B		LAMA D		LAMA E		LAMA F	
	rho	p	rho	p	rho	p	rho	P
Oral Production	-0.388	0.606	1.078	0.004	0.582	0.236	-0.038	1.846
Written Production	-0.448	0.468	0.772	0.07	0.91	0.024	-0.012	1.946
Elicited Imitation	-0.856	0.036	0.034	1.856	0.898	0.026	0.458	0.448

أثر توظيف تعقيد المهام مع إعادة الصياغة في تحسين اكتساب قواعد اللغة الإنجليزية لدي المصريين متعلمي اللغه

د/ الهام سويلم

قسم المناهج وطرق التدريس ـ كلية التربية جامعة ـ ٦ اكتوبر

المستخلص

كان الهدف الرئيسي من الدراسة الحالية هو استكشاف تأثير استخدام تعقيد المهام عند اقترانها بإعادة الصياغة علي قواعد اللغة الإنجليزية في تنمية المتعلمين في مصر ، بالإضافة إلى تأثيرها على العلاقة بين الكفاءة وتنمية مهارات متعلمي اللغة بعد إعادة الصياغة. لتحقيق الغرض من الدراسة ، تم تقسيم ١٢٠ من متعلمي FL إلى مجموعتين تجريبيتين: إحداهما أنهت مهام صنع القرار المعقدة ، بينما أنجزت المجموعة الأخرى مهام بسيطة لنقل المعلومات. تم تقديم إعادة الصياغة كتغذية راجعة على أخطاء المشاركين عندما استخدموا قواعد اللغة الانجليزية المتمثله في أشكال الفعل الفردي الحالية بصيغة الشخص الثالث. لجمع البيانات ، استخدم الباحث اختبارات التقليد المستنتجة والإنتاج الشفوي والإنتاج المكتوب لقياس تحسن FL، بالإضافة إلى)اختبار التعرف علي الصوت LLAMA D وأختبار توافق الرمز الصوتي LLAMA E واختبار الاستدلال النحوي LLAMA F) لتقييم كفاءة المتعلمين. أظهرت النتائج أنه كان من المفيد أكثر تطبيق المهام ذات المتطلبات المعرفية الأقل صعوبة. في ظروف المهمة المعقدة ، تم التنبؤ بالتطور من خلال نتائج المشاركين في LLAMA D ، والتي قاست الإنتاج الشفوي والكتابي. أيضا ، تم توقع المكاسب في FL من خلال أداء المتعلمين على LLAMA E ، والذي تم قياسه على التقليد المستحث وكذلك التنبؤ بالتطور من خلال أداء الطلاب في LLAMA F والتي قاست القواعد النحوية.

الكلمات المفتاحية : المتعلمين المصريين الابتدائية FL , تعقيد المهام مع إعادة الإرسال , قواعد اللغة الإنجليزية