



ORBIT - Online Repository of Birkbeck Institutional Theses

Enabling Open Access to Birkbeck's Research Degree output

Scrutinizing the effects of the 4/3/2 activity: repetition, increasing time pressure, accuracy enhancement and cognitive individual differences

<https://eprints.bbk.ac.uk/id/eprint/40440/>

Version: Full Version

Citation: Tran, Mai Ngoc (2019) Scrutinizing the effects of the 4/3/2 activity: repetition, increasing time pressure, accuracy enhancement and cognitive individual differences. [Thesis] (Unpublished)

© 2020 The Author(s)

All material available through ORBIT is protected by intellectual property law, including copyright law.

Any use made of the contents should comply with the relevant law.

[Deposit Guide](#)
Contact: [email](#)

SCRUTINIZING THE EFFECTS OF THE 4/3/2 ACTIVITY: ROLES OF
REPETITION, INCREASING TIME PRESSURE, ACCURACY ENHANCEMENT
AND COGNITIVE INDIVIDUAL DIFFERENCES

Mai Ngoc Tran

A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy
to

Department of Applied Linguistics and Communication
Birkbeck, University of London

2019

Declaration

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Mai Ngoc Tran

ACKNOWLEDGEMENTS

I would not have completed this dissertation without the advice, encouragement and support of many people. First of all, I would like to express my special appreciation to my supervisor, Dr. Kazuya Saito for his prompt and constant assistance. I am impressed by his expertise and devotion to his students. He is always willing to share his deep and profound knowledge with me. When I first started the program, he recommended many research papers which helped me to develop my research ideas. He guided me through all the necessary steps from designing research methods, collecting data, writing and improving my dissertation during the revision. What he did really inspired me to be a better student, a more dedicated teacher and a passionate researcher in the future. Without his tremendous advice, continuous support, and excellent supervision I would never have been successful in obtaining my degree.

Additionally, I am grateful for the support of the examination committee members, Prof. Jean-Marc Dewaele, and Dr. Andrea Révész. Their extremely constructive feedback and valuable suggestions helped me to improve my dissertation and broaden my knowledge.

I am also thankful to the Dean of the Faculty of Foreign Languages, Dr. Nguyen Hoang Tuan who offered me an opportunity to recruit participants in the University of Thu Dau Mot, Vietnam. He also provided me with the best conditions, and support that enabled me to complete my data collection on time. Much appreciation is also extended to the students who volunteered to participate in my study. Their enthusiastic participation allowed me to obtain valuable data and have interesting results.

Much appreciation is also extended to one of my supervisor's friends, Dr. Viktoria Magne, and my friends Yui Suzukida, Hui Sun, and Meltem Ilkan. They assisted me with preparation and practice prior to my presentations at conferences and the viva. Together with my supervisor, they have invested their time to listen to me and give me useful feedback, academic advice, and warm encouragement throughout my doctoral program.

Last but not least, I would also like to express my sincere and deepest gratitude to my family. They not only encouraged me to pursue my PhD degree in the UK but also gave me unconditional love, constant care and countless support which kept me going throughout the three-year journey and contributed to the successful completion of my dissertation.

ABSTRACT

A number of scholars have examined the pedagogical potential and theoretical relevance of adding time pressure to task repetition (i.e., the 4/3/2 activity). However, such fluency enhancement can impact on fluency but not necessarily accuracy aspects of L2 speech (e.g., Thai & Boers, 2016).

To help L2 learners improve both fluency and accuracy, the present study examined the effects of task repetition, fluency enhancement and accuracy enhancement on the development of L2 fluency and accuracy. Furthermore, the study explored the extent to which such gains could be ascribed to learners' cognitive individual differences, operationalized as four different constructs of foreign language aptitude: (a) associative memory, (b) phonemic coding, (c) language analytic ability and (d) sound sequence recognition.

A total of 48 university-level students participated in three 20-minute dyadic sessions. They were randomly divided into four groups: (a) Control; (b) fluency enhancement (FE); (c) accuracy enhancement (AE) and (d) fluency enhancement + accuracy enhancement (FE+AE). Whereas those in the FE and FE+AE groups repeated a monologue task with increasing time pressure (4 → 3 → 2 minutes), those in the AE and FE+AE groups received corrective feedback from the researcher (i.e., accuracy enhancement). After the end of the treatment, all the participants took the LLAMA test (Meara, 2005).

According to the results of statistical analyses, those who engaged in both FE and AE attained significantly more fluent and accurate L2 speech after the treatment at a broad level. However, when AE is introduced to elicit L2 learners' focus on form, certain aspects of their fluency and accuracy development, especially those related

to linguistic encoding (reduction in pauses between clauses and regular past tense forms), remain unchanged.

Finally, the results of the language aptitude test scores suggest the complex relationship between cognitive individual differences, task conditions and L2 fluency and accuracy development.

Key words: Task repetition; corrective feedback; speech production, fluency, accuracy, cognitive individual differences

TABLE OF CONTENT

ACKNOWLEDGEMENTS.....	3
ABSTRACT	5
LIST OF TABLES.....	11
LIST OF FIGURES	13
CHAPTER 1: INTRODUCTION	14
1.1 Background to the Study	14
1.2. Significance of the Study	18
1.3. Thesis Outline	19
CHAPTER 2: LITERATURE REVIEW.....	21
2.1 Task Repetition	21
2.2 Increasing Time Pressure as Fluency Enhancement (4/3/2 Activity)	40
2.3 SUMMARY: Task Repetition and Fluency Enhancement	49
2.4 Corrective Feedback as Accuracy Enhancement	50
2.4.1 CF types.....	52
2.4.2 CF Effectiveness	54
2.4.3 Affecting Factors for CF Effectiveness.....	58
2.4.4 CF Timing	64
2.4.5 SUMMARY of CF Literature.....	79
2.4.6 Gaps to Fill in TBLT and CF Literature	80
2.5 The relationship between fluency and accuracy development.....	82
2.6 Cognitive Individual Differences in Instructed SLA.....	84
2.6.1 What Characterizes Foreign Language Aptitude	85
2.6.2 Aptitude and Treatment Interaction	88
2.6.3 SUMMARY of Aptitude Literature Review.....	108
2.7 Motivation for Current Study	109
2.8 Current Study	112
2.8.1 Research Questions	112
2.8.2 Predictions	114
CHAPTER 3: METHODOLOGY	119

3.1 Participants	119
3.2 Procedure	119
3.3 Content of Treatment.....	122
3.3.1 Task Repetition	122
3.3.2 Fluency Enhancement (Increasing Time Pressure).....	123
3.3.3 Accuracy Enhancement (Corrective Feedback).....	123
3.4 Speech Analyses	128
3.4.1 Materials	128
3.4.2 Measuring Fluency	128
3.4.3. Measuring Accuracy.....	129
3.5 Inter-Coder Reliability.....	130
3.6 Aptitude Measures	131
CHAPTER 4: DATA ANALYSIS.....	134
4.1 Fluency	135
4.1.1 Speed Fluency	135
4.1.2 Breakdown Fluency.....	136
4.1.3 Repair Fluency	139
4.2 Accuracy.....	141
4.2.1 Accuracy in the use of the English past tense.....	141
4.2.2 General accuracy.....	142
4.3 Cognitive Individual Differences (Aptitude Scores).....	146
4.3.1. Aptitude x Control Group.....	149
4.3.2. Aptitude x FE.....	151
4.3.3. Aptitude x AE	153
4.3.4. Aptitude x FE+AE	155
CHAPTER 5: DISCUSSION	158
5.1 Task Repetition	160
5.2 Corrective Feedback (Accuracy Enhancement).....	161
5.3 Incorporating both Fluency and Accuracy Enhancement.....	163
5.4 Roles of Cognitive Individual Differences	165
CHAPTER 6: CONCLUSION	169
6.1 SUMMARY	169
6.2 Limitations and Future Directions.....	172

6.3 Implications	177
References.....	178
Appendices.....	213
Appendix A: Information Sheet and Consent Form for EFL Students	213
Appendix B: Background Questionnaire for Students.....	217
Appendix C: Speaking Tasks.....	220
Appendix D: Examples of Repeated Monologue in the Shrinking Time Condition (4/3/2).....	225
Appendix E: Examples of Repeated Monologue in the Constant Time Condition (3/3/3).....	229
Appendix F: Obligatory Contexts for Past Tense Use.....	233
Appendix G: Intercoder Reliability Analysis	236

LIST OF TABLES

<i>Table 1: Summary of 10 Key Task Repetition Studies</i>	28
<i>Table 2: Summary of 6 Key 4/3/2 Activity Studies</i>	43
<i>Table 3: Summary of 10 Key Corrective Feedback Studies</i>	67
<i>Table 4: Summary of 10 Key Aptitude Studies</i>	90
<i>Table 5: Participant Information by Group</i>	120
<i>Table 6: Group Means, Standard Deviations, Significant Improvement and Effect Size for Articulation Rate</i>	136
<i>Table 7: Group Means, Standard Deviations, Significant Improvement and Effect Size for Mid-clause Pause Ratio</i>	137
<i>Table 8: Group Means, Standard Deviations, Significant Improvement and Effect Size for Final-clause Pause Ratio</i>	138
<i>Table 9: Group Means, Standard Deviations, Significant Improvement and Effect Size for Repair Ratio</i>	139
<i>Table 10: Group Means, Standard Deviations, Significant Improvement and Effect Size for Repetition Ratio</i>	140
<i>Table 11: Group Means, Standard Deviations, Significant Improvement and Effect Size for Targetlike Use of Regular Past Tense</i>	141
<i>Table 12: Group Means, Standard Deviations, Significant Improvement and Effect Size for Targetlike Use of Irregular Past Tense</i>	142
<i>Table 13: Group Means, Standard Deviations, Significant Improvement and Effect Size for Global Accuracy</i>	143
<i>Table 14: Group Means, Standard Deviations, Significant Improvement and Effect Size for Semantic Accuracy</i>	144
<i>Table 15: Group Means, Standard Deviations, Significant Improvement and Effect Size for Morphosyntactic Accuracy</i>	145
<i>Table 16: Descriptive Statistics of Aptitude Scores by Groups</i>	146
<i>Table 17: Descriptive Statistics of Aptitude Scores by Subtests</i>	147
<i>Table 18: Interrelationships between the Four Subtests (D, B, E, F) in LLAMA</i>	148
<i>Table 19: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the Control Group</i>	150
<i>Table 20: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the FE Group</i>	152
<i>Table 21: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the AE Group</i>	154
<i>Table 22: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the FE + AE Group</i>	156

Table 23: Summary of Overall Results of Fluency and Accuracy Analyses
..... 159

LIST OF FIGURES

Figure 1. Levelt's (1989, 1993, 1999) model of speech production 24
Figure 2. A taxonomy of CF strategies 54
Figure 3. Summary of research design: control, FE, AE and FE+AE 127

CHAPTER 1: INTRODUCTION

This chapter provides an overview of the thesis in three sections. The first section introduces background to the present study. The second section is dedicated to the significance of the study. The third section outlines the structure of the thesis.

1.1 Background to the Study

Complexity, accuracy and fluency are the three linguistics subcomponents of L2 performance (Vercellotti, 2017). They have been used as “major research variables” and for “measuring progress in language learning” (Housen & Kuiken, 2009, p.461) in many studies in the field of applied linguistics and second language acquisition (SLA) in general and task-based language teaching (TBLT) in particular. There has been a debate about cognitive limitations in language performance in the literature. Researchers who supported the competition in attentional resources claimed that complexity, fluency and accuracy compete with each other when tasks are high in cognitive demands (Bygate, 1999; Skehan, 1998; Skehan & Foster, 1999, 2008). Therefore, focusing on one of the three dimensions results in the detriment of the other two. The trade-offs could involve accuracy and complexity (Foster & Skehan, 1996) or fluency and accuracy (Michel, Kuiken & Vedder, 2007; Yuan & Ellis, 2003). Robinson (2001a, 2003) held a different view. He demonstrated that attentional resources are non-competitive. Therefore, where “tasks are made increasingly complex simultaneously along dimensions which draw on different resources pools” (Robinson, 2001, p. 307), simultaneous improvement of accuracy and complexity is possible (for more information about Skehan (1998)’s Limited Capacity Hypothesis and Robinson (2001)’s Cognition Hypothesis, see Literature Review in section 2.5. The relationship between fluency and accuracy development).

In general, the aim of L2 learning is to master complexity, accuracy and fluency (Vercellotti, 2017). Research has been carried out to examine the development of these three aspects of L2 performance. While a number of studies focused on the effects of age, instruction, individual differences, and learning context on L2 development (DeKeyser, 2012; Derwing & Rossiter, 2003; Freed, Segalowitz & Dewey, 2004; Muñoz, 2006), TBLT research examined the extent to which the implementation of tasks affects L2 performance (Nation, 1989; Ahmadian, 2012a; Fukuta, 2015; Bui & Huang, 2016; Lambert, Kormos & Minn, 2016; Levkina & Gilabert, 2012; Révész & Han, 2006). The results of these TBLT studies have shown that certain task types or task conditions could lead to the increase in complexity, accuracy or fluency (Révész & Han, 2006; Yuan & Ellis, 2003).

Over the past several decades, a growing number of TBLT studies have examined the facilitative effects of task repetition—a technique in which students will repeat a meaningful task multiple times (Bygate, 1996, 2001, 2009; Gass, Mackey, Fernande, & Alvarez-Torres, 1999; Kim, 2013; Pinter, 2005, 2007a, 2007b; Wang, 2014). It has been suggested in the literature that the repetition of the task provides students with an ideal opportunity to improve their task performance, especially their fluency (de Jong & Perfetti, 2011, Ahmadian, 2013; Fukuta, 2016). This is consistent with Levelt (1989)'s model of speech production (1989) and de Bot (1992)'s adapted model for bilingual production. According to these two models, speech production involves three stages of processing: conceptualization, formulation and articulation. In the first stage, learners plan what to say. In the second stage, they engage in lexical, grammatical and phonological encoding. In the last stage, they turn the encoded content to speech. By repeating a task, students keep what they already

planned in their memory for future use. Therefore, they produced smoother and faster speech (Thai & Boers, 2016).

In recent years, a number of scholars have been deeply interested in the role of increasing time pressure in task repetition (Arevart & Nation, 1991; De Jong & Perfetti, 2011; De Jong, 2012; Maurice, 1983; Nation, 1989; Sato & Lyster, 2012). One well-researched repetition is the 4/3/2 activity (Arevart & Nation, 1991; Boers, 2014; De Jong & Perfetti, 2011; De Jong, 2012; Maurice, 1983; Thai & Boers, 2016) which was developed by Maurice (1983). It is a classroom activity specifically designed for oral fluency development. In this activity, L2 learners are asked to give a monologue on a familiar topic three times with an increasing amount of time pressure. They are given four minutes for the first delivery, three minutes for the second delivery and just two minutes for the third delivery.

To date, the 4/3/2 activity has been widely used in many L2 English classrooms worldwide, since the activity is believed to help learners express their ideas more quickly, more efficiently, and with fewer hesitations and shorter pauses. In general, it has been shown in previous studies that the 4/3/2 activity exerts sizable positive effects on L2 learners' fluency enhancement (e.g., Arevart & Nation, 1991; Boers, 2014; De Jong & Perfetti, 2011; De Jong, 2012; Nation, 1989). At the same time, certain scholars have pointed out that the relationship between the 4/3/2 activity and L2 accuracy development remains unclear (e.g., Thai & Boers, 2016). These results were explained using Skehan & Foster (1997)'s trade-off hypothesis. According to this hypothesis, if students devote much attention to fluency, they fail to attend to other aspects of language performance (accuracy or complexity). Also, these findings are in line with what Robinson (2001a) predicts about tasks that are made complex along dispersing dimensions. The question remaining involves which

modifications should be made to this activity to help students improve both fluency and accuracy. In order to fill this gap in the literature, the current study takes an exploratory approach towards examining how adding accuracy enhancement as a form of delayed corrective feedback can help $N = 48$ Vietnamese learners of English improve both fluency and accuracy aspects of their L2 speech. This is the first attempt to further maximize the pedagogical potential of the 4/3/2 activity by adding corrective feedback.

Corrective feedback (CF) has become a field of interest to SLA researchers for over 40 years. While earlier studies are mostly descriptive (Lyster & Ranta, 1997; Sheen, 2004, Seedhouse, 2004), recent studies are experimental (Ammar & Spada, 2006; Sheen, 2010; Kartchava & Ammar, 2014, Saito & Lyster, 2012; Thomas, 2008). The results of these experimental studies have been consistent in showing that CF has a sizable effect on L2 learning and accuracy development. Previous CF studies have focused on variables that moderate CF effectiveness including CF types, the timing of CF, CF targets, learner variables, research settings, and length of treatment (Carroll & Swain, 1993; Lyster, 2004; Nakata, 2014; Yilmaz, 2012). Interestingly, these studies exclusively examined the facilitative effects of CF on L2 learning. However, few studies have investigated the effectiveness of CF in the context of task repetition. The present study, therefore, extends the scope of CF research by adding CF to the sequence of task repetition.

Also in various SLA studies, foreign language aptitude has been found to be a very important individual difference variable. A broad line of these studies focused on its impacts on the effectiveness of instructional treatment (Li, 2015b). The results of these studies suggested that the same type of instruction benefits learners with different aptitude profiles differently. Therefore, the current study will further explores

whether, to what degree and how learners' gains are associated with their cognitive individual differences, operationalized as four different constructs of foreign language aptitude: (a) associative memory, (b) phonemic coding, (c) language analytic ability and (d) sound sequence recognition.

1.2. Significance of the Study

The findings of the current study will be of interest to not only researchers from various fields including SLA, Applied Linguistics, TBLT, Language Teaching and Learning but also teachers for many reasons. Firstly, while a number of studies have been conducted on the effects of task repetition and time pressure (i.e., the 4/3/2 activity) on fluency development, other studies have been carried out to investigate the impacts of corrective feedback on accuracy development. These are somewhat independent topics in the literature. To my knowledge, this study is the first attempt to add a corrective feedback component to the 4/3/2 activity. Therefore, the major contribution that this study makes to the recent literature on L2 task performance is the findings about the combined effects of task repetition, time pressure and corrective feedback on L2 fluency and accuracy development. Moreover, in terms of outcome measures, while previous research has mostly used broad measures of accuracy, the present study examines accuracy in more detail by including both measures of overall accuracy and accuracy in the use of a specific grammatical structure (i.e., the English past tense). For gauging fluency, this study uses the up-to-date fluency analysis which allows me to see students' speech production processes while engaging in 4/3/2 (or 3/3/3 activity). These are the unique characteristics of the present study that make it novel. In addition, the present study also looks at the relationship between cognitive individual differences and task

conditions and L2 development. The findings, therefore, will provide teachers with valuable information about which tasks to choose and how to best combine them based on students' profiles (i.e., aptitude scores). Furthermore, the obtained results will also have some pedagogical implications which allow teachers to decide which conditions for task repetition are appropriate for which intended outcomes of the activity. Finally, this study draws on Levelt (1989)'s model of speech production, De Bot (1992)'s bilingual production model, and Skehan (1998)'s trade-off hypothesis as well as Robinson (2001)'s cognition hypothesis to predict and interpret the results. Therefore, the findings can make theoretical contributions to the existing literature.

Taken together, the current study is original in different aspects. Firstly, no other studies in the literature have investigated the combined effects of task repetition, time pressure, and corrective feedback on L2 performance. Secondly, the research design and methods employed (including population, setting, and data collection procedure) were unique. Thirdly, the data used in this study had a high degree of originality because it was collected by the researcher and has not been published anywhere. Finally, the implications were based on the findings of this research so they were unique.

1.3. Thesis Outline

This thesis consists of six chapters including (1) Introduction, (2) Literature Review, (3) Methodology, (4) Data Analysis (5) Discussions, and (6) Conclusion.

Chapter 1 (Introduction) introduces the background to the present study, the focus of study, significance and the organization of the study.

Chapter 2 (Literature Review) first covers work that has been done in the fields of task repetition, fluency enhancement, accuracy enhancement and cognitive

individual differences. It then provides the motivation for the present study. The last part of the chapter is dedicated to the research questions and predictions.

Chapter 3 (Methodology) presents the research methodology employed for this empirical work. This chapter includes sections on participants, piloting procedure, data collection procedure, content of the treatment, speech analyses, and inter-coder reliability.

Chapter 4 (Data Analysis) reports the results of the statistical analyses which are presented in three main sections. The first section deals with fluency (speed, breakdown and repair fluency) while the second section focuses on accuracy (accuracy in the use of irregular and regular English past tense and overall accuracy). In the last section, the correlations between the participants' cognitive individual differences (measured via LLAMA: Meara, 2005) and L2 development are reported.

Chapter 5 (Discussion) interprets the results then discusses them in relation to the research questions, predictions and the results of previous research in the field.

Chapter 6 (Conclusion) summarizes the main findings of the study, addresses the contributions of the work and discusses some pedagogical implications. This chapter also includes limitations and recommendations for further study.

CHAPTER 2: LITERATURE REVIEW

This chapter provides a critical review and analysis of the literature directly relevant to the current study. The topics covered here include: (1) task repetition, (2) Increasing time pressure as fluency enhancement (4/3/2 activity), (3) corrective feedback as accuracy enhancement, (4) The relationship between fluency and accuracy development, and (5) cognitive individual differences and language aptitude. The chapter also discusses motivation for the current study. Finally, the research questions and corresponding predictions are introduced.

2.1 Task Repetition

Task repetition, “the repetition of the same or slightly altered task – whether the whole task, or parts of a task” (Bygate & Samuda, 2005, p. 43), facilitates students’ L2 speech development (Bygate, 2001). Bygate (2006) emphasized that repetition here was not referred to as “word-for-word repetition” (p. 167). It involved the repetition of both “content and form” (Bygate, 2006, p.167). A great deal of attention has been directed toward the content or procedure of task repetition. In some studies (i.e., Pinter, 2005), students repeated the same task procedure but with different content (procedural repetition). In other studies (Arevart & Nation, 1991; Ahmadian & Tavakoli, 2010; Ahmadian, 2011, 2012a; Boers, 2014; Lambert et al., 2016; Nation, 1989; Thai & Boers, 2016), students repeated the same task in terms of both content and procedure (task repetition). Some other studies compared the impact of these two different types of task repetition (Bygate, 2001; De Jong & Perfetti, 2011; De Jong, 2012; Gass et al., 1999; Kim & Tracy-Ventura, 2013; Kim & Payant, 2014) on L2 development. Overall, the previous studies on task repetition

have shown that task repetition is particularly effective for developing L2 fluency (Bygate, 2001; De Jong & Perfetti, 2011; De Jong, 2012; Gass et al., 1999).

According to Ellis (2003, 2005, 2008), task repetition was regarded as a type of task planning. He claimed that the first enactment of the task provided students with a chance to prepare for further performances. The beneficial effects of task repetition have been explained by many researchers with reference to Levelt's (1989, 1993, 1999) model of speech production (Ahmadian & Tavakoli, 2010; Ahmadian, 2011, 2012a, 2013; Bygate, 2001; Bygate & Samuda, 2005; Kim & Tracy-Ventura, 2013; Kim & Payant, 2014; Van de Guchte, Braaksma, Rijlaarsdam & Bimmel, 2016; Wang, 2014). According to Levelt's model (see figure 1), when speakers intend to speak, they undergo a process comprised of three different stages: (a) conceptualization (conceptualizing the content of speech), (b) formulation (searching for adequate linguistic forms), and (c) articulation (producing language without too many hesitations and/or pauses). The products of the Conceptualizer are called preverbal messages which are "conceptual structures that can be accepted as input by the Formulator" (Levelt, 1989, p.10). There are two phases in the conceptualization: macro-planning and micro-planning. During the macro-planning phase, the speaker elaborates communicative goals/intentions and chooses what to include in the message, the order in which the information is conveyed and how to present it. Micro-planning is "the speaker's elaboration of a communicative intention by selecting the information whose expression may realize the communicative goals" (Levelt, 1989, p.5). The preverbal messages are then sent to the Formulator where they are converted into a phonetic plan. In the Formulator, grammatical encoding which is "the process by which a message is mapped onto a surface structure" takes place (Levelt, 1989, p.235). The Formulator involves grammatical encoding and

phonological encoding. The necessary information for encoding is retrieved in the lexicon which consists of lemmas (semantic and syntactic properties) and lexemes (or morphological and phonological forms). Here, while the speaker selects semantic and syntactic lexical representations, the corresponding morphological and phonological representations are activated and encoded. After receiving the product of the Formulator (input) in the form of the phonetic plan, the Articulator converts it into overt actual speech. In the model, there is also a speech-comprehension system which is connected with an auditory system. The product of the speech-comprehension system is parsed speech. Another important component in Levelt's model is the monitoring component. According to Levelt, monitoring takes place not only at the Conceptualizer or prior to the Articulation but also at different phases of the speech production processes when necessary. It monitors the speaker's internal speech. Bygate (2001) asserts that task repetition assists language performance precisely because "part of the work of conceptualization, formulation, and articulation carried out on the first occasion is kept in the learners' memory store and can be reused on the second occasion" (p. 29). Also, based on Levelt's model, Yuan & Ellis (2003, p.3) argued that online planning is "the process by which speakers attend carefully to the formulation stage during speech planning and engage in pre-production and post-production monitoring of their speech acts". This line of thoughts is relevant to the present study because in this study, the students in the FE groups were provided with less online planning time for task completion when they repeated their talks. Research from previous online planning studies have found its positive impacts on accuracy development (Yuan & Ellis, 2003; Ahmadian, 2012b; Ahmadian & Tavakoli, 2010)

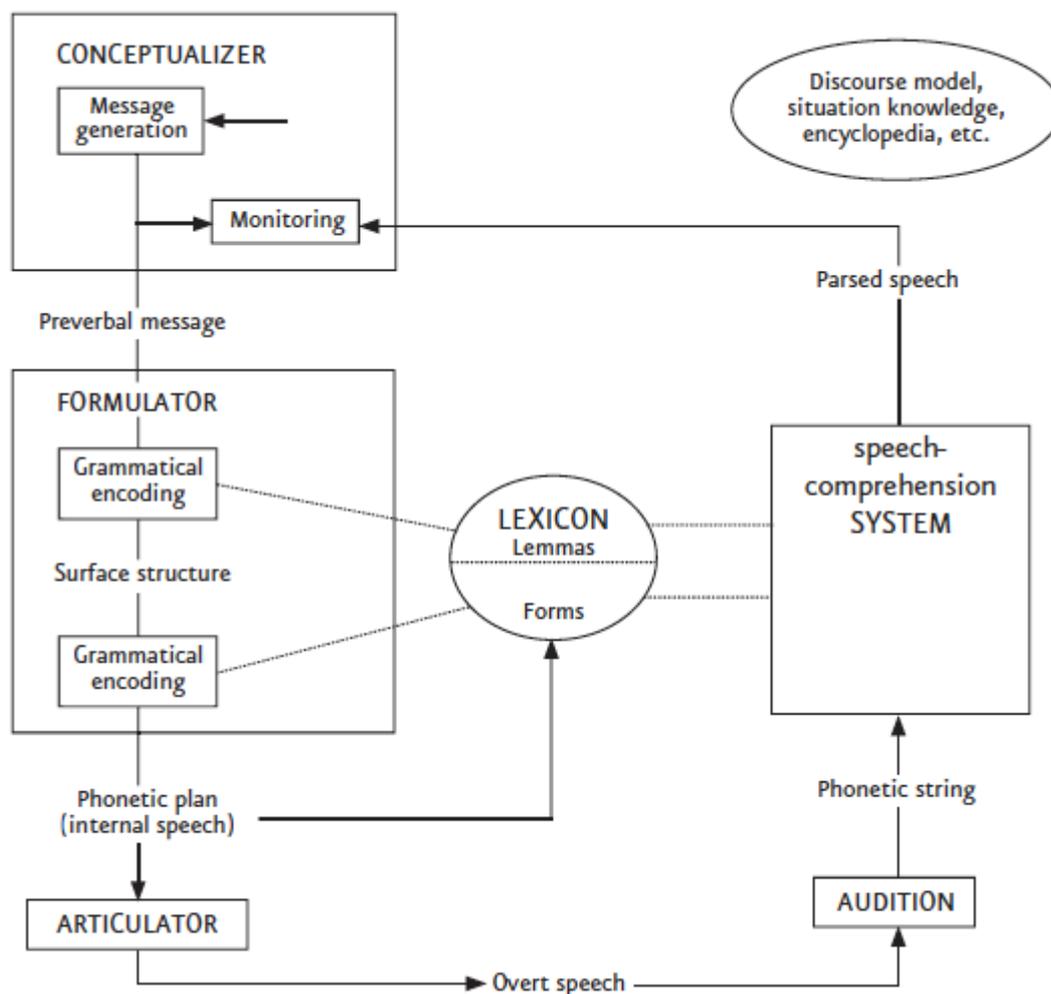


Figure 1. Levelt's (1989, 1993, 1999) model of speech production

De Bot (1992) adapted Levelt's L1 model to bilingual speech production and made only little changes to the original one. While most of the features of the two models are the same, there are some differences between them. Firstly, while Levelt's (1989) believed that all activities of the Conceptualizer are language specific, de Bot assumes that only the ones in the microplanning stage are. Secondly, there are different processing components for different languages in the Formulator but there is only one lexicon where lexical items of both languages are stored together. Finally, there is only one Articulator for both languages which is not language specific.

Several empirical studies on task repetition (Ahmadian & Tavakoli, 2010; Ahmadian, 2011, 2012a, 2013; Arevart & Nation, 1991; Boers, 2014; Bygate, 2001; Bygate & Samuda, 2005; De Jong & Perfetti, 2011; De Jong, 2012; Kim & Tracy-Ventura, 2013; Kim & Payant, 2014; Nation, 1989; Sato & Lyster, 2012; Thai & Boers, 2016; Van de Guchte et al., 2016; Wang, 2014) have been conducted to examine how task repetition affects different aspects of oral performance (i.e. complexity, accuracy and fluency). While many of these studies focused on various variables (i.e. content of repetition, numbers of repetitions, and length of intervals) impacting language development, some other studies have explored whether the facilitative effects of task repetition will carry over to new contexts (see Table 1 for a comprehensive summary of key task repetition studies).

One of the first attempts to investigate the effectiveness of task repetition was Bygate (1996). In his study, a learner narrated a story while watching a short video cartoon. After 3 days, s/he engaged in the same task again. The learner showed significant improvement on fluency and accuracy. Later in 2001, Bygate conducted a larger-scale study involving 48 ESL students. In this study, all the participants carried out two types of tasks (a narrative and an interview) twice with an interval of 10 weeks in between. In addition, the three treatment groups engaged in practice on a fortnightly basis before they repeated the tasks. The analysis revealed significant changes in fluency for the interview task and complexity for both tasks. However, these changes did not carry over to a new context. In addition, no significant differences in accuracy were observed. In terms of transferability, these results were partially in line with the findings of Gass's similar research (Gass et al., 1999).

In Gass et al. (1999), the effects of repetition on accuracy and lexical sophistication of learners' production were examined. A total of 103 learners of

Spanish engaged in the repetition either of the same content (task repetition) or different content (procedural repetition) four times. The findings revealed that task repetition resulted in significant improvement on overall proficiency, accuracy in the use of *estar*, and lexical complexity. However, these results did not carry over to the new version of the task with different content.

Another study that compared the effects of task repetition and procedural repetition on L2 development was Kim and Tracy-Ventura (2013). Thirty six ESL students were assigned to one same content group, and one different content group. Both groups repeated the tasks three times with one day of interval in between. The findings showed that both types of repetition promoted gains in the use of task-induced linguistic features but only procedural repetition resulted in syntactic complexity development. Concerning speech rate, no changes were found for any groups.

Similarly, Kim and Payant (2014), compared the effects of task repetition (repeating both the procedure and the content) and procedural repetition (repeating the procedure with different content) under two different task conditions (simple vs. complex). In this study, task complexity was operationalized as the presence and absence of reasoning demands. A total of 92 Korean junior high school female participants repeated a task on three occasions. They were randomly assigned to four treatment conditions. Group 1 (simple/task repetition) repeated a simple task with the same content. Group 2 repeated a simple task with different content (simple/procedural repetition). Group 3 (complex/task repetition) repeated a complex task with the same content. Group 4 (complex/procedural repetition) repeated a complex task with different content. The results of this study suggested that task complexity did not significantly affect the occurrence of lexical and grammatical

language-related episodes but task repetition did. More precisely, students in the procedural repetition groups produced significantly more grammatical and lexical language-related episodes.

One of the most recent task repetition studies, conducted by Lambert et al. (2016), involved immediate repetition of tasks. In their study, 32 Japanese EFL learners in a public university in Japan engaged in each of the three oral communication tasks (instruction, narration and opinion) 6 times. This study specifically focused on the effects of task repetition on diverse dimensions of L2 fluency (i.e. speed, breakdown, and repair). Lambert et al. (2016) observed that task repetition was uniquely related to certain stages of the speech production process differently. Students' speech rate improved most markedly over the first three performances. However, the decrease in final-clause pauses happened until the second performance, the ratio of mid-clause pauses declined up to the fourth performance, and self-repairs decreased after the fourth performance.

For a comprehensive summary of more task repetition studies, see Table 1.

Table 1: Summary of 10 Key Task Repetition Studies

Bygate (1996)	
Participants	1 English language learner
Instructional treatment	The student watched a video cartoon and retold the story. Two days later, she repeated the same task.
Number and interval of repetitions	2 repetitions (interval: 3 days)
Outcome measures	<u>Task</u> - Oral narrative task (monologic) - The analysis focused on complexity, fluency and accuracy.
Findings	Task repetition had a significant effect on fluency and accuracy of students' performances.
Gass et al., (1999)	
Participants	103 native speakers of English studying Spanish as a second language
Instructional treatment	Length: approximately 2 weeks

While watching Mr. Bean videos students simultaneously told stories in Spanish. They talked about what they saw as it was happening.

- same content group (students watched the same video 3 times and a new video at time 4, n = 32)
- different content group (students watched different videos each time and a new video at time 4, n = 33)
- control group (students saw videos only at Time 1 and Time 4, n = 38)

Number and interval of repetitions 4 repetitions (interval: 2-3 days)

Outcome measures Task
Telling stories

Judgement:

2 native speakers of Spanish were asked to judge whether participants demonstrated “better Spanish” in Time 3 and Time 4 than in Time 1. The analysis focused on:

- holistic change across testing sessions
 - change in the area of morphosyntax
 - lexical change
-

- Findings
- Task repetition had a significant effect on overall proficiency, accuracy in the use of *estar*, and lexical sophistication.
 - These improvements did not carry over to a new context.

Lynch & Maclean (2000)

Participants	14 students from six European countries learning English for specific purposes
Instructional treatment	<p>Length: approximately 90 minutes</p> <p>Participants worked in pairs making a poster based on a research article. One participant (A) answered visitors' questions. One student (B) visited the posters of other pairs and asked questions. When A participants came back, they received questions and B participants went visiting other posters.</p>
Number and interval of repetitions	6 repetitions (interval: immediate repetition)
Outcome measures	<p><u>Task</u></p> <p>A poster carousel</p> <p>14 sets of six interactions between two hosts and 12 visitors were analysed. The analysis focused on:</p> <ul style="list-style-type: none"> - Subject-verb structures

- Lexico-grammatical accuracy and performance
- Pronunciation
- Explanation of a complex concept

- Findings
- Task repetition helped learners develop different areas of their interlanguage.
 - They produced more fluent and accurate speech.

Bygate (2001)

Participants 48 non-native speakers of English studying in the UK

Instructional Length: 10 weeks

treatment All the participants were given one narrative and one interview task at Time 1.

- After 2 weeks, participants in the narrative group were given 2 narrative tasks on a fortnightly basis at Time 2, 3, and 4.
- After 2 weeks, participants in the interview group were given 2 interview tasks on a fortnightly basis at Time 2, 3, and 4.
- Participants in the control group received no treatment

At the tenth week, all the participants were given two narrative tasks and two interview tasks. One of the narratives and one of the interviews are the same as the ones they had undertaken at Time 1.

2 repetitions (interval: 10 weeks)

Number and
interval of
repetitions

Outcome
measures

Task

Narratives and interviews

The analysis focused on complexity (the number of words per t-unit), fluency (the number of unfilled pauses per t-unit) and accuracy (error-free t-units).

Findings

- Task repetition resulted in greater fluency and complexity, but not accuracy.
- These improvements did not carry over to a new task.

Pinter (2005)

Participants

10 pairs of 10-year-old Hungarian children

Instructional
treatment

Length: 4 weeks

The children worked in pairs carrying out two information gap tasks. After 3 to 4 days, they repeated similar versions of the tasks.

Number and
interval of
repetitions

3 repetitions (interval: 3 to 4 days)

Outcome	<u>Task</u>
measures	<ul style="list-style-type: none"> - Spot the differences - Follow the route on the map <p>The analysis focused on fluency development (children's pace on the tasks, the amount of silence, the total number of words delivered, and speech rate).</p>
Findings	<ul style="list-style-type: none"> - Students' pace on the tasks increased (less amount of silence, less time and language used). - Speech rate increased.

Kim & Tracy-Ventura (2013)

Participants	36 female Korean junior high school students
Instructional treatment	<p>Length: 4 weeks</p> <ul style="list-style-type: none"> - task repetition (students repeated the same information-exchange task procedure with the same content three times, n = 18) - procedure repetition (students repeated the same information-exchange task procedure with different content three times, n = 18) <p>3 repetitions (interval: 1 day)</p>

Number and
interval of
repetitions

Outcome

Task

measures

Monologic picture-based oral narrative task

Students' speeches were measured in terms of complexity (syntactic and lexical complexity), accuracy (error-free AS-units, clauses, and simple past verbs), and fluency (number of syllables and reformulations per minute).

Findings

- Students engaged in procedural repetition showed significant syntactic complexity development.
- Both groups showed significant accuracy improvement in the use of simple past tense.
- No significant improvements in speech rate were found for any groups.

Ahmadian (2013)

Participants

42 Iranian intermediate EFL learners

Instructional
treatment

All the participants took a working memory test. After that, they engaged in an oral narrative task twice with an interval of two weeks.

Number and interval of repetitions 2 repetitions (interval: 2 weeks)

Outcome measures Task
 An oral narrative task
 Students' speeches were measured in terms of complexity (syntactic complexity and variety), accuracy (error free clauses and verb forms), and fluency (number of and meaningful syllables produced).

Findings - Participants with greater working memory capacity showed more improvement in fluency and accuracy when they repeated the task.

Lambert et al. (2016)

Participants 32 Japanese learners of English at three proficiency levels of high, mid and low.

Instructional treatment Length: 90 minutes
 Students performed three oral communication tasks (instruction, narration and opinion), each task six times.

6 repetitions (interval: Immediate repetition)

Number and
interval of
repetitions

Outcome

Task

measures

Three communication tasks: instruction, narration, and opinion
Students' fluency (speed, breakdown, and repair) was
measured.

Findings

- Task repetition is beneficial to fluency development regardless of proficiency level or task type.
- Students' speech rate improved most significantly over the first three performances.
- Final-clause pauses decreased until the second performance.
- Mid-clause pauses decreased up to the fourth performance.
- Self-repairs decreased after the fourth performance.

Fukuta (2016)

Participants

28 Japanese students learning English as a foreign language

Instructional
treatment

Length: Approximately 10 minutes

- All participants narrated six-frame cartoons. After one week,
- The experimental group performed the same narrative task again.

- The comparison group performed the same task but with a different picture.

Number and interval of repetitions 2 repetitions (interval: 1 week)

Outcome measures Task
Narrative tasks of six-frame cartoons
Students' speeches were measured in terms of complexity (syntactic complexity and lexical variety), accuracy (error-free AS-units), and fluency (pruned words per minute).

Findings - The students focused more on the syntactic encoding process and less on the conceptualizing process when they repeated the same content the second time.

Van de Guchte et al. (2016)

Participants 48 ninth-grade students learning German as a foreign language

Instructional treatment Length: 9 weeks
There were 2 interventions, each of which focused on a different target structure and was spread out 3 weeks. For each intervention, both groups performed the main task, and received

form-focused feedback on the target structures. Two weeks later,

- repetition group repeated a similar task after two weeks (n = 24)
- no-repetition group performed a filler task (n = 24)

Number and interval of repetitions 2 repetitions (interval: 2 weeks)

Outcome measures Task

- Metalinguistic knowledge tests
- Written accuracy tests
- Oral accuracy tests
- Oral fluency tests

Findings

- The repetition group outperformed the no-repetition group on written accuracy and metalinguistic knowledge.
- There were no significant differences between conditions on oral fluency and accuracy.

Overall, a large and growing number of previous studies have been conducted and reported the beneficial effects of task repetition on diverse aspects of L2 oral performance including overall proficiency (Gass et al., 1999), complexity (Bygate, 2001; Gass et al., 1999; Kim & Tracy-Ventura, 2013; Wang, 2014), accuracy (Bygate, 1996; Gass et al., 1999; Kim & Tracy-Ventura, 2013; Lynch & Maclean, 2000, 2001; Van de Guchte et al., 2016; Wang, 2014) and fluency (Bygate, 1996, 2001; Lambert et al., 2016; Lynch & Maclean, 2000, 2001; Pinter, 2005; Sample & Michel, 2014; Wang, 2014).

With regard to the relation between task repetition and L2 speech processing, Lambert et al (2016) suggested that the first repetition was considered to be connected to conceptualizing stage and help students conceptualize the content while the subsequent task repetitions after the fourth performance supported learners in linguistic encoding. In terms of transferability, results from task repetition studies have shown that the gains achieved during task repetition did not transfer to speeches about new topics (Gass et al., 1999; Bygate, 2001). When task complexity was concerned, various task repetition studies (i.e., Kim & Payant, 2014) found that complex tasks did not promote greater language-related episodes (i.e., more focus on form).

In this current study, I will focus on one specific kind of task repetition activity—i.e., 4/3/2. In what follows, I will carefully outline pedagogical/theoretical underpinnings of the activity and previous empirical studies to provide readers with a better understanding of the topic.

2.2 Increasing Time Pressure as Fluency Enhancement (4/3/2 Activity)

The 4/3/2 activity (Maurice, 1983) is one unique kind of task repetition, whereby students repeat a monologue task three times, but with an increasing amount of time pressure (4 → 3 → 2 minutes). According to Nation (1989), this activity has two important pedagogical features:

- Students repeat the same monologue three times, which reduces the need for planning and promotes their confidence in speaking; and
- The time allowed is reduced each time the talk is delivered, which increases the time pressure believed to support fluency *enhancement*.

There have been a number of empirical studies on the effectiveness of the 4/3/2 activity on L2 speech learning (e.g., Arevart & Nation, 1991; Boers, 2014; De Jong & Perfetti, 2011; De Jong, 2012; Nation, 1989; Thai & Boers, 2016). First, Nation (1989) conducted a study to examine the complexity, accuracy, and fluency aspects of L2 speech among six advanced adult ESL learners who engaged in the 4/3/2 activity. Comparing the participants' first and third speeches, Nation found significant gains especially in their fluency.

Arevart and Nation (1991) replicated Nation's (1989) study with 20 intermediate ESL learners in New Zealand. The results of their study were in agreement with those of the original study in indicating that the 4/3/2 technique allowed learners to reach a higher level of fluency. More specifically, in the third delivery of their talks, the number of hesitations reduced significantly. In addition, their speech rate was much faster. However, both Nation (1989) and Arevart and Nation (1991) reported the effects of task repetition within the 4/3/2 sequence. It is

worth emphasizing that these two studies made no attempt to examine the long-term development.

More recently, De Jong and Perfetti (2011) illustrated in depth the relationship between the 4/3/2 activity and L2 fluency development, controlling for the content of task repetition as one independent variable. While most of previous studies measured the immediate effects of the 4/3/2 activity, De Jong and Perfetti's (2011) study was conducted over a longer period. Moreover, they included both immediate posttests and delayed posttests. In their laboratory study, 24 high intermediate-level ESL students at a university in the United States were divided into two groups. Whereas the repetition group gave a monologue on the same topic three times during the 4/3/2 activity, the non-repetition group talked about three different topics during each delivery. According to the results, the repetition group not only significantly improved their fluency over time, but also successfully transferred such gains to the contexts of new topics. In contrast, the non-repetition group's fluency was clearly observed only when their performance was tested via trained topics.

It is crucial to note that in these three studies (Nation, 1989; Arevart & Nation, 1991; De Jong & Perfetti, 2011), although the 4/3/2 task was reported to have a clear positive effect on fluency development, it remained unclear which aspects of the treatment—task repetition or time pressure—could be beneficial for the development of fluency.

De Jong (2012) aimed to provide some tentative answers to this question by examining the effects of these two variables (task repetition and time pressure) separately. A total of 32 ESL learners involved in their study were assigned to two different treatment conditions. In the decreasing time condition, learners were given 180, 135, and 90 seconds, respectively to complete the task three times while in the

constant time condition, learners were given 135 seconds each time. The former condition was found to promote gains in fluency significantly while no significant improvement was observed for complexity in either condition.

Particularly relevant to the current study is Boers's research (Boers, 2014; Thai & Boers, 2016). In these seminal studies, the researchers focused on the importance of time pressure by comparing the differential impact of the increasing pressure (4/3/2) and constant pressure (3/3/3) activities on L2 speech development. Interestingly, whereas the 4/3/2 group demonstrated more robust gains in their fluency than the 3/3/3 group did, no significant change was observed among the participants' accuracy and complexity aspects of L2 speech. The results suggest that the increasing time pressure (i.e., 4/3/2) could be instrumental to the development of L2 fluency rather than L2 accuracy, since the time pressure variable induces L2 learners to prioritize fluency over accuracy in L2 speech. See Table 2 for a summary of key studies on 4/3/2 activity.

Table 2: Summary of 6 Key 4/3/2 Activity Studies

Nation (1989)	
Participants	Six advanced adult ESL learners
Instructional treatment	Length: 1 session Students talked about an interesting event happened to them three times under increasing time pressure.
Outcome measures	<u>Task</u> : 4/3/2 activity The participants' first and third speeches were measured in terms of fluency (speed rate, hesitations, repetitions, and false start), grammatical accuracy, and control of the content.
Findings	- Significant gains in fluency were found (the rate of speaking increased, the numbers of false starts, repetition and hesitations reduced). - No significant improvement in accuracy was detected.
Arevart & Nation (1991)	
Participants	20 ESL learners of intermediate proficiency level and various first language backgrounds taking an English proficiency course in a university in New Zealand
Instructional treatment	Length: 1 session

Students told the same story three times under increasing time pressure.

Outcome	<u>Task</u> : 4/3/2 activity
measures	The participants' first, second and third speeches were measured in terms of speaking speed, and number of hesitations.
Findings	- Students delivered their speeches faster with fewer hesitations in the second and third deliveries.

De Jong & Perfetti (2011)

Participants	24 ESL students of a university in the United States of America
Instructional treatment	Length: 3 sessions over a period of 2 weeks - repetition group: gave a monologue on the same topic three times during the 4/3/2 activity - no repetition group: performed the same task but talked about three different topics
Outcome measures	<u>Task</u> : 2-min monologues Students' speeches were measured in terms of fluency (mean length of fluent runs, mean length of pauses, phonation/ time ratio and articulation rate).

- Findings
- Fluency increased in both groups between the pre-test and post-test.
 - This increase in fluency only transferred to a speech about a new topic when the students had repeated their speeches in the 4/3/2 training.

De Jong (2012)

Participants 32 adult ESL learners in the United States

Instructional Length:
treatment Participants were provided with a wordless six-panel picture story and had to narrate the story three times. However, they were assigned to either of the two conditions:

- decreasing time condition: Participants were given 180, 135, and 90 seconds, respectively to narrate the story three times (n = 15)
- constant time condition: Participants also narrated the story three times but they were given 135 seconds each time (n = 17)

Outcome Task: 180/135/90 activity and 135/135/135 activity

measures

Students' speeches were measured in terms of complexity (Phrasal complexity, and subordination), and fluency (phonation/time ratio, number of reformulated words per 100 words, and articulation rate).

- Findings
- Phonation/time ratio improved across deliveries only in the decreasing time condition.
 - Complexity did not increase in either condition.

Boers (2014)

- Participants 10 adult ESL learners of varying levels of proficiency and various first language backgrounds studying at a University in New Zealand
- Instructional treatment Length: 1 session
Students talked about each of the two different topics three times either under increasing time pressure (i.e. 4/3/2) or under constant time condition (i.e. 3/3/3)
- Outcome measures Task: 4/3/2 activity, and 3/3/3 activity
The participants' first and third speeches were measured in terms of complexity (lexical sophistication and subordination), accuracy, and fluency (the mean number of dysfluencies, speech rate).
- Findings
- Both task conditions resulted in fluency improvement.
 - Participants' speech rate increased more in the shrinking time condition.
 - There was a lack of improvement in complexity and accuracy.
-

Thai & Boers (2016)

Participants	Twenty 10-grade EFL students at a high school in Vietnam
Instructional treatment	Length: 1 session - 4/3/2 group: talked about one topic three times under increasing time pressure - 3/3/3 group: talked about one topic three times under constant time condition
Outcome measures	<u>Task</u> : 3/2/1 activity, and 2/2/2 activity The participants' first, second and third speeches were measured in terms of complexity (syntactic and lexical complexity), accuracy and fluency (speech rate, and nonphonation/time ratio).
Findings	- Fluency was enhanced more markedly in the shrinking time condition than in the constant time condition. - No significant improvement was found for complexity and accuracy.

To date, research has shown that 4/3/2 can be highly effective for L2 fluency development (Arevart & Nation, 1991; Boers, 2014; De Jong & Perfetti, 2011; De Jong, 2012; Nation, 1989; Thai & Boers, 2016); however, little is known about what kinds of remedial techniques can best help L2 learners improve both fluency and accuracy when doing the 4/3/2 activity. More research is needed to determine to what degree and how enhancing L2 learners' awareness towards accuracy could be an effective strategy, especially when it is integrated into the context of task repetition. In this regard, the following two studies provide some crucial implications

Hawkes (2011) introduced conscious-raising activities before task repetition. 13-14- year old Japanese EFL students participate in this study. According to the results of the interviews, it was shown that Japanese EFL students were guided to pay attention to their accurate use of pronunciation, grammar and vocabulary forms while using language for meaning under task repetition conditions. Based on the findings, Hawkes (2011) claimed that during the repeated performance, more attention was paid to form. In Van de Guchte et al. (2016)'s experimental study, 48 ninth-grade Dutch participants learning German as a foreign language first received explicit metalinguistic information on the accurate use of target grammatical structures (the German dative case after a preposition and the German comparatives). Subsequently, the participants were divided into two group conditions (Task Repetition, No Repetition). According to the results of pre- and posttests, the repetition group outperformed the non-repetition group especially in terms of written accuracy. This suggests that the awareness activity was successful in drawing learners' attention to form to some extent.

2.3 SUMMARY: Task Repetition and Fluency Enhancement

Taken together, the aforementioned literature provides four broad conclusions. First, repeating the same task multiple times significantly impacts L2 learners' fluency in particular (e.g., Bygate, 1996, 2001; Lambert et al., 2016; Lynch & Maclean, 2000, 2001; Pinter, 2005; Sato & Lyster, 2012). Second, if such task repetition is carried out under increasing time pressure conditions (4/3/2), it could further enhance the transferability, generalizability and robustness of L2 fluency development (e.g., De Jong & Perfetti, 2011; De Jong, 2012; Thai & Boers, 2016). Third, the 4/3/2 activity benefits fluency but not accuracy aspects of L2 development, since it induces learners to prioritize fluency over accuracy (Boers, 2014; Thai & Boers, 2016). Fourth, there is an indication that some awareness activities may be needed in the 4/3/2 activity in order to lead to both fluency and accuracy development (Hawkes, 2011; Van de Guchte et al., 2016).

For these reasons, in this study, I intend to introduce corrective feedback as accuracy enhancement during the 4/3/2 activity and scrutinize the differential effects of task repetition, time pressure, and CF on L2 fluency and accuracy development. To my knowledge, this is the first attempt to increase the pedagogical potential of 4/3/2 activity through "accuracy" enhancement as a form of providing CF to draw learners' attention to "accuracy" aspects of language while they are working on "fluency" development. In the next section, I will provide an overview on L2 research on CF which has evolved somehow "independently" from the TBLT literature.

2.4 Corrective Feedback as Accuracy Enhancement

Corrective feedback has been an area of interest to both second language teachers and researchers since the inception of the field of SLA (Nassaji, 2016). The term CF is defined as “the form of responses to learner utterances containing an error” (Ellis, 2006, p. 28). Whereas the main function of CF is to signal that learners have made errors (i.e., negative evidence), different types of CF can include target forms (i.e., positive evidence) with a different degree of explicitness (e.g., less explicit recasts to more explicit metalinguistic information).

From a theoretical perspective, CF is believed to be effective for L2 learning, especially when it is provided during communicatively authentic tasks, since it encourages learners to notice, understand and acquire form while maintaining their primary focus on meaning (Ellis, 2016; Long, 2007a; Mackey, 2012). The proponents of the Skill Acquisition Theory also argue that communicative focus on form via CF can promote learners’ gradual transition from controlled to automatic use of language (DeKeyser, 2012). To date, there have been ample studies evidencing the effect of CF on L2 accuracy development in the context of grammar learning (e.g., Lyster, Saito, & Sato, 2013 for a narrative review; Li, 2010 for a meta-analytic review).

A large volume of published studies on CF have focused on different CF types and their effectiveness (Ammar & Spada, 2006; Baralt, 2013; Egi, 2007a, 2007b, 2010; Ellis, 2007; Ellis, Loewen & Erlam, 2006; Loewen & Erlam, 2006; Goo, 2012, 2016; Gooch, Saito & Lyster, 2016; Li, 2013; Loewen & Nabei, 2007; Kartchava & Ammar, 2014; Lyster, 2004; Lyster & Ranta, 1997; Mackey & Philp, 1998; Révész, 2009; Révész & Han, 2006; Sato & Lyster, 2012; Sheen, 2010; Thomas, 2018; Yang & Lyster, 2010; Yilmaz, 2012, 2013; Yilmaz & Yuksel, 2011; Van de Guchte et al.,

2016). While Lyster and Ranta (1997) introduced six different types of CF, some other studies made a distinction between two broad categories of CF including implicit and explicit CF (Carroll & Swain, 1993; Ellis et al., 2006; Li, 2013; Yilmaz, 2013). Whereas many studies explored the efficacy of the most frequently used type of CF which is recasts (Baralt, 2013; Doughty & Varela, 1998; Egi, 2007a, 2007b, 2010; Ellis & Sheen, 2006; Han, 2002; Ishida, 2004; Leeman, 2003; Loewen, & Philp, 2006; Long, Inagaki & Ortega, 1998; Lyster, 1998b; Mackey & Philp, 1998; Mackey, Gass & McDonough, 2000; McDonough & Mackey, 2006; Révész, 2009; Révész & Han, 2006; Sheen, 2006; Carpenter, MacGregor & Mackey, 2006; Yilmaz & Yuksel, 2011), other studies compared the benefits of recasts with other types of CF (Ammar, 2008; Ammar & Spada, 2006; Carroll & Swain, 1993; Dilans, 2010; Ellis et al., 2006; Ellis, 2007; Goo, 2012; Gooch et al., 2016; Kartchava & Ammar, 2014; Li, 2013; Loewen & Nabei, 2007; Loewen & Philp, 2006; Lyster, 2004; Lyster & Izquierdo, 2009; Lyster & Mori, 2006; Ranta & Lyster, 2007; Sato & Lyster, 2012; Sheen, 2010; Sheen, 2007a; Thomas, 2018; Yang & Lyster, 2010; Yilmaz, 2012, 2013).

Another line of research is concerned with variables mediating the effectiveness of CF such as CF saliency (Egi, 2007a; Ellis & Sheen, 2006; Leeman, 2003; Loewen & Philp, 2006; Philp, 2003; Sheen, 2006; Yilmaz & Yuksel, 2011), CF implicitness (Egi, 2007a; Ellis & Sheen, 2006; Leeman, 2003; Loewen & Philp, 2006; Philp, 2003; Sheen, 2006), CF linguistic targets (Ellis, 2007; Ishida, 2004; Iwashita, 2003; Kartchava & Ammar, 2014; Leeman, 2003; Yang & Lyster, 2010; Yilmaz, 2012), content familiarity (Révész & Han, 2006), task types (Révész & Han, 2006), learners' levels (Carroll, Swain, & Roberge, 1992; Iwashita, 2001; Mackey & Philp, 1998; Philp, 2003; Lyster, 2004), and language aptitude (Benson & DeKeyser, 2018;

Mackey, Adams, Stafford, & Winke, 2010; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Miyake & Friedman, 1998; Sheen, 2007a, 2007b; Sagarra, 2007; Trofimovich, Ammar, & Gatbonton, 2007; Yilmaz & Granena, 2016). In addition, the timing of CF (i.e. whether correction happens immediately during communicative activities or after the completion of the tasks) has also been a subject of interest for researchers (i.e. Li, Zhu, & Ellis, 2016; Nakata, 2015; Nassaji, 2007, 2011; Rolin-Ianziti, 2010).

2.4.1 CF types

An extensive body of research has empirically examined the efficacy of different types of feedback for L2 development (Adams, Nuevo & Egi, 2011; Ammar & Spada, 2006; Benson & DeKeyser, 2018; Carroll et al, 1992; Carroll & Swain, 1993; Dilans, 2010; Ellis, 2007, 2009; Ellis et al., 2006; Fu & Nassaji, 2016; Fujii, Ziegler & Mackey, 2016; Goo, 2012, 2016; Gooch et al., 2016; Iwashita, 2003; Kartchava & Ammar, 2014; Li, 2010, 2013; Li et al., 2016; Loewen & Nabei, 2007; Lyster, 2004; Lyster & Mori, 2006; Lyster & Ranta, 1997; Mackey, 2006, 2012; Mackey & Philp, 1998; Nassaji, 2009, 2011; Nataka, 2014; Révész 2009; Révész & Han, 2006; Saito & Lyster, 2012a, 2012b; Sato & Lyster, 2007, 2012; Sauro, 2009; Sheen, 2004, 2007a, 2007b, 2010; Shintani & Ellis, 2015; Thomas, 2018; Yang & Lyster, 2010; Yilmaz, 2012, 2013; Yilmaz & Granena, 2016; Yilmaz & Yuksel, 2011). In Lyster and Ranta's (1997) descriptive study, they identified six types of CF including recasts, explicit correction, elicitation, metalinguistic cues, clarification requests, and repetition. In this study, Lyster and Ranta (1997) examined the frequency and distribution of these types of feedback used by four teachers in meaning-focused L2 classrooms in a primary school in Montreal, Canada. They reported that there was a tendency for teachers to rely heavily on recasts. However, recasts were found to be the least effective in eliciting uptake and repair. Ranta and

Lyster (2007) subsequently grouped these CF types into two broad categories: reformulations and prompts. Reformulations include “recasts and explicit correction because both these moves supply learners with target reformulations of their non-target output” (Ranta & Lyster, 2007, p. 152). Prompts (i.e. elicitation, metalinguistic cues, clarification requests, and repetition) include “a variety of signals, other than alternative reformulations, that push learners to self-repair” (Ranta & Lyster, 2007, p. 152).

CF types can also be distinguished in terms of the degree of their explicitness (Carroll & Swain, 1993; Ellis et al., 2006). According to Carroll and Swain (1993), explicit CF would overtly state that an error has been made and include direct correction. In contrast, implicit CF “refers to any corrective move that does not overtly inform the learner of the unacceptability of his/her erroneous production” (Li, 2010, p. 337). Li (2010) characterized explicit correction and metalinguistic feedback as explicit techniques and considered recasts, clarification requests, elicitation, and repetition as implicit feedback types. Besides, CF can also be input-providing when the correct forms were provided to learners or output prompting when a correction was elicited from the learners (Ellis, 2009; Sheen & Ellis, 2011).

Lyster and Saito (2010) and Lyster et al. (2013) introduced a taxonomy of CF strategies in which a distinction can be made between not only explicit and implicit CF but also reformulations and prompts (See Figure 2). As being illustrated in the taxonomy, prompts (output eliciting) include five different strategies such as clarification requests, repetition, paralinguistic signal, elicitation and metalinguistic clue while reformulations (input providing) include recasts, explicit correction and explicit correction with metalinguistic explanation. These CF types range from

implicit to explicit along a continuum. Furthermore, conversational recast can also be distinguished from didactic recast. The former is more implicit than the latter.

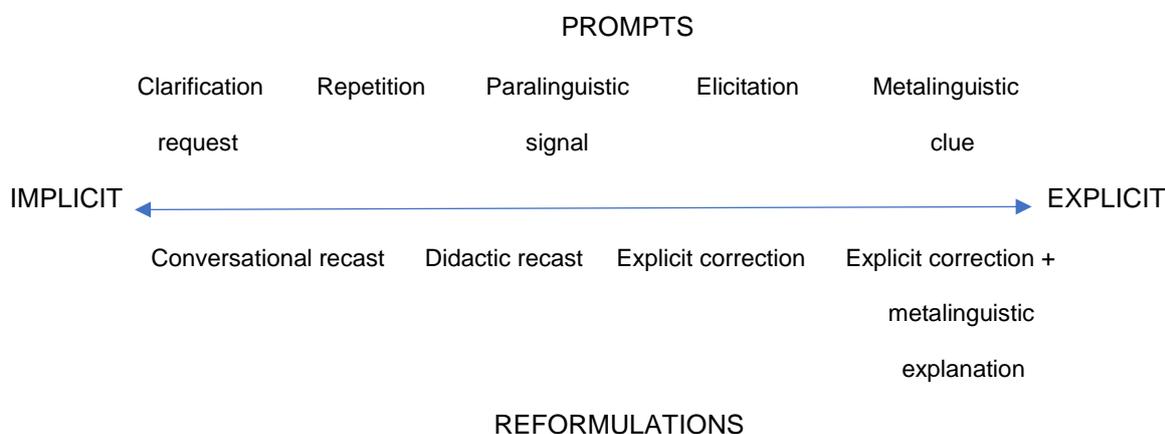


Figure 2. A taxonomy of CF strategies (Lyster & Saito, 2010, p. 278; Lyster et al., 2013, p.5)

2.4.2 CF Effectiveness

The role of CF in L2 grammar acquisition have been the focus of a considerable number of empirical studies carried out in various classroom and laboratory contexts (for reviews see Brown, 2016; Li, 2010; Lyster & Saito, 2010a, 2010b; Lyster et al., 2013; Mackey & Goo, 2007; Nassaji, 2016; Russell & Spada, 2006).

According to Doughty and Varela (1998), CF is an effective focus-on form technique because it draws learners' attention to linguistics form while they are performing a communicative task. Carroll and Swain (1993) conducted a study involved 100 adult ESL Spanish learners in Toronto, Canada to investigate the effects of explicit and implicit forms of feedback on the use of English dative

alternation. According to the results of the analysis, all the treatment groups performed significantly better than the comparison group who did not receive feedback. In addition, the group that received explicit metalinguistic feedback significantly outperformed all the other groups.

Similar results have been obtained in Ellis et al. (2006). In their study, Ellis et al. (2006) reported the effects of metalinguistic feedback (explicit CF) and recasts (implicit CF) on the acquisition of English past tense -ed. A total of 34 adult low-intermediate ESL learners studying at a private language school in New Zealand were divided into two experimental groups and a control group. The two experimental groups retold stories based on picture sequences and received either explicit or implicit CF depending on the condition in response to their erroneous utterances while the control group did not complete the tasks or receive any feedback. Ellis et al. (2006) found that the metalinguistic feedback group was superior to the recast group and the control group.

Apparently, of the different types of CF, recasts have received the most amount of research attention in the literature (Ammar & Spada, 2006; Baralt, 2013; Braidı, 2002; Carpenter et al., 2006; Doughty & Varela, 1998; Egi, 2007a, 2007b; Ellis & Sheen, 2006; Ellis et al., 2006; Goo, 2012; Goo & Mackey, 2013; Gooch et al., 2016; Han, 2002; Hawkes & Nassaji, 2016; Ishida, 2004; Kartchava & Ammar, 2014; Leeman, 2003; Li, 2013; Loewen, & Philp, 2006; Lyster, 1998a; 1998b; Lyster, 2004; Lyster & Ranta, 1997; Mackey & Philp, 1998; Mackey et al., 2000; Révész 2009; Révész & Han, 2006; Sato & Lyster, 2012; Sheen, 2006, 2010; Yang & Lyster, 2010; Yilmaz & Yuksel, 2011). While some studies focused only on recasts (Braidı, 2002; Doughty & Varela, 1998; Egi, 2007a, 2007b; Ellis & Sheen, 2006; Goo & Mackey, 2013; Han, 2002; Hawkes & Nassaji, 2016; Ishida, 2004; Kim & Han, 2007;

Leeman, 2003; Li, 2015b; Loewen, & Philp, 2006; Lyster, 1998b; Mackey & Philp, 1998; Mackey et al., 2000; McDonough & Mackey, 2006; Révész, 2009, 2012; Révész & Han, 2006; Révész, Sachs & Mackey, 2011; Sheen, 2006; Carpenter, et al., 2006; Yilmaz & Yuksel, 2011), other studies attempted to examine whether recasts were more beneficial to L2 learning than other types of CF such as prompts (Ammar, 2008; Ammar & Spada, 2006; Carroll & Swain, 1993; Dilans, 2010; Ellis et al., 2006; Ellis, 2007; Gooch et al., 2016; Kartchava & Ammar, 2014; Loewen & Philp, 2006; Lyster, 2004; Lyster & Izquierdo, 2009; Lyster & Mori, 2006; Ranta & Lyster, 2007; Sato & Lyster, 2012; Van De Guchte, Braaksma, Rijlaarsdam & Bimmel, 2015; Yang & Lyster, 2010), metalinguistic explanation (Carroll & Swain, 1993; Ellis et al., 2006; Loewen & Nabei, 2007; Sheen, 2010; Thomas, 2018), explicit correction (Lyster, 1998b; Yilmaz, 2012, 2013), and metalinguistic corrections (Goo, 2012; Li, 2013; Sheen, 2007a).

When the effects of prompts were compared with those of recasts, results from most previous CF studies have been strongly consistent in showing that prompts were more facilitative for L2 acquisition than recasts (Ammar, 2008; Ellis et al., 2006; Ellis, 2007; Loewen & Philp, 2006; Lyster, 2004; Yang & Lyster, 2010) while some studies have shown that prompts and recasts had similar beneficial effects (Kartchava & Ammar, 2014; Lyster & Izquierdo, 2009) on L2 learning. When the students' proficiency level was concerned, mixed results were found. For high-proficiency learners, prompts and recasts were equally effective (Ammar & Spada, 2006). However, for low-proficiency learners, the effects of prompts were significantly superior to those of recasts (Ammar & Spada, 2006).

Sheen (2007a) compared the effects of recasts and metalinguistic corrections on the use of English articles and found that the metalinguistic group significantly

outperformed both the recast group and the control group. However, Goo (2012) obtained different results. In a study conducted with 54 Korean learners learning English as a foreign language, he provided evidence for the equal effects of recasts and metalinguistic feedback.

Yilmaz (2012) investigated the effectiveness of recasts and explicit correction. The participants of this study were learners of Turkish who were assigned to either recast group or explicit correction group. Their errors in the use of locative and plural Turkish morphemes were corrected during communication games. The results of this study revealed that the explicit correction group outperformed the recast group.

Lyster (1998a) and Mackey et al. (2000) argued that recasts containing morphosyntactic errors are ineffective in facilitating learning because they may be so implicit and nonsalient that learners cannot notice their corrective force. However, recasts could be more or less explicit first depending on the linguistic focus, recast length and the number of changes (Egi, 2007a; Philp, 2003; Loewen & Philp, 2006; Ellis & Sheen, 2006; Nicholas, Lightbown & Spada, 2001; Sheen, 2006). Regarding linguistic targets, recasts that involve a particular linguistic item are more explicit than the ones that focus on multiple linguistic items (Egi, 2007a; Philp, 2003; Sheen, 2006).

Importantly, recasts of lexical and phonological targets are more likely to be recognized than those of morphosyntactic targets (Lyster, 1998b; Mackey et al., 2000; Sheen, 2006; Carpenter et al., 2006). Moreover, whether recasts are more or less implicit/explicit depends on their length. Recasts can be full (the reformulations involve the whole erroneous utterance) or partial (only a part of the original erroneous utterance is reformulated). The latter is more salient and explicit than former (Egi, 2007a; Loewen & Philp, 2006; Lyster, 1998a; Sheen, 2006). In addition,

recasts that involve fewer changes are more explicit (Egi, 2007a; Loewen & Philp, 2006; Philp, 2003; Sheen, 2006).

2.4.3 Affecting Factors for CF Effectiveness

The results of CF studies have shown that in second language classrooms, CF is consistently more effective than no CF (see Lyster et al, 2013 for a narrative review). However, there are a number of factors influencing the effectiveness of CF (Loewen, 2012). These moderating factors include types of CF, target features, learner levels and learner cognitive abilities/aptitude.

Types of CF. The degree of explicitness and saliency of recasts in turn affected L2 learning differently (Braidı, 2002; Egi, 2007a; Ellis & Sheen, 2006; Leeman, 2003; Loewen & Philp, 2006; Philp, 2003; Sheen, 2006). Loewen and Philp (2006) explored the nature of recasts provided in adult L2 classrooms and different factors affecting the effectiveness of recasts measured by both posttests and successful uptake. One hundred and eighteen learners and 12 teachers from a private language school in Auckland, New Zealand participated in this study. Thirty two hours of their classroom interaction during meaning-focused lessons was observed. The results of this study indicated that certain characteristics of recasts had a significant impact on the learners' successful uptake and test scores. With regard to uptake, recasts with declarative intonation, fewer changes, stress and multiple feedback moves resulted in successful uptake. With regard to accuracy, on posttests, intonation, morpheme length, and number of changes affected the learners' accuracy greatly. Especially, short recasts with few changes and interrogative intonation were found to be more effective. The results of Loewen and Philp's (2006) study concurred with those reported in other studies (Egi, 2007a;

Philp, 2003; Sheen, 2006) in demonstrating that shorter recasts with fewer changes will be of more benefit.

Target features. Different CF types are differentially facilitative of acquisition in relation to the target structures (Ellis, 2007; Ishida, 2004; Iwashita, 2003; Kartchava & Ammar, 2014; Leeman, 2003; Yang & Lyster, 2010; Yilmaz, 2012; Van de Guchte et al., 2016). This is the factor that the current study particularly focused on—i.e., the role of CF in different aspects of L2 grammar acquisition.

Yang and Lyster (2010) conducted a classroom-based research to explore whether and to what degree recasts and prompts can differentially affect Chinese EFL learners' acquisition of regular and irregular past tense. With regard to the acquisition of the regular past tense, learners benefited more from prompts than from recasts. Comparatively, prompts and recasts had similar effects in the acquisition of irregular past tense.

Ellis (2007) compared the effects of recasts and prompts on the acquisition of two grammatical structures (the English past tense –ed and comparative –er). He found superior effects for prompts over recasts. Moreover, prompts were more beneficial for the acquisition of the comparative than for the acquisition of the past tense forms.

Kartchava and Ammar (2014) scrutinized the effects of two types of CF (prompts and recasts) targeted two different structures (simple past and questions in the past). The analysis showed that CF on the past tense promoted more noticing. On the noticing of the past tense, the prompt and mixed groups were significantly more superior the recast group. In addition, the accuracy scores for the past tense increased more than those for questions.

Leeman (2003) attempted to investigate how recasts promote the acquisition of two target structures (number agreement and gender agreement). The results obtained revealed that recasts benefited learners' acquisition of Spanish number agreement to a greater extent than Spanish gender agreement.

On the whole, the primary studies reviewed here suggest the complex interaction between linguistic structures, their inherent complexity and learning difficulty. Corrective feedback techniques are likely to be effective, when they target more salient, rule-based and relatively difficult linguistic features—a topic that the current study is designed to further explore.

Learner proficiency levels. The effects of CF also depended on different learner variables (Carroll et al., 1992; Mackey & Philp, 1998, Nicholas et al., 2001; Philp, 2003; Lyster, 2004). One of which is students' proficiency levels (Ammar & Spada, 2006; Carroll et al., 1992; Mackey & Philp, 1998; Philp, 2003; Lyster, 2004) which include developmental level (Carroll et al., 1992; Mackey & Philp, 1998; Philp, 2003) and proficiency level (Ammar & Spada, 2006; Iwashita, 2001; Lyster, 2004).

Carroll et al. (1992) examined the effects of explicit correction. Their study involved 79 adult native speakers of Canadian English who were at intermediate and advanced levels of proficiency in French. Participants were divided into experimental and comparison subgroups based on their proficiency levels. All of them were trained on two different word formation rules of French suffixation. However, when the experimental groups made an error, they were given feedback while the comparison groups were never corrected. The results suggested that the experimental groups outperformed the comparison groups. Moreover, CF seemed to be more helpful for the advanced experimental group because they retained the information they had learned better.

Philp (2003) explored the learners' noticing of recasts. They focused on the developmental level of the learner as one of the factor affecting noticing. The results suggested that recasts might be less beneficial to low-level learners because they were not ready to acquire the question forms and had limited ability to notice recasts. The findings summarised here were similar to those of Mackey and Philp (1998).

Lyster (2004) investigated the differential effects of prompts and recasts on the acquisition of grammatical gender in French for both low and high proficiency level learners. He found out that high proficiency learners benefited equally from the two CF techniques while low proficiency learners benefited much more from prompts than from recasts. These findings were in line with the results of Ammar and Spada's (2006) study.

Learner cognitive abilities/aptitude. In addition to learners' levels, in some CF studies, language aptitude has also been examined as one individual difference factor influencing the extent to which learners benefit from CF (Li, 2013, 2015; Mackey, et al., 2010; Mackey et al., 2002; Sheen, 2007a; Sagarra, 2007; Trofimovich et al., 2007, Yilmaz, 2013).

Sheen (2007a) examined the effects of CF on the acquisition of articles and the extent to which language aptitude mediates those effects. This study involved 91 ESL intermediate-level adult learners at a community college in the United States. The participants were divided into a direct-only correction group, a direct metalinguistic correction group, and a control group. Sheen (2007a) found that the two CF groups outperformed the control group. Besides, there was a significant association between the students' gains and their language aptitude under the metalinguistic feedback condition, but not under the other condition. Learners with

higher language analytic ability showed more accuracy improvement. This finding was supported by other studies (Trofimovich et al., 2007, Yilmaz, 2013).

Trofimovich et al. (2007) examined how cognitive factors such as attention control, phonological and working memory and analytical ability affected learners' ability to notice and benefit from recasts targeted English possessive determiners and transitive verbs. A total of 32 adult Francophone ESL learners of English performed picture description tasks then received recasts. Then they described the pictures again. The results of this study suggested that learners with more extensive stronger analytical ability skills and more flexible attention control were more likely to benefit more from recasts. However, there was a weak and non-significant association between phonological, working memory and accuracy gains.

Mackey et al. (2010) examined the relationship between learners' working memory capacity and their modified output. Forty-two college-level English students of Spanish in the United States interacted with a native speaker of Spanish on four communicative tasks and received CF in the form of clarification requests and repetitions during interaction. The results of this study showed that learners with higher working memory capacities tended to produce significantly more modifications. These results were compatible with those of previous CF studies (Sagarra, 2007). Later in 2013, in a study involving 48 adult native speakers of English studying Turkish, Yilmaz found that both working memory capacity and language analytic ability affected the efficacy of CF positively.

Révész and Han (2006) examined the impact of task content familiarity and task type on the efficacy of recasts. Their study involved 36 adult ESL learners who were randomly divided into 4 groups: the Same Video group, the Different Video Group, the Same Notes group, and the Different Notes group. All the groups

received recasts targeted the past progressive form while performing the tasks.

Results pointed to the positive effects of task content familiarity on the accurate use of the target structure. In particular, learners who received recasts while performing tasks with familiar content outperformed learners who performed different content tasks. With regard to task type, the results of the oral tests confirmed that different types of task affected the degrees of accuracy development differently. The video treatment groups significantly outperformed the notes groups.

2.4.4 CF Timing

Finally, I will provide an overview on another crucial factor—immediate vs. delayed CF—that the current study took into consideration in methodology. Immediate type of CF could be distinguished from the delayed one (Li et al., 2016; Lyster et al., 2013; Nakata, 2015; Nassaji, 2007; 2011; Rolin-Ianziti, 2010).

According to Sheen and Ellis (2011), immediate CF involved online attempts to immediately correct students' erroneous utterances while they are speaking whereas delayed CF involved offline attempts to delay the correction until the oral activity was completed.

According to Scrivener (2005), correction should only occur after the completion of communicative activities since providing feedback during task performance will affect fluency negatively. This view was supported by Harmer (2007) who proposed that immediate CF will interrupt the learner's utterance and should be avoided if the errors do not seriously misinterpret the meaning or make the message difficult to understand.

Hedge (2000) introduced some delay CF techniques as follows:

- Recording an activity then asking students to listen and do the correction on their own
- Making notes of students' errors while they are speaking and correcting them afterwards

For many years, the above categories of corrective feedback—delayed techniques—have not received the same attention in the literature. While a large volume of published studies have examined the effects of immediate corrective

feedback on L2 learning (Ammar & Spada, 2006; Ellis, 2007; Ellis et al., 2006; Kartchava & Ammar, 2014; Lyster, 2004; Lyster & Ranta, 1997; Mackey & Philp, 1998; Révész 2009; Sato & Lyster, 2012; Sheen, 2010; Thomas, 2018; Yang & Lyster, 2010; Yoshida, 2008), there are relatively few studies examining the effectiveness of delayed feedback (Nassaji, 2007; 2011) or directly comparing immediate and delayed feedback (Nakata 2014; Li et al, 2016).

In her descriptive study, Rolin-lanziti (2010) examined the organization of delayed corrective feedback sequence. The participants of this study were four teachers of French and their students in an Australian tertiary institution. While the students were performing their tasks, the teachers made notes of their errors. The delayed correction sequences occurred after learners completed a communicative task. The teachers and their students discussed erroneous utterances orally. Rolin-lanziti (2010) classified two different approaches of delayed correction. The first one was the “teacher-initiated/ completed correction” approach in which the teacher located the error and replaced it with the correct version. The second one was the “teacher-initiated, student-correction” approach in which the teacher initiated the correction then left it to the student to complete the repair.

Li et al. (2016) compared the effects of immediate and delayed corrective feedback targeted the English past passive on oral production. One hundred and twenty EFL learners at a public school in China were randomly assigned to four conditions: immediate feedback, delayed feedback, task-only, and control. The three experimental groups received two hours of instruction consisting of 2 sessions. In every session, they listened to a narrative, practiced retelling it in pairs, and individually told the story to the rest of the class. The task-only group only performed the tasks while the two feedback groups performed the tasks and received either

immediate or delayed corrective feedback in the form of a prompt, followed by recasts. The control group did not receive any treatment but took the pretests and posttests. The results of this study showed that both the immediate and delayed feedback resulted in gains in grammaticality judgment test scores.

For a summary of key corrective feedback studies, see Table 3 below.

Table 3: Summary of 10 Key Corrective Feedback Studies

Lyster & Ranta (1997)	
Participants	- 104 students - 4 teachers from a primary school in Montreal
Target of instruction	No specific language forms were targeted (either phonological, lexical, or grammatical)
Instructional treatment	Length: 18.3 hours Data were collected from 14 subject-matter lessons and 13 French language arts lessons.
CF techniques	- Explicit correction - Recasts - Clarification requests - Metalinguistic cues - Elicitation - Repetition
Outcome measures	Communicative interaction
Findings	- Recasts tended to be used the most frequently by teachers.

- Recasts were least effective in terms of eliciting student-generated repair.

Mackey & Philp (1998)

Participants 35 adult ESL learners in Sydney, Australia

Target of
instruction Question formation

Instructional
treatment Length: 2 hours and 30 minutes

- recast ready group (n = 9)
- recast unready group (n = 8)
- interactor ready group (n = 6)
- interactor unready group (n = 6)
- control group (n = 6)

The two recast groups engaged in negotiated interaction tasks and received intensive recasts of their nontarget-like utterances. The two interactor groups performed the same task but did not receive recasts. Ready groups include students who were developmentally ready to acquire the question forms while unready groups include students who were not. The control group only took the tests.

CF techniques Recasts

Outcome Picture differences
measures

Findings The analysis of the sustained stage development of each participant showed that the recast ready group outperformed the other 4 groups and was the only group that showed significant improvement.

Lyster (2004)

Participants 179 fifth-grade immersion students

Target of French grammatical gender
instruction

Instructional Length: 9 hours
treatment - form-focused instruction (FFI) + recast group (n = 38)
 - FFI + prompt group (n = 49)
 - FFI only group (n = 41)
 - control group (n = 51)

Treatment: The FFI only group received FFI but no particular type of feedback. The two CF groups received FFI and either prompts or recasts while the control group did not receive any special FFI.

CF techniques - Prompts

	- Recasts
Outcome measures	- 2 written tasks (a binary-choice test and a text completion Test) - 2 oral tasks (an object-identification test and a picture-description test)
Findings	- The three treatment groups (FFI-recast group, FFI-prompt group and FFI-only group) outperformed the control group in both written and oral tasks. - In written tasks, FFI was more effective when it was combined with corrective feedback and prompts were more effective than recasts in FFI.

Ellis et al. (2006)

Participants	34 adult lower intermediate ESL students in New Zealand
Target of instruction	The regular past tense -ed.
Instructional treatment	Length: 1 hour - implicit feedback group (n = 12) - explicit feedback group (n = 12) - control group (n = 10)

Treatment: The two treatment groups were given picture sequences, read written account of the story and retold it. They received either recasts or feedback in the form of metalinguistic information while they were performing the tasks. The control group neither completed the tasks nor received any feedback.

CF techniques	- Recasts - Metalinguistic explanation
Outcome	- An oral imitation test
measures	- An untimed grammaticality judgment test - A metalinguistic knowledge test
Findings	- Explicit feedback was more effective than implicit feedback and no feedback.

Ammar & Spada (2006)

Participants	64 sixth grade ESL students in Montreal, Canada
Target of instruction	Third-person singular English possessive determiners
Instructional treatment	Length: 6 hours - recast group - prompt group

- control group

All the groups were involved in 11 practice sessions in which they carried out communicative activities. During the activities the experimental groups received either recasts or prompts (depending on the condition) when they made errors in the use of English possessive determiners while in the control group, the teacher did not react to these errors.

CF techniques	<ul style="list-style-type: none"> - Prompts - Recasts
Outcome measures	<ul style="list-style-type: none"> - Passage correction - Oral picture description - MEQ test - Vocabulary tests
Findings	<ul style="list-style-type: none"> - The two CF groups outperformed the control group. - For high-proficiency learners, prompts were as effective as recasts. - For low -proficiency learners, prompts were significantly more effective than recasts.

Révész (2009)

Participants	90 adult elementary and pre-intermediate EFL learners
--------------	---

Target of instruction	The past progressive form
Instructional treatment	<p>Length: 6 weeks</p> <ul style="list-style-type: none"> - recast + photo group (completed a description task with photo support and received recasts, n = 18) - recast – photo group (completed a description task and received recasts, n = 18) - nonrecast + photo group (completed a description task with photo support, n = 18) - nonrecast – photo group (completed a description task, n = 18) - control group (n = 18) <p>The four experimental groups took part in three treatment sessions while the control group only took the pretest and posttests.</p>
CF techniques	- Recasts
Outcome measures	<ul style="list-style-type: none"> - A written description task - An oral description task with photo support - An oral description task without photo support
Findings	

- The recast – photo group showed more improvement than the recast + photo group.
- The nonrecast + photo group outperformed the nonrecast – photo group.

Sheen (2010)

Participants	143 adult intermediate ESL learners of various L1 backgrounds from a college in the United States
Target of instruction	English definite and indefinite articles
Instructional treatment	<p>Length: 1 hour</p> <ul style="list-style-type: none"> - oral recast group ($n = 26$) - oral metalinguistic group ($n = 26$) - written direct correction group ($n = 31$) - written direct metalinguistic group ($n = 32$) - control group ($n = 28$) <p>All the 4 treatment groups engaged in two 30-min communicative narrative tasks in which the two oral CF groups read stories, then retold them, and received CF while the written CF groups read stories, rewrote them, and were given CF. The control group only took the pretest and two posttests without completed the tasks.</p>

CF techniques	<ul style="list-style-type: none"> - Recasts - Oral metalinguistic CF - Written direct correction - Written direct metalinguistic CF
Outcome measures	<ul style="list-style-type: none"> - Speeded dictation test - Written narrative test - Error correction test
Findings	<p>Three CF groups (oral metalinguistic, written direct correction, and written direct metalinguistic) significantly outperformed the control group.</p> <ul style="list-style-type: none"> - Implicit oral recasts targeted article errors were ineffective in helping learners improve their grammatical accuracy.

Yang & Lyster (2010)

Participants	72 Chinese undergraduate EFL students
Target of instruction	Regular and irregular English past tense
Instructional treatment	<p>Length: 2 hours</p> <ul style="list-style-type: none"> - prompt group - recast group

- control group

Treatment: All the groups engaged in form-focused oral practice activities but the two CF received CF in response to their past tense errors while the control group received feedback on the content.

CF techniques

- Prompts
- Recasts

Outcome measures

- Retelling stories based on word cues (oral)
- Story narration (written)

Findings

- All the three groups showed significant improvement between pre-test and post-tests.
- In the use of regular past tense forms, prompts were more effective than recasts.
- In the use of irregular past tense forms, prompts and recasts had similar effects.

Kartchava & Ammar (2014)

Participants 99 high-beginner college level ESL learners

Target of instruction

- Simple past
- Questions in the past

Instructional treatment	<p>Length: 4 hours</p> <ul style="list-style-type: none">- recast group (n = 31)- prompt group (n = 25)- mixed group (n = 23)- control group (n = 20) <p>The participants in the three treatment groups engaged in a communicative task promoting the use of both linguistic targets and received CF from the teachers while the control group only took the tests.</p>
CF techniques	<ul style="list-style-type: none">- Prompts- Recasts
Outcome measures	<ul style="list-style-type: none">- Picture description- Spot-the-differences
Findings	<ul style="list-style-type: none">- CF on the past tense was noticed more.- For the past tense, the prompt and mixed groups were able to notice significantly more CF than the recast group.- The accuracy levels for the past tense increased more than those for questions.- There were no significant differences between the two targets across groups.

Participants	49 upper secondary Swedish EFL learners
Target of instruction	English subject–verb agreement
Instructional treatment	Length: 2 hours - metalinguistic group (n = 15) - analogy-based group (n = 13) - explicit correction group (n = 12) - control group (n = 9)
CF techniques	- Explicit correction - Metalinguistic corrective feedback - Analogy-based corrective feedback
Outcome measures	- Timed grammaticality judgment task - Untimed Grammaticality Judgment Task - Sentence completion tasks
Findings	- All the three CF groups showed significant delayed improvement on the untimed grammaticality judgment task for ungrammatical items. - Analogy-based CF had lowest means on the immediate posttest compared to the other two types of CF.

2.4.5 SUMMARY of CF Literature

In sum, four broad conclusions emerge from this section. First, in meaning-oriented language classrooms, many different types of CF have been used and found to have positive effects on L2 development (Ammar & Spada, 2006; Ellis, 2007; Ellis, et al., 2006; Goo, 2012; Kartchava & Ammar, 2014; Li, 2013; Loewen & Nabei, 2007; Lyster, 2004; Lyster & Ranta, 1997; Mackey & Philp, 1998; Révész, 2009; Sheen, 2010; Thomas, 2018; Yang & Lyster, 2010).

Second, recasts have been identified as the most frequently used type (Ellis, Basturkmen, & Loewen, 2001; Lyster & Ranta, 1997; Panova & Lyster, 2002; Zyzik & Polio, 2008) but less effective than other CF types (i.e. prompts, metalinguistic feedback, and explicit correction) in facilitating L2 learning (Ammar, 2008; Carroll & Swain, 1993; Ellis et al., 2006; Ellis, 2007; Loewen & Philp, 2006; Lyster, 2004; Sheen, 2007a; Yang & Lyster, 2010). However, in some studies, recasts were found to be as effective as prompts (Ammar & Spada, 2006; Kartchava & Ammar, 2014; Lyster & Izquierdo, 2009) or metalinguistic feedback (Goo, 2012).

Third, there are many factors affecting CF effectiveness. The first factor is CF implicitness. Specifically, explicit types of CF are more effective (Carroll & Swain, 1993; Ellis et al., 2006; Ellis, 2007; Sheen, 2007a). The second factor is CF saliency. Recasts that are more salient (Egi, 2007a; Loewen & Philp, 2006; Philp, 2003; Sheen, 2006; Yilmaz & Yuksel, 2011) are more beneficial. The third factor is target structure. Different CF targets affected CF effectiveness differently (Ellis, 2007; Ishida, 2004; Iwashita, 2003; Kartchava & Ammar, 2014; Leeman, 2003; Yang & Lyster, 2010; Yilmaz, 2012). The next factor is learners' level (Ammar & Spada, 2006; Carroll et al., 1992; Iwashita, 2001; Mackey & Philp, 1998, Nicholas et al., 2001; Philp, 2003; Lyster, 2004). Especially, learners with higher levels of

developmental readiness (Carroll et al., 1992; Mackey & Philp, 1998; Philp, 2003) and proficiency (Ammar & Spada, 2006; Iwashita, 2001; Lyster, 2004) benefited more from CF than learners with lower levels. Another factor is students' language aptitude (Li, 2013; Mackey et al., 2002; Mackey et al., 2010; Sheen, 2007a; Miyake & Friedman, 1998; Sagarra, 2007; Trofimovich et al., 2007). It has been shown in the literature that students with higher aptitude scores appeared to benefit more from CF (Li, 2013; Mackey et al., 2002; Mackey et al., 2010; Sheen, 2007a; Sagarra, 2007; Trofimovich et al., 2007, Yilmaz, 2013). The last factor was task variables (Révész & Han, 2006). In terms of task familiarity, recasts were more effective when they were given while students were performing tasks with familiar content. In terms of task type, video treatment was more effective than notes treatment. Fourth, with regard to the timing, both immediate and delayed CF led to significant developmental gains (Li et al, 2016; Nakata 2014; Nassaji, 2007; 2011).

2.4.6 Gaps to Fill in TBLT and CF Literature

Interestingly, a majority of CF studies have exclusively focused on how providing CF could facilitate L2 grammar learning when learners are involved with communicatively authentic tasks. To my knowledge, however, few studies have tested the efficacy of CF in the context of task repetition and L2 fluency development. As shown in Task-Based Language Teaching literature, L2 learners can greatly improve their fluency (but not accuracy) through repeating the same task multiple times, especially under increasing time pressures (4, 3, 2 minutes) (e.g., Thai & Boers, 2016). Thus, it would be interesting to delve into the extent to which adding CF (accuracy enhancement) to the 4/3/2 activity (fluency enhancement) could impact L2 fluency *and* accuracy development. Here, I tested this topic by conducting a quasi-experimental study with a pre-and-posttest design.

In the previous literature, the outcomes of instructional treatment have been predicted using Skehan (1998, 2009)'s Limited Capacity Hypothesis and/or Robinson (2001a)'s Cognition Hypothesis. They have some debates on the relationship between fluency & accuracy development. In the next section, an overview of these two theories will be provided.

2.5 The relationship between fluency and accuracy development

According to Skehan (2009), parallel processing ability is one of the features that differentiates native speakers and L2 speakers. L2 speakers have to make decisions about which aspect of language performance to prioritize. Skehan's trade-off hypothesis emphasizes that complexity, accuracy and fluency "enter into competition with one another, given the limited attentional capacities of second language users" (Skehan & Foster, 1997, p. 185) and "committing attention to one area, other things being equal, might cause lower performance in others" (Skehan, 2009, p.511). For example, increase in fluency in L2 performance will come at the expense of development of accuracy or complexity. Foster & Skehan (1996) argued for the trade-offs between accuracy and complexity. Conversely, Yuan and Ellis (2003) proposed that the model involved fluency and accuracy. Overall, in spite of demonstrating different research findings, previous studies indicated that the trade-off effects negatively impact L2 language performance. As a result, the simultaneous enhancement in all these aspects of speech production is challenging (Thai & Boers, 2016). Research into complexity, accuracy and fluency has examined different task conditions that help balance the development of these aspects (Housen & Kuiken, 2009, Ahmadian, 2012b, Ahmadian & Tavakoli, 2010).

Robinson (2001a, b, 2011) accepted the trade-off effects but rejected the single-source view of attention. He argued that trade-off effects are due to attentional control. His Cognition Hypothesis suggested that complexity, accuracy and fluency will not compete with each other for attentional resources. In contrast, during L2 task performance L2 learners can draw on multiple and non-competitive attentional resource pools for different aspects of performance (Robinson, 2001a). According to Robinson (2011), the complexity of a task can be increased along two dimensions:

resource-directing and resource-dispersing. Tasks that are complex along resource-directing dimensions push learners to focus their attention and effort on particular aspects of the language system that can be used to complete the task. As a result, L2 performance will be promoted. For example, when the intentional-reasoning demands of tasks are increased, learners will make greater effort at product controlling and have better output monitoring. In contrast, making tasks more complex along resource - dispersing dimensions increases the performative and procedural demands on learners' attentional and memory sources but does not direct them to any particular features of the linguistic systems. Attentional resources can be dispersed by removing planning time before/ during task performance, giving no access to available relevant prior knowledge or increasing the number of concurrent tasks. As a results, all aspects of language production will negatively be affected. This may lead to the trade-off effects. In general, although Skehan and Robinson have debates on the competition in attentional resources, they both agreed that resource-dispersing dimensions of complexity affect language performance negatively and expected the trade-off effects to be found. In the present study, the tasks are made complex by removing online planning time (4/3/2 activity). Although this study does not test the Trade-of Hypothesis and Cognition Hypothesis directly, its results will shed light on the theoretical implications of these two theories.

In addition to investigating the facilitative effects of accuracy enhancement and fluency enhancement on L2 development, the present study also examined the extent to which foreign language aptitude mediated these effects. In the sections that follow, I will introduce the definitions of aptitude, its components and measurement. Then I will review research on interaction between aptitude and individual differences observed in instructed second language acquisition.

2.6 Cognitive Individual Differences in Instructed SLA

In this section, I will provide a comprehensive overview on another crucial aspects of the present study—foreign language aptitude. In the current study, I examined how L2 learners could differentially benefit from a combination of task repetition, time pressure and corrective feedback. However, it is reasonable to assume that such L2 gains could be also susceptible to a great deal of individual variation. In the field of instructed SLA (featuring both TBLT and CF literature), a growing amount of attention has been given to the role of cognitive individual differences in the effectiveness of instruction on L2 acquisition (see Li, 2015a for a meta-analytic review; Skehan, 2015 for a critical/narrative overview).

As I reviewed in previous sections, it is true that what learners experience during instruction (type and quantity of experience) matters for successful L2 learning. However, research has also shown that even if two L2 learners spend the same amount of time on the same kind of practice activities, they will most likely end up with different levels of L2 proficiency in a target language. This is arguably because certain learners are more cognitively adept at learning L2 speech, even if input is limited in quantity and quality. To further examine the complex underlying mechanism of successful L2 learning, scholars have explored, in particular, foreign language aptitude. In the current study, therefore, I not only examined different types of instruction (task repetition, increasing time pressure, delayed CF) on L2 development, but also compared participants' different improvement patterns in accordance with their aptitude profiles.

2.6.1 What Characterizes Foreign Language Aptitude

Foreign language aptitude has been considered as a very important individual differences variable in L2 learning (Li, 2015a). Moreover, it has been described as the best predictor of successful foreign language learning outcomes (Carroll, 1981; Ellis, 1994; Gardner & MacIntyre, 1992). According to DeKeyser (2012), every treatment involves a learning process that requires certain levels of specific aptitudes. Therefore, a specific instructional treatment is effective when the aptitude components it demands are present.

Over the past decades, several definitions of language aptitude have been proposed in the literature. In general, the term language aptitude commonly concerns “specific talent for learning foreign languages that exhibits considerable variation between learners” (Dornyei & Skehan, 2003, p. 613). More specifically, Carroll and Sapon (2002) referred to it as a set of cognitive abilities that are ‘predictive of how well, relative to other individuals, an individual can learn a foreign language in a given amount of time and under given conditions’ (p. 23).

Foreign language aptitude can predict not only L2 learning rate, (Carroll, 1981; Saito, 2017) but also the extent to which learners can attain nativelike ability (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000; Granena & Long, 2013a, 2013b; Kormos, 2013; Saito, 2017) in both instructed contexts and in non-instructional settings. Language aptitude has been viewed as a fairly stable trait of an individual (Carroll, 1973, 1993; Li, 2015a; Saito, 2017; Skehan, 1998).

Emphasizing the role of instructional context, Robinson (2005) defined second language learning aptitude as “strengths individual learners have—relative to their population—in the cognitive abilities information processing draws on during L2 learning and performance in various contexts and at different stages” (p. 46).

Foreign language aptitude has been described as a multi-componential concept, since it comprises a set of cognitive abilities (Carroll, 1981, 1993; Hummel, 2009; Linck, Hughes, Campbell, Silbert, Tare, Jackson, Smith, Bunting & Doughty, 2013; Granena, 2013b; Skehan, 1998, 2012; Sparks & Ganschow, 1991, 1995; Sparks, Patton, Ganschow & Humbach, 2011; Wen, Biedroń & Skehan, 2017). As proposed by Carroll (1981), language aptitude has four components including phonetic coding ability, associative memory, grammatical sensitivity, and inductive language learning ability. Phonetic coding ability refers to the ability to distinguish sounds to associate these sounds with symbols representing them, and memorize these associations. Associative memory is the ability to learn and remember associations between sounds and meaning efficiently and to retain these associations. Grammatical sensitivity involves the ability to recognise the syntactic functions of words in sentences. Inductive language learning ability is viewed as the “ability to infer or induce the rules governing a set of language materials, given samples of materials that permit such inferences” (Carroll, 1981, p. 105). These four components have been extensively investigated in previous studies (e.g., Abrahamsson & Hyltenstam, 2008; de Graaff, 1997; Ehrman & Oxford, 1995; Harley & Hart, 1997; Saito, Suzukida & Sun, 2018).

Subsequently, Skehan (1986, 1989, 1998, 2002) introduced a new model of language aptitude in which the two components suggested by Carroll (1981) - grammatical sensitivity and inductive language learning ability were combined into one category called language analytical ability. According to him, three main abilities making up language aptitude are phonetic coding ability, language analytic ability, and memory ability. Later, in some reviews (DeKeyser & Koeth, 2011; Skehan, 2012) and empirical studies (Erlam, 2005; Harrington & Sawyer, 1992; Juffs &

Sawyer, 2011; Kormos & Sáfár, 2008; Li, 2013; Mackey et al., 2002; Miyake & Friedman, 1998; Suzuki & Dekeyser, 2017), working memory was included as one of the components of aptitude (see Li, 2016 for a meta-analysis).

To measure aptitude, various aptitude tests have been used over the last decades (for a review, see Wen et al., 2017). One of the most popular tests that has been widely used in SLA research is the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959). The MLAT comprises five subtests (Number Learning, Phonetic Script, Hidden Words, Words in Sentences, and Paired Associates) which are designed to tap into the abilities of phonetic coding, grammatical sensitivity, and rote learning. None of these subtests measures inductive language learning ability (Skehan, 2012). Validation studies (e.g., Carroll, 1965) have found that MLAT scores are predictive of students' short-term success, primarily in the early stages of L2 learning and in classroom contexts.

In recent years, language aptitude has been conceptualized as a multifaceted construct in which explicit and implicit language learning aptitudes are distinguished (Saito, 2017). The former involves conscious learning process while the later involves incidental learning process. As a result, several aptitude test batteries have been developed and validated. In Linck et al.'s (2013) validation study for the Hi-LAB (High-level Language Aptitude Battery), potential cognitive predictors of high levels of L2 proficiency were examined. Results of the analysis revealed that associative memory, implicit learning, and phonological short-term memory successfully predicted high-level attainment.

The most recent well-known aptitude test is the free computerized LLAMA test (Meara, 2005). It includes four subtests measuring sound sequence recognition (LLAMA D), associative memory (LLAMA B), phonemic coding (LLAMA E), and

language analytic ability (LLAMA F). According to the results of Granena's (2013a) validation study for the LLAMA test, whereas three LLAMA subtests (LLAMA B, E, and F) were concerned with explicit learning, LLAMA D measures implicit learning. In the current study, all the four LLAMA subtests (D, B, E and F) were used.

There are two broad lines of aptitude research in the literature (Li, 2015a). The first line of research concerns the correlations between learners' aptitude and their ultimate second language attainment (Abrahamsson & Hyltenstam, 2008; Bylund, Abrahamsson & Hyltenstam, 2012; Cochran, McCallum, & Bell, 2010; DeKeyser, 2000, 2010; Granena, 2014; Granena & Long, 2013a; Saito, 2017). The second one involves the extent to which individual differences in language aptitude mediates the effects of different types of treatment (Benson & Dekeyser, 2018; Dekeyser, 1995; De Graaff, 1997; Dornyei & Skehan, 2003; Ehrman & Oxford, 1995; Erlam, 2005; Goo, 2012; Harley & Hart, 1997; Hwu & Sun, 2012; Li, 2013; Robinson, 1995, 1997; Sáfár & Kormos, 2008; Sawyer & Ranta, 2001; Sheen, 2007b; Shintani & Ellis, 2015; Yalçın & Spada, 2016; Yilmaz, 2013; Yilmaz & Granena, 2016; Yilmaz, Granena & Meyer, 2016; Yilmaz & Koylu, 2016; Van Patten & Borst, 2012; Van Patten, Collopy, Price, Borst & Qualin, 2013). The focus of the present study is the later. In particular, I aim to examine how foreign language aptitude mediates the effects of task repetition, fluency enhancement and accuracy enhancement. In what follows, I will review previous studies that examined the correlations between aptitude and instructional treatments.

2.6.2 Aptitude and Treatment Interaction

In previous aptitude-treatment interaction studies, various instructional conditions have been examined (De Graaff, 1997; Erlam, 2005; Hwu & Sun, 2012) including deductive and inductive instruction (Dekeyser, 1995; Erlam, 2005; Hwu &

Sun, 2012; Hwu, Pan & Sun, 2012) or explicit and implicit instruction (De Graaff, 1997). In some other studies, the extent to which language aptitude affects the acquisition of easy and difficult structures was investigated (Robinson, 1997; Yalçın & Spada, 2016). Another line of research focused on the extent to which language aptitude mediates the effects of different types of corrective feedback (Benson & Dekeyser, 2018; Goo, 2012; Li, 2013, 2015b; Sheen, 2007a, 2007b; Shintani & Ellis, 2015; Trofimovich et al., 2007; Yilmaz, 2013; Yilmaz & Granena, 2016; Yilmaz et al., 2016; Yilmaz & Koylu, 2016), such as implicit and explicit feedback (Goo, 2012; Li, 2013; Yilmaz, 2013; Yilmaz & Granena, 2016; Yilmaz et al., 2016) or direct feedback and metalinguistic feedback (Sheen, 2007b; Shintani & Ellis, 2015; Benson & Dekeyser, 2018). See Table 4 for a summary of key language aptitude studies.

Table 4: Summary of 10 Key Aptitude Studies

De Graaff (1997)	
Participants	56 native speakers of Dutch studying at a university in Germany
Design	<p>Participants followed a self-study course consisting of ten 1.5 hour lessons in the artificial language <i>eXperanto</i>. They were divided into two experimental groups:</p> <ul style="list-style-type: none"> - Explicit condition (students were provided with short dialogs in <i>eXperanto</i> and engaged in comprehension activities. Subsequently, they received explanation on grammatical structures). - Implicit condition (after the short dialogs and comprehension activities, students received a rehearsal of some example sentences).
Length of treatment	15 hours
Outcome measures	<ul style="list-style-type: none"> - Sentence judgment task - Gap-filling task - Contextualized Dutch-<i>eXperanto</i> vocabulary translation task - Sentence judgment and correction task
Aptitude measures	- The <i>Words in Sentences</i> subtest of the MLAT was used to measure sensitivity to grammatical structure.

- The *Paired Associates* subtest of the MLAT was used to measure rote memory.

- A lexical inferencing task was developed to measure the capacity to infer the meaning of *eXperanto* words from context.

Findings

- Students in the explicit instruction group outperformed those in the implicit instruction group.

- Aptitude equally affected test performance in both conditions. In both groups, students with higher aptitude demonstrated more gains.

- Explicit instruction for the complex syntactic structure was much more effective than for the simple one.

Erlam (2005)

Participants 60 students studying French as a second language in a secondary school in New Zealand

Design All the participants attended three 45-minute instructional sessions on French direct object pronouns. They were randomly divided into three groups:

- Deductive instruction group (students received explicit instruction involving rule explanation, then engaged in form-focused activities in which they applied these rules).

- Inductive instruction group (students engaged in input-based activities involving hypothesis-testing about the direct object

pronouns. Subsequently, they produced the pronoun forms and identified whether these pronouns were used correctly in given contexts).

- Structured input instruction group (students received explicit information and rule explanation about the target structure then processed spoken and written input, interpreted them correctly and identified errors in them).

Length of
treatment

135 minutes

Outcome
measures

- Listening comprehension test
- Reading comprehension test
- Written production test
- Oral production test

Aptitude
measures

- The *Words in Sentences* subtest of the MLAT was used to measure language analytic ability.
- The *Sound Discrimination* test of the PLAB (Pimsleur's Language Aptitude Battery, Pimsleur, 1966) was used to measure phonemic coding ability.
- A working memory test was designed to assess the processing of information.

- Findings
- The deductive instruction group outperformed the other two groups.
 - In the deductive instruction group, L2 gain scores were not related to language aptitude.
 - In the inductive instruction group, students with greater language analytical ability showed more gains.
 - In the structured input group, students with greater language analytical ability and working memory capacity gained more.

Sheen (2007b)

Participants 91 intermediate ESL learners of various L1 backgrounds in the USA

Design Participants were divided into three groups:

- Direct-only correction group (the location of an error was indicated, then the correct form was provided)
- Direct metalinguistic correction group (the location of an error was indicated, the correct form was provided followed by metalinguistic explanation)
- Control group

The two feedback groups attended two treatment sessions in which they completed narrative tasks and received corrective feedback targeted the English articles. The control group only took the tests.

Length of treatment	2 sessions
Outcome measures	<ul style="list-style-type: none">- A speeded dictation test- A narrative writing test- An error correction test
Aptitude measures	Language Analytic Ability Test developed by Otto (2002)
Findings	<ul style="list-style-type: none">- Both feedback groups outperformed the control group on the immediate posttests.- The direct metalinguistic group outperformed the direct-only correction group.- Learners with high language analytic ability benefited more from corrective feedback.- In the direct metalinguistic group, the correlation between language analytic ability and learners' acquisition of articles was stronger.

Hwu & Sun (2012)

Participants	93 native English-speaking students studying Spanish at a university in the USA
--------------	---

Design	<p>Participants followed a course consisting of five lessons in which they received all instruction via the Internet in an electronic format. Each lesson lasted from 40 to 50 minutes. They were randomly assigned to one of the three groups:</p> <ul style="list-style-type: none">- Deductive group (students were presented with the rule then encountered exemplars followed by explanations)- Explicit inductive group (students observed exemplars and had to answer multiple-choice questions involving the rule).- No-instruction group (students only took the tests) <p>The target structure was the Spanish psych verb <i>gustar</i>.</p>
Length of treatment	Approximately 4 hours
Outcome measures	<ul style="list-style-type: none">- Written sentence production task- Written sentence correction task
Aptitude measures	<ul style="list-style-type: none">- A task involving recall of fifteen Indonesian grammar rules was used to measure memory for text.- The <i>Words in Sentences</i> subtest of the MLAT was used to measure grammatical sensitivity.- The <i>Paired Associates</i> subtest of the MLAT was used to measure associative memory.

- Findings
- The two instructional techniques equally affected learning performance.
 - Under explicit inductive condition, learners with good memory for text significantly gained more.
 - Under deductive condition, learners with poor textual memory tended to perform better. However, the gains were statistically insignificant.

Li (2013)

Participants 78 Chinese L2 learners from two universities in the US

Design Participants were assigned to one of the three conditions:

- Implicit feedback condition (recasts)
- Explicit feedback condition (metalinguistic correction)
- No feedback condition

All the participants completed the treatment tasks (picture description and spot the difference) but only the two treatment groups received feedback in response to their errors in the use of Chinese classifiers.

Length of treatment 45 minutes

Outcome measures

- A grammaticality judgment test
- An elicited imitation test

Aptitude measures	<ul style="list-style-type: none"> - The <i>Words in Sentences</i> subtest of the MLAT was used to measure language analytic ability - A listening span test was used to measure working memory
Findings	<ul style="list-style-type: none"> - Both types of feedback promoted learners' gains. - Under the implicit feedback condition, language analytic ability was a significant predictor. - Under the explicit feedback condition, working memory was a significant predictor.

Yalçın & Spada (2016)

Participants	66 pre-intermediate level EFL learners at a private secondary school in Turkey
Design	Participants were provided with instruction on the passive (a difficult structure) and the past progressive (an easy structure) for four hours each.
Length of treatment	8 hours
Outcome measures	<ul style="list-style-type: none"> - A written grammaticality judgment task - An oral production task

Aptitude - LLAMA test (D, B, E, F)
measures

Findings - Learners with greater levels of grammatical inferencing ability (LLAMA F) demonstrated more gains on the passive on the written measure.
- Learners with better associative memory (LLAMA B) gained more on the past progressive on the oral measure.

Yilmaz & Granena (2016)

Participants 48 L2 learners of English studying at a university in Turkey

Design Participants were randomly assigned to one of the three groups:
- explicit group (a repetition of the learner's error and an indication that the previous utterance was erroneous followed by a correction of the error)
- implicit group (recasts)
- no-feedback group
They engaged in two treatment sessions included three oral production tasks. Learners in the two feedback treatment groups received corrective feedback in response to their errors in the use of English indefinite article while learners in the control group did not.

Length of treatment	Approximately 50 minutes
Outcome measures	A guided oral-production task A spot-the-difference task A story-retelling task
Aptitude measures	LLAMA Language Aptitude subtests (B, E, F)
Findings	Under the explicit feedback condition, learners with higher language aptitude benefitted more from feedback.

Suzuki & Dekeyser (2017)

Participants	40 Japanese ESL learners at beginner-level
Design	Participants attended two training sessions in which they were trained on an element of Japanese morphosyntax - the <i>-te</i> form of the verb used to express the present progressive. They were divided into 2 groups: <ul style="list-style-type: none">- Distributed practice group (the interval between 2 sessions was 7 days)- Massed practice group (the interval between 2 sessions was 1 day).

4/3/2 ACTIVITY REVISITED

100

Length of treatment Approximately 1 hour and 40 minutes

Outcome measures - Rule application test
- Picture sentence completion test

Aptitude measures - LLAMA F was used to measure language analytic ability
- Ospan tasks were used to measure working memory capacity

Findings - Language analytic ability was correlated with learning gains under distributed practice condition.
- Working memory capacity was correlated with learning gains under massed practice condition.

Saito (2017)

Participants 50 young adult Japanese EFL learners at varied proficiency levels

Design All the participants first took the LLAMA tests then performed a picture narrative task and finally completed a language background questionnaire.

Length of treatment 50 minutes

4/3/2 ACTIVITY REVISITED

101

Outcome	<u>Task</u>
measures	<p>A picture narrative task</p> <p>The analysis focused on:</p> <ul style="list-style-type: none">- Pronunciation: segmental and prosodic (word stress, intonation) accuracy- L2 fluency (pauses, speech rate and articulation rate)- Vocabulary usage (lexical appropriateness, lexical richness)- Grammar usage (morphological accuracy; grammatical complexity) <p><u>Judgment</u></p> <p>Five native-speaking raters evaluated:</p> <ul style="list-style-type: none">(a) segmental errors(b) word stress(c) intonation
Aptitude	LLAMA test (D, B, E, F)
measures	
Findings	<p>There were significant correlations between:</p> <ul style="list-style-type: none">- LLAMA E and segmentals- LLAMA B and grammatical complexity (clause to AS-unit ratio)- LLAMA F and lexical richness (diversity and frequency) <p>Besides, there were marginally significantly correlations between:</p>

- LLAMA E and other pronunciation (word stress, intonation) and morphological accuracy

- LLAMA B and fluency (articulation rate)

There were no significant aptitude–proficiency links between LLAMA D and lexical appropriateness.

Benson & Dekeyser (2018)

Participants 151 ESL learners

Design Participants were divided into three groups:

- Direct feedback group
- Metalinguistic feedback group
- Control group

The two feedback groups received corrective feedback in response to their errors in the use of the simple past tense and the present perfect tense whereas the control group received general comments on content.

Length of treatment 3 hours

Outcome measures - Essay writing task
- Form-focused grammar test

The analysis focused on accuracy in the use of the simple past tense and the present perfect tense.

Aptitude measures LLAMA F was used to measure language analytic ability

Findings

- Both feedback groups outperformed the control group for the two verb tenses.
- Only the direct feedback group maintained significant gains for the simple past tense at the time of the delayed posttest.
- For the simple past tense, learners with greater language analytic ability benefited more from direct feedback; for the present perfect, learners with lower language analytic ability benefited more from metalinguistic feedback.

In the aforementioned aptitude studies, the effects of instruction were examined in relation to one (Benson & Dekeyser, 2018; Sheen, 2007b; Van Patten & Borst, 2012; Yilmaz & Koylu, 2016) or more aptitude subcomponents (De Graaff, 1997; Erlam, 2005; Hwu & Sun, 2012; Li, 2013; Saito, 2017; Suzuki & Dekeyser, 2017; Yalçın & Spada, 2016; Yilmaz & Granena, 2016). In most studies, each subcomponent was examined separately (Benson & Dekeyser, 2018; Erlam, 2005; Hwu & Sun, 2012; Li, 2013; Saito, 2017; Sheen, 2007b; Suzuki & Dekeyser, 2017; Yalçın & Spada, 2016) while in a smaller body of research, global aptitude scores were calculated (De Graaff, 1997; Yilmaz & Granena, 2016).

2.6.2.1 Deductive and Inductive Instruction

To investigate the role of language aptitude in determining the effects of deductive and inductive instructional methods on the learning of grammar rules, Hwu and Sun (2012) conducted a study involving 93 university-level students of L1 English learning Spanish as a foreign language in the USA. Participants followed an online course consisting of five lessons that focused on the Spanish psych verb *gustar*. They were randomly divided into three groups. The deductive group was presented with the rule then encountered exemplars followed by explanations. The inductive group observed exemplars and had to figure out the rule. The no-instruction group only took the tests. The results of this study suggested that in the inductive group, learners with good memory for text performed significantly better. In the deductive group, learners with poor textual memory tended to gain more but these gains were statistically insignificant.

Hwu et al. (2014) also examined the interaction between language aptitude and instructional approaches (deductive and inductive). They found that under the

deductive instruction condition, learners with high aptitude performed significantly better than high-aptitude learners. However, these results were contrary to Erlam's (2005) findings which revealed that under inductive condition, students with greater language analytical ability demonstrated more gains. Under deductive condition, L2 gain scores were not related to language aptitude.

2.6.2.2 Explicit and Implicit Instruction

One of the earliest studies examining the interaction between explicit and implicit grammar instruction and language aptitude was De Graaff's (1997) study. Fifty-six native speakers of Dutch studying the artificial language *eXperanto* at a university in Germany were randomly assigned to two treatment conditions. Under the explicit condition, short dialogs were given to students in *eXperanto*. After that, they engaged in comprehension activities and received explanation on grammatical structures. Under the implicit condition, students were provided with a rehearsal of some example sentences after the short dialogs and comprehension activities. According to the results of the analyses, under both conditions, students with higher aptitude showed more learning. These results confirmed the findings of Robinson (1997).

Focusing on 104 non-native speakers of English, Robinson (1997) investigated how foreign language aptitude mediates the effects of four instructional treatments (i.e., incidental, implicit, explicit, and rule-search) on the learning of easy (i.e., the plural morpheme) and hard grammar rules (i.e., *Wh*-question formation). The results of Robinson's (1997) also demonstrated that foreign language aptitude similarly predicted learning outcomes across different instructional types.

2.6.2.3 Easy and Difficult Structures

Yalçın and Spada (2016) investigated the relationship between language learning aptitude and the learning of two English structures. Sixty-six EFL learners at pre-intermediate level at a private secondary school in Turkey received instruction on a difficult structure (the passive) and an easy structure (the past progressive) for 8 hours in total. The results of Yalçın and Spada's (2016) study showed that learners with higher grammatical inferencing ability showed demonstrated more gains on the passive on the written measure while better associative memory helped learners gain more on the past progressive on the oral measure. In contrast, Robinson (1997) found that memory scores predicted accuracy on hard rules while grammatical sensitivity scores predicted accuracy on the easy rule.

2.6.2.4 Different Types of Corrective Feedback

In a study involving 78 Chinese L2 learners in the USA, Li (2013) investigated the extent to which explicit language aptitude mediated the effects of two types of CF (explicit and implicit feedback) on the acquisition of Chinese classifiers. Participants were randomly assigned to one of the three treatment conditions: explicit feedback, implicit feedback and no-feedback. The results of the study suggested that in the implicit feedback group, language analytic ability was a significant predictor while in the explicit feedback group, working memory was a significant predictor. These results were different from those of Goo (2012) which revealed that working memory mediated the effects of implicit feedback. Yilmaz and Granena (2016) also examined the differential effects of explicit feedback and implicit feedback and the extent to which language aptitude mediated their effects. However, unlike Li (2013), Yilmaz and Granena (2016) focused on a broad type of language aptitude which is explicit language aptitude. It covered associative memory, phonemic coding and language

analytic ability. According to the results of this study, learners with higher language aptitude benefitted more from explicit feedback.

In some other CF studies, the relationship between the effectiveness of direct feedback and metalinguistic feedback and learning aptitude was established (Sheen, 2007b; Shintani & Ellis, 2015; Benson & Dekeyser, 2018). In Sheen's (2007b) study, 91 intermediate-level ESL learners at a community college in the United States were assigned a direct-only correction group, a direct metalinguistic correction group and a control group. The two feedback groups received feedback targeted the English articles. Findings revealed that in both CF treatment groups, higher language analytic ability resulted in greater gains. However, the correlation between language analytic ability and learners' gain scores was stronger for the direct metalinguistic group.

Benson and Dekeyser (2018) also investigated the extent to which language aptitude is related to learning outcomes under these two feedback conditions. However, in their study, CF targeted two structures (the simple past tense and the present perfect tense). In addition, delayed posttests were included. For the simple past tense, results showed a strong relation between greater language analytic ability and learners' gains in the direct feedback group. In contrast, there was a strong relation between lower language analytic ability and learners' gains in the metalinguistic feedback group. With respect to long term effects, only under the direct feedback condition, learners could maintain significant gains for the simple past tense.

2.6.3 SUMMARY of Aptitude Literature Review

Taken together, three broad conclusions can be drawn from language aptitude studies. First, language aptitude is an important individual difference variable predicting L2 development (Benson & Dekeyser, 2018; De Graaff, 1997; Dornyei & Skehan, 2003; Ehrman & Oxford, 1995; Erlam, 2005; Goo, 2012; Harley & Hart, 1997; Hwu & Sun, 2012; Li, 2013; Robinson, 1997; Sheen, 2007b; Shintani & Ellis, 2015; Yalçın & Spada, 2016; Yilmaz, 2013; Yilmaz & Granena, 2016).

First and foremost, learners with higher aptitude for explicit learning can gain more (Robinson, 2005). Second, various aptitude components including associative memory (Hwu & Sun, 2012; Yalçın & Spada, 2016), phonemic coding (Erlam, 2005), and language analytic ability (Benson & Dekeyser, 2018; Erlam, 2005; Li, 2013; Sheen, 2007b; Suzuki & Dekeyser, 2017) have been found to moderate L2 development differently as per different learning conditions. Third, not all treatments can be effective for particular individuals. Some instructional techniques (e.g., explicit vs. implicit instruction) can be more or less effective for learners since they have different levels of cognitive abilities.

2.7 Motivation for Current Study

As reviewed above, there is a clear gap between the two somewhat independently developing fields—(a) Task-Based Language Teaching and (b) Corrective Feedback. In the former literature (Task-Based Language Teaching), we have known how L2 learners' performance (complexity, accuracy, fluency) changes thanks to different types of task repetition techniques (Ahmadian, 2011, 2012a, 2013; Ahmadian & Tavakoli, 2010; Arevart & Nation, 1991; Boers, 2014; Bygate, 1996, 2001; De Jong & Perfetti, 2011; De Jong, 2012; Fukuta, 2016; Gass et al., 1999; Kim, 2013; Kim & Tracy-Ventura, 2013; Lambert et al., 2016; Lynch & Maclean, 2000, 2001; Maurice, 1983; Nation, 1989; Pinter, 2005; Sample & Michel, 2014; Thai & Boers, 2016).

In the latter literature (Corrective Feedback), it has been demonstrated that corrective feedback is facilitative of L2 learners' accurate use of grammar (Ammar, 2008; Ammar & Spada, 2006; Carroll & Swain, 1993; Doughty & Varela, 1998; Egi, 2007a; Ellis, 2007; Ellis et al., 2006; Han, 2002; Ishida, 2004; Kartchava & Ammar, 2014; Leeman, 2003; Loewen & Philp, 2006; Lyster, 2004; Lyster & Izquierdo, 2009; Lyster & Mori, 2006; Mackey & Philp, 1998; McDonough & Mackey, 2006; Ranta & Lyster, 2007; Révész, 2009; Sauro, 2009; Sheen, 2007a, 2010; Thomas, 2018; Yang & Lyster, 2010), pronunciation (Gooch et al., 2016; Lee & Lyster, 2016; Saito, 2013, 2015; Saito & Lyster, 2012a, 2012b) and vocabulary (De la Fuente, 2002; Dilans, 2010; Ellis & He, 1999; Nakata, 2014, 2015). However, very few scholars have examined whether, to what degree and how the effectiveness of CF varies under different task conditions.

One such exception is Révész (2009) which examined the differential effects of recasts on L2 grammar acquisition when learners engaged either in simple (with contextual support) or in complex (without contextual support) tasks. Révész (2009) conducted an experimental study involving 90 adult EFL high school students at elementary and pre-intermediate levels in Hungary. She examined the effects of recasts in conjunction with task complexity on the acquisition of the past progressive forms. In this study, task complexity was operationalized as the presence or absence of contextual support. Participants were divided into four experimental groups and a control group. The control group only took the pretest and posttests while the four other groups took part in three treatment sessions in which they described the pictures. The two recasts groups received recasts on their inaccurate use of the past progressive forms while the two non-recast groups did not receive any feedback. The four treatment groups also differed in whether they were allowed to view the photos while describing them.

The study found that the students who received recasts without contextual support (complex tasks) showed more improvement than students who also received recast with contextual support (simple tasks). This indicated that recasts were more effective when students engaged in tasks that were more complex. In contrast, for the two non-recasts groups, the students who viewed photos outperformed who did not. These findings support the claim that tasks that are less demanding along the resource dispersing dimensions promote greater L2 development.

The results were explained in line with Skehan's (1998, 2002, 2012) Limited Capacity Hypothesis which states that L2 learners access the limited amount of cognitive resources while using language for meaningful purposes. When learners participate in simple tasks, they could save much of their cognitive resources in order

to notice, understand and make the most of recasts. In contrast, when they engage in complex tasks, they may not have enough cognitive resources left for processing recasts in an efficient and effective way.

In sum, Révész (2009) have provided important insights on the relationship between the nature of task conditions and L2 speech learning. Overall, L2 learners can make the most of accuracy enhancement (e.g., recasts), especially when they engage in simple rather than complex tasks. As such, L2 learners can allocate their limited cognitive resources with a view of perceiving the corrective intension of recasts and paying selective attention to form while using language for meaning (i.e., focus on form). Comparatively, cognitively demanding tasks likely take up learners' cognitive resources which they could otherwise use in order to notice, understand and process corrective feedback.

The current study was designed to further pursue this topic by closely examining how providing CF could be facilitative of SLA when L2 learners participate in task repetition activities but with different levels of complexity and demand—with and without increasing time pressure (see below). Additionally, in conjunction with the growing amount of literature on the role of cognitive indifferences in instructed SLA (e.g., Suzuki & DeKeyser, 2017; Yalçın & Spada, 2016; Yilmaz & Granena, 2016), the study also explored whether, to what degree and how learners' different improvement patterns could be ascribed not only to task conditions (task repetition, increasing time pressure, delayed corrective feedback), but also to their aptitude profiles, measured via LLAMA (Meara, 2005) in terms of associative memory, phonemic coding, language analytic ability and sound sequence recognition.

2.8 Current Study

2.8.1 Research Questions

In the current study, a total of 48 university-level students participated in three 20-minute dyadic sessions. In each session, they were given a different topic and repeated a monologue three times. Importantly, the way they repeated the task differed according to the following four group conditions: (a) task repetition only (control); (b) task repetition + fluency enhancement (increasing time pressure); (c) task repetition + accuracy enhancement (corrective feedback) and (d) task repetition + fluency enhancement + accuracy enhancement. As for outcome measures, fluency was analyzed from the perspectives of speed (articulation rate), breakdown (the frequency and location of pauses) and repair (the ratio of self-repetitions and repairs). Accuracy was evaluated in terms of the participants' targetlike use of irregular and regular English past tense forms. In addition, three overall measures of accuracy were also adopted (global accuracy, semantic accuracy, and morphosyntactic accuracy). The present study was conducted to examine the extent to which adding fluency (increasing time pressure) and accuracy (corrective feedback) enhancement increase the pedagogical potential of task repetition. Therefore, three research questions were formulated:

- R1: To what degree does task repetition impact L2 fluency and accuracy development?
- R2: To what degree does task repetition + FE impact L2 fluency and accuracy development?

- R3: To what degree does task repetition + AE impact L2 fluency and accuracy development?
- R4: To what degree does task repetition + FE+ AE impact L2 fluency and accuracy development?

Finally, the study explored the role of cognitive individual differences in the effectiveness of task-based instruction and corrective feedback on L2 development; the following research question was formulated:

- R5: To what degree are learners' improvement patterns related to their cognitive individual differences in associative memory, phonemic coding, language analytic ability and sound sequence recognition?

2.8.2 Predictions

According to Robinson (2001a, 2003, 2015), task complexity can be conceptualized according to two dimensions: *resource-directing* variables and *resource-dispersing* variables.

In line with the former dimension, tasks are considered more complex, when they entail greater reasoning demands, more elements, and more explicit reference to present/past time framework but without contextual support. In terms of the latter dimension, which is more relevant to the current study, tasks are considered more complex, when they include less planning time and less clear structures. According to this theoretical rationale, the cognitive demand of the task activities used in this study could be ranked in the following order: task repetition only (Control) < adding FE (4/3/2) or AE (delayed CF) < combining FE and AE.

With respect to resource-dispersing aspects of task complexity, it has been claimed that more complex tasks (e.g., less planning time) lead learners to use an L2 less accurately and fluently (cf. see theoretical debates on the relationship between resource-directing aspects of task complexity vs. L2 speech performance and learning, see Robinson, 2001a, 2003, 2015 vs. Skehan, 1998, 2009, 2014). In what follows, predictions were presented in accordance with previous relevant literature in task-based language teaching (e.g., Révész, 2009) and corrective feedback (e.g., Sato & Lyster, 2012) vis-à-vis the suggested task complexity hierarchy (Robinson, Skehan).

R1: To what degree does task repetition impact L2 fluency and accuracy development?

With regard to fluency development, previous studies have been consistent in showing that task repetition would affect fluency positively (Ahmadian & Tavakoli, 2010; Bygate, 1996, 2001; Lambert et al., 2016; Lynch & Maclean, 2000, 2001; Pinter, 2005). However, In terms of transferability, it has been shown in a large volume of published studies that these gains may not transfer to performances of a new topic (De Jong & Perfetti, 2011; Gass et al. 1999; Bygate, 2001). With regard to the effects of task repetition on accuracy development, the results were mixed with some studies reporting accuracy improvement (Bygate, 1996; Gass et al., 1999; Kim & Tracy-Ventura, 2013; Lynch & Maclean, 2000, 2001), and some reporting no accuracy gains (Bygate, 2001; Boers, 2014; Thai & Boers, 2016). Drawing on the findings of these studies, it is reasonable to predict that a significant improvement in fluency, but not in accuracy, will be found, when students repeat the task under constant time condition.

R2: To what degree does task repetition + FE impact L2 fluency and accuracy development?

Building on the empirical evidence of previous studies on task repetition and 4/3/2 activity (Boers, 2014; De Jong, 2012; De Jong & Perfetti, 2011; Nation, 1989; Thai & Boers, 2016), it is predicted that task repetition + FE (increasing time pressure) will greatly enhance the fluency of students' oral production because time pressure will push students to speed up their speech rate. However, under this condition, no compelling evidence of changes in accuracy will be found since time

pressure will make students prioritize meaning over form (Ahmadian & Tavakoli, 2010; Ahmadian, 2012b; Boers, 2014; De Jong & Perfetti, 2011; Nation, 1989; Thai & Boers, 2016; Yuan & Ellis, 2003). This could be explained using Skehan (1998)'s trade-off hypothesis. Moreover, according to Robinson (2001a, 2003, 2015), less online-planning time makes the tasks more complex along dispersing dimensions. Therefore, their speech performance will be negatively affected.

R3: To what degree does task repetition + AE impact L2 fluency and accuracy development?

Drawing on the findings from previous CF studies, it is predicted that adding accuracy enhancement will help students improve their accuracy most markedly (Doughty & Varela, 1998; Li, 2010; Sato & Lyster, 2012), because CF is assumed to draw the students' attention to form. However, according to Skehan (1998, 2014), participants at beginner to intermediate levels cannot attend simultaneously to both fluency and accuracy with their limited cognitive resources. Therefore, it is also predicted that fluency will not be enhanced due to the trade-off effects between fluency and accuracy (Skehan, 1998, 2014; Skehan & Foster, 2008).

R4: To what degree does task repetition + FE+ AE impact L2 fluency and accuracy development?

In keeping with the results of previous studies on 4/3/2 activity and CF, it is predicted that simultaneously engaging in FE and AE would lead to the enhancement of both fluency and accuracy of students' oral production because

both task repetition and time pressure assist fluency development (Boers, 2014; De Jong, 2012; Thai & Boers, 2016). In addition, CF facilitates accuracy (Doughty & Varela, 1998; Li, 2010; Sato & Lyster, 2012). However, in this condition, the task is complex along dispersing dimensions (less online planning time) so students cannot attend to all aspects of language at the same time. Therefore, some debilitating effects will be found (Robinson, 2001a, 2003, 2015; Skehan, 1998, 2014; Skehan & Foster, 2008). In particular, only some dimensions of fluency and accuracy will be enhanced.

R5: To what degree are learners' improvement patterns related to their cognitive individual differences in associative memory, phonemic coding, language analytic ability and sound sequence recognition?

According to the literature review on aptitude x treatment interaction (e.g., Suzuki & DeKeyser, 2017; Yalçın & Spada, 2006; Yilmaz & Granena, 2016), it is predicted that learners with greater aptitude can demonstrate more gains in terms of the accuracy and fluency aspects of L2 development. More specifically, L2 learners who have high-level phonemic coding and language analytic ability may enhance both accuracy and fluency to a great degree, as they can decode grammar information in speech and analyze patterns in an efficient and effective way. Thanks to their linguistic competence for accuracy, such talented learners can use their remaining cognitive resources to fluency development (Suzuki & DeKeyser, 2017). As shown in the previous literature (e.g., Saito, 2017), I may not find any significant effects of incidental learning aptitude (sound sequence recognition), because such

learning does not likely happen during the highly explicit language training provided in the current study.

This chapter has provided a broad overview of the theoretical background to the present study in four major sections: task repetition, increasing time pressure as fluency enhancement, corrective feedback as accuracy enhancement and cognitive individual differences in instructed SLA. In addition, motivation for the current study, research questions and predictions have also been presented. The next chapter proceeds to describe in detail the research methods employed in this study.

CHAPTER 3: METHODOLOGY

This chapter contains six sections. The first section gives an overview of the participants of the present study. The second section addresses the procedure for data collection. The third section describes the content of treatment. The fourth section deals with speech analysis. The fifth section reports the results of inter-coder analysis. The last section is concerned with aptitude measures.

3.1 Participants

This study was conducted at a university in Vietnam. The participants were composed of 48 Vietnamese EFL learners (7 males and 41 females) who volunteered to participate in the study. Their ages ranged from 18 to 21 (*M age* = 20.2 years). All of the participants reported that they had been learning English for at least 7 years prior to the project. None of them had any experience abroad. Not surprisingly, the participants had few opportunities to use English for communicative purposes outside their classrooms. The participants majored in a wide range of fields: education, information technology, history, law, literature and social work and had to take two General English courses per week at the university. Based on their self-reported TOEIC scores, the participants' English proficiency levels ranged from pre-intermediate (370 out of 990) to intermediate (690) (*M* = 504.4). Table 5 provides the details of participant information by group.

3.2 Procedure

Prior to the treatment sessions, the participants were informed the nature of the study, what will be involved in taking part in this study and how the data would be used. Then they signed a consent form prepared following the ethical guidelines of Birkbeck, University of London (the full consent form is available in Appendix A).

Table 5: Participant Information by Group

	Control group (<i>n</i> =12)	FE group (<i>n</i> =12)	AE group (<i>n</i> =12)	FE+AE group (<i>n</i> =12)
Gender	1 male/ 11 females	4 male/ 8 females	0 male/ 12 females	2 male/ 10 females
Age	<i>M</i> = 20.1 (<i>SD</i> = 1.1)	<i>M</i> = 20.1 (<i>SD</i> = .5)	<i>M</i> = 19.9 (<i>SD</i> = 1.4)	<i>M</i> = 20.6 (<i>SD</i> = .7)
TOEIC	<i>M</i> = 446.3 (<i>SD</i> = 85.5)	<i>M</i> = 502.92 (<i>SD</i> = 89.6)	<i>M</i> = 537.9 (<i>SD</i> = 61.4)	<i>M</i> = 530.4 (<i>SD</i> = 84.5)

After that, they filled in a language background questionnaire (for the full questionnaire, see Appendix B) which is a modified version of the language contact profile developed by Freed, Dewey, Segalowitz and Randall (2004). They were asked to provide information such as their background information (i.e, gender, age, hometown, native language, languages use at home, languages for precollege education, whether they have been to an English-speaking country, their English ability, when they first started learning English, their major), and their use of English (the frequency of communicating in English and practicing English inside and outside the classroom, their IELTS, TOEIC or TOFEL scores).

The participants were randomly divided into the four different groups, each of which consisted of 12 students (a) Control, (b) Fluency Enhancement (FE), (c) Accuracy Enhancement (AE) and (d) FE+AE. All of them engaged in three individual tutoring sessions (Sessions 1, 2, 3) which took place in a quiet room at the university. Each session lasted for 20 minutes, and the participants repeated a monologue task on one of three different topics (Topics A, B, C: for details see below). The interval between these sessions was three days. The nature of instructional treatment was different between the four group conditions as follows:

1. **Control ($n = 12$):** In each session, learners repeated a monologue task three times; a total of three different topics were prepared.
2. **Fluency Enhancement ($n = 12$):** Learners repeated each monologue task three times with increasing time pressure (4, 3, 2 minutes).
3. **Accuracy Enhancement ($n = 12$):** Learners engaged in task repetition under constant time conditions (3/3/3); they received corrective feedback on

grammatical accuracy from the researcher between their first and second deliveries during task repetition.

4. **Fluency and Accuracy Enhancement ($n = 12$):** Learners received corrective feedback (AE) during each 4/3/2 activity (FE).

Two weeks before the full-scale study was conducted, a pilot study had been carried out with 8 same profile students. They were also divided into 4 groups. The aims of piloting was to check whether the participants could talk non-stop about a topic in English for 4 minutes (See Appendix D for examples of repeated monologue in the shrinking time condition and Appendix E for examples of repeated monologue in the constant time condition). Piloting was also carried out to test the methodology. It was observed that the tasks are not challenging but for these learners. They could easily fill 4 minutes with talks. In addition, the methods chosen for the experiment are appropriate.

3.3 Content of Treatment

In this subsection, I detail the way task repetition, fluency enhancement and accuracy enhancement were operationalized.

3.3.1 Task Repetition

All of the participants joined three task repetition sessions. For each session, the following three topics were used in this study: (a) *the last favorite movie you watched*, (b) *your last summer vacation*, and (c) *your 16th birthday celebration* (for the actual materials used in the study, see Appendix C). To eliminate any order effects, these topics were used in counterbalanced orders. In every group, the first half of the students started with “favorite movie,” then “last summer vacation” and

“16th birthday” while the other half started with “last summer vacation,” then “16th birthday” and finally “favorite movie.”

For every session, students first had three minutes of pre-task planning time to organize their talks. To help them conceptualize what to say, they were given six prompts. For example, when students were asked to talk about their favorite movie, the following questions were provided to them: *What was it called? What kind of movie was it? When and where did you watch it? Who were the main characters? What happened in the movie? Why did you like it?* As part of this pre-task planning time, all participants were encouraged to take notes about what they wanted to say by using key words, instead of complete sentences. They were allowed to review their notes only between the first and second deliveries (1 minute) but not between the second and third deliveries. For the same methodological decision, see De Jong and Perfetti (2011) and Thai and Boers (2016).

3.3.2 Fluency Enhancement (Increasing Time Pressure).

Whereas the participants in the Control and AE groups repeated the task on an equal time interval (3 min → 3 min → 3 min), those in the FE and FE+AE group were given an increasing amount of time pressure for the first delivery (4 min), second delivery (3 min), and third delivery (2 min).

3.3.3 Accuracy Enhancement (Corrective Feedback)

To avoid hindering the communicative flow of the participants' speech, a decision was made to provide delayed CF only between the first and second deliveries, not during their monologue delivery. Recently, a growing number of CF scholars have advocated the effectiveness of delayed CF (e.g., Li et al., 2016; Nakata 2015; Nassaji, 2007, 2011). Different from immediate CF (correcting errors

immediately after they occur), the delayed CF technique allows teachers/researchers to delay the correction until students finish performing their conversations/tasks (e.g., Ellis, 2009; Rolin-lanziti, 2010). To date, certain studies have explored the provision and effectiveness of delayed CF (Nassaji, 2007, 2011) or directly compared the different effects of immediate and delayed CF (Li et al., 2016; Nakata 2015).

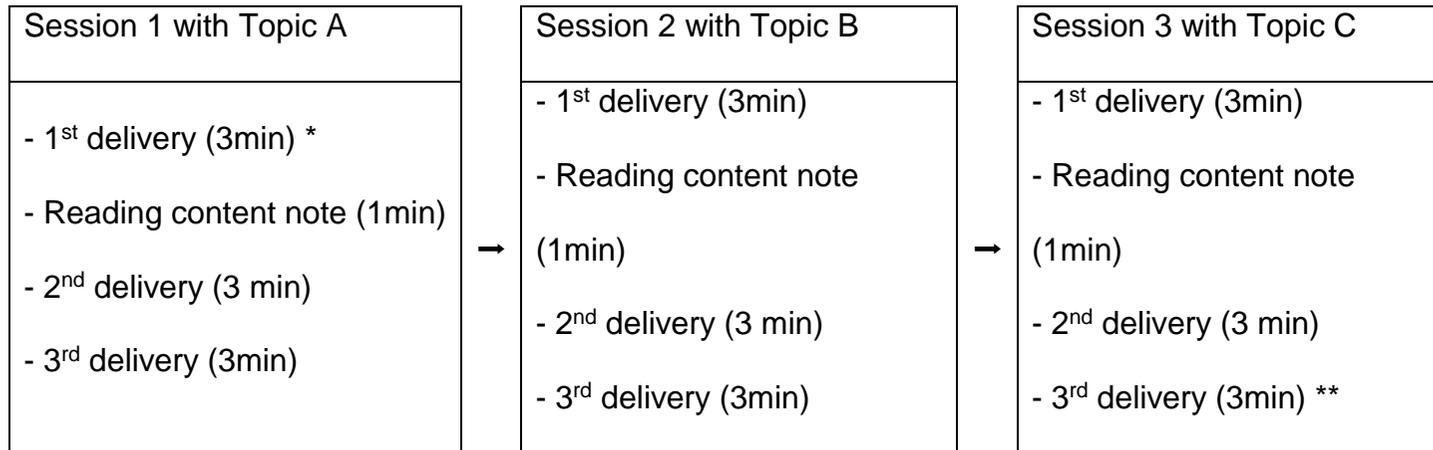
Following the delayed CF procedure adopted in Rolin-lanziti (2010), the researcher wrote down the participants' grammatical errors (past tense) during their first delivery and provided the error logs with metalinguistic explanation to the participants when they finished their monologue. Unlike the participants in the Control and FE group, those who engaged in the accuracy enhancement (delayed CF) did not review their notes on the content. Rather, they reviewed the language-focused feedback from the researcher so that they could reflect on their grammatical accuracy before their second delivery (1 minute).

In terms of the focus of CF, I chose one of the most extensively researched instances: the acquisition of English past tense form (e.g., Ellis et al., 2006; Li et al., 2016; Loewen & Erlam, 2006; Nobuyoshi & Ellis, 1993; Yang & Lyster, 2010). This decision was made for the following reasons. First, CF has been found to be effective for the proletarianization of partially acquired knowledge rather than the acquisition of completely new knowledge (Ellis & Sheen, 2006). English past tense is considered to be an ideal testing ground, since all the participants were already familiar with the target structure. English past tense is generally introduced at the very beginning of many textbooks used in Vietnam. Second, despite learners' familiarity and metalinguistic awareness of English past tense, this particular feature has been considered difficult due to the fact that it is perceptibly and communicatively non-salient to many EFL learners (Ellis et al., 2006; Yang & Lyster,

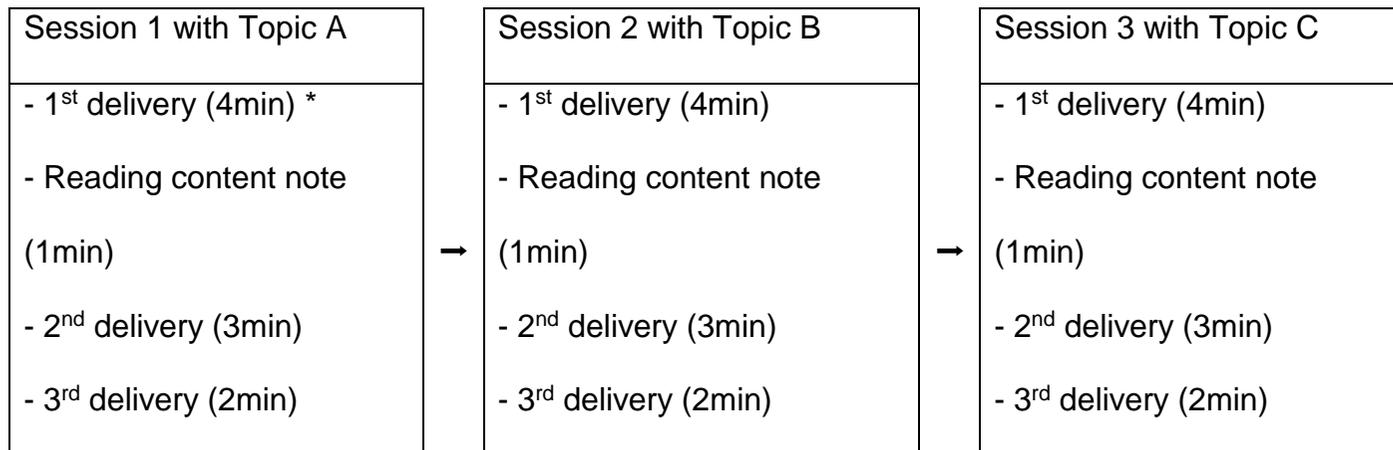
2010). In fact, many Vietnamese EFL students likely fail to master the past tense during speaking despite many years of exposure to it (McDonald, 2000). Third, the previous CF literature has shown that CF treatment can be effective for the acquisition of the English past tense in many classroom settings (Ellis et al., 2006; Ellis, 2007; Doughty & Varela, 1998; Kartchava & Ammar, 2014; Takashima & Ellis, 1999; Yang & Lyster, 2010).

The research design and treatment type was also visually summarized in Figure 3

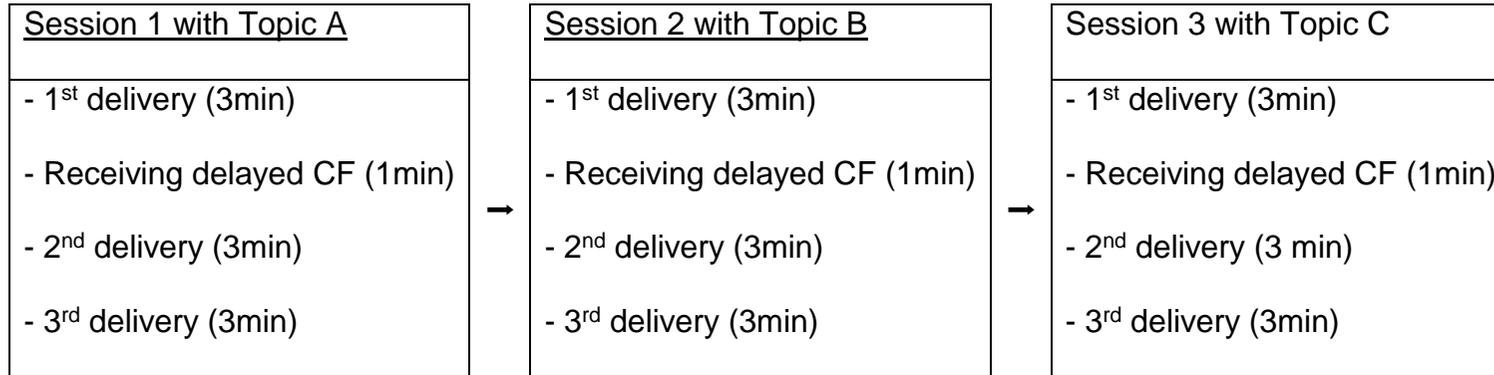
Control Group



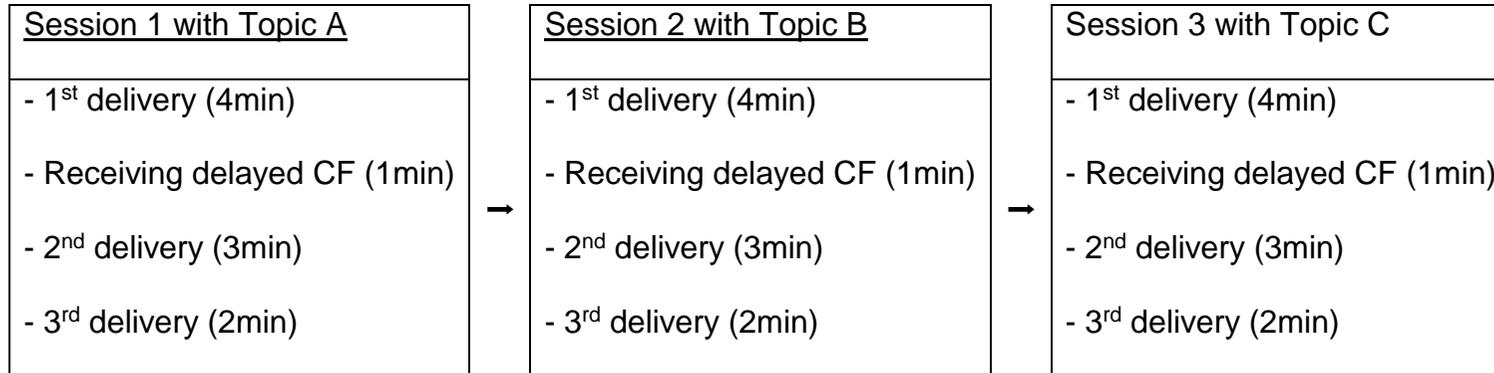
Fluency Enhancement Group



Accuracy Enhancement Group



Fluency + Accuracy Enhancement Group



Note. * indicates a speech sample used for T1; ** indicates a speech sample used for T2.

Figure 3. Summary of research design: control, FE, AE and FE+AE

3.4 Speech Analyses

3.4.1 Materials

To measure any change or improvement in the students' accuracy and fluency aspects of L2 speech over time, their task performance was audio-recorded. In total, 432 monologues were recorded (48 participants × 3 sessions × 3 deliveries) constituting 21 hours and 36 minutes of oral production data. Each of the 48 participants produced 27 minutes of speech. For the purpose of the analyses, I used the first two minutes from the first delivery in the first session (T1); and the first two minutes of the third delivery in the last session (T3) (48 talkers × T1 and T2 = 96 samples). The audio-recordings were then transcribed, coded and analyzed.

3.4.2 Measuring Fluency

Fluency in the broad sense is understood as general proficiency (Bui & Huang, 2016; De Jong & Perfetti, 2011; Housen & Kuiken, 2009; Housen, Kuiken & Vedder, 2012; O'Brien, 2014; Tavakoli, Campbell & McCormack, 2016; Wood, 2010) while in the narrow sense it deals with temporal measures such as speech rate, hesitations, and pausing (De Jong & Perfetti, 2011, Housen et al., 2012; O'Brien, 2014; Tavakoli et al., 2016). According to Segalowitz (2012, p. 240), "for most, the qualities that make speech fluent include fast speech, and the relative absence of undue hesitations, pausing, repetition, and self-repairs." The current study focused on the narrow sense of speed, breakdown and repair fluency. Following the precursor fluency studies (Bosker, Pinget, Quené, Sanders & De Jong, 2013; Housen & Kuiken; 2009; Lambert et al., 2016; Tavakoli & Skehan, 2005), each aspect of fluency was analyzed as follows:

- **Speed Fluency:** The aspect of speed fluency was assessed in terms of articulation rate by dividing the total number of syllables produced by phonation time. The phonation time was analysed by subtracting all the fillers (ah, oh, eh) and extensive silence (250ms) from the total length of each sample.
- **Breakdown Fluency:** Pausing behavior was assessed in terms of the frequency of filled and unfilled pauses by dividing the number of pauses by the total number of words. Following Tavakoli and Skehan's (2005) recommendation, I separately calculated breakdown fluency for pauses in the middle and end of clauses. Whereas the frequency of mid-clause pauses is assumed to indicate L2 learners' efficiency in linguistic encoding processes, the ratio of final-clause pauses is supposed to reflect L2 learners' conceptualization processes (Lambert et al., 2016).
- **Repair Fluency:** Repair was analysed in terms of the number of repetitions and the number of self-repairs. This factor is assumed to correspond to L2 learners' monitoring process (Lambert et al., 2016).

3.4.3. Measuring Accuracy

According to Yuan and Ellis (2003), accuracy concerns "the extent to which the language produced conforms to target language norms" (Yuan & Ellis, 2003, p. 2). Similarly, Housen and Kuiken (2009) defined accuracy as the ability to produce error-free speech. In order to assess accuracy, different measures have been used in different studies. Numerous scholars used overall measures of accuracy, such as percentage of error-free clauses (Ahmadian & Tavakoli, 2010; Ahmadian 2012b;

Bygate, 1999; Foster & Skehan, 1996; Sato & Lyster, 2012; Skehan & Foster, 1997, 1999; Tavakoli et al., 2016; Thai & Boers, 2016; Vercellotti, 2017; Wang, 2014; Yuan & Ellis, 2003;). More recently, Saito (2018) suggested new ways of overall accuracy judgments in which three different measures could be distinguished. The first measure was global accuracy which involved overall ease of understanding. The second measure was semantic accuracy involving the selection of appropriate words in contexts. The last one was morphosyntactic accuracy involving the accurate use of tense, aspects, agreement, plurality, and word order. In other studies, to trace changes in accuracy more precisely, the error frequency of a certain specific feature has been analysed via obligatory context analyses (Ahmadian, 2012b; Ellis et al., 2006; Van de Guchte et al., 2016; Yang & Lyster, 2010).

The present study employed overall measures of accuracy used in Saito (2018)'s study. The researcher read the transcript and assessed global accuracy, semantic accuracy and morphosyntactic accuracy on a 9-point scale. Besides, this study also investigated the effects of CF on the participants' accurate use of the target structure (English past tense) in obligatory contexts. To this end, any instances of past tense in obligatory contexts were grouped into either regular or irregular forms (See Appendix F for obligatory contexts for past tense use). It has been shown that L2 learners acquire irregular and regular forms at a different rate with the former being mastered more easily and promptly than the latter—a similar developmental pattern in first language acquisition (see Yang & Lyster, 2010).

3.5 Inter-Coder Reliability

To ensure inter-coder reliability, 20 out of 96 samples (20.83%) were randomly selected and coded independently by a second coder (see Appendix G).

Two-tailed Pearson correlation analyses demonstrated a high degree of reliability for articulation rate ($r = .955$, $n = 20$, $p < .001$), mid-clause pauses ($r = .992$, $n = 20$, $p < .001$), final-clause pauses ($r = .983$, $n = 20$, $p < .001$), repairs ($r = .811$, $n = 20$, $p < .001$), self-repetitions ($r = .904$, $n = 20$, $p < .001$), global accuracy ($r = .822$, $n = 20$, $p < .001$), semantic accuracy ($r = .839$, $n = 20$, $p < .001$), and morphosyntactic accuracy ($r = .869$, $n = 20$, $p < .001$). Besides, Cohen's kappa was used as a measure of inter-coder reliability for the coding of past tense forms as this procedure included categorical variables. The analyses demonstrated a high degree of reliability for regular past tense ($k = .890$, $n = 20$, $p < .001$), as well as irregular past tense ($k = .894$, $n = 20$, $p < .001$).

3.6 Aptitude Measures

The LLAMA test consists of four subtests measuring various domains of L2 aptitude. The entire session took approximately 30 minutes in the following order: LLAMA-D → B → E → F. The tests were automatically scored out of 70 for LLAMA-D and 100 for LLAMA-B, E and F.

LLAMA-D. This refers to a subtest that measures the ability to recognize novel or old items after listening to sound strings only once without any practice phase. To avoid learners' any intention of learning during the listening sessions, the participants completed this subtest as a part of a sound test.

After hearing a set of 10 words in an unfamiliar language unintentionally, test-takers participate in the test phase in which they will listen to these words again together with words that they have not heard before. Their task is to decide which words are repeated and which ones are new. Each time test-takers have a correct answer, they score five points. If they have a wrong answer, they lose five points.

LLAMA-B. This refers to a subtest that measures the ability to learn written forms of new vocabulary items by associating word strings and drawings (similar to the paired-associates test in MLAT). Unlike LLAMA-D, the participants were first explicitly told about the purpose of the test (i.e., vocabulary learning followed by recollection).

In the first phase, test-takers are presented with 20 objects on the main panel. When they click on each object, its name will be displayed. Within 120 seconds, they have to learn as many names as possible. Test-takers can click on the objects as many times as they wish in the time available and remember the set of paired associates. Participants are not allowed to take notes. In the test phase, each time test-takers click on the arrow, the name of each of the 20 objects will be displayed. Their task is to identify the correct object by clicking on it.

LLAMA-E. This refers to a subtest that measures the ability to learn new sound-symbol correspondence (phonemic coding ability) by associating sound strings and unfamiliar alphabetical symbols (similar to the phonetic script test in MLAT).

In the first phase, test-takers are presented with spellings of 24 syllables of a language on the main panel. They have to learn the spelling system of the language by clicking on the syllables and hear the sounds. They have 2 minutes in total to explore the language. Notetaking is allowed in this phase. In the test phase, each time test-takers click on the arrow, a sound is played and two possible spellings for a two-syllable word is displayed. They have to click on the spelling that they think is correct to score points. They score five points for each correct answer. In contrast, five points are deducted if they have a wrong answer.

LLAMA-F. This refers to a subtest that measures the ability to induce the grammatical rules of an unfamiliar language (similar to the grammatical sensitivity task in MLAT).

In the first phase, each time the test-takers click on a small button in the main panel, a picture and a short sentence describing it will be displayed. Their task is to work out the relation between the pictures and the sentences in the time available. They can take notes if they wish. In the test phase, each time they click the arrow, a picture and two sentences describing it will be displayed. One of the sentences is grammatical correct while the other is erroneous. Test-takers have to click on the sentence that they think is correct to score points. They are awarded five points for each correct answer. If they have a wrong choice, they lose five points.

In this chapter, I have thus far presented and justified the methodology used for conducting the current research project including information on participants' profiles, procedure for collecting data, research design, speech analysis, and measurement of foreign language aptitude. Furthermore, the results of correlation analysis for interrater reliability has also been summarized. In chapter 4, the results of the statistical analyses will be reported.

CHAPTER 4: DATA ANALYSIS

In this chapter, I aim to examine the extent to which the participants improved fluency (speed, breakdown, repair) and accuracy (regular, irregular) between T1 (1st delivery of 1st session) and T2 (3rd delivery of 3rd session) according to the four group conditions—Control (repetition only), Fluency Enhancement (4/3/2), Accuracy Enhancement (corrective feedback) and Fluency + Accuracy Enhancement. For each linguistic dimension (fluency, accuracy), a set of one-way ANOVAs was first conducted on T1 test scores to check whether there were statistically significant differences across the four groups. The analysis yielded no significant results, (articulation rate: $F(3,44) = .929, p = .435$; mid-clause pauses: $F(3,44) = 1.322, p = .279$; final-clause pauses: $F(3,44) = 2.029, p = .124$; self-repairs: $F(3,44) = .416, p = .743$; self-repetitions: $F(3,44) = .580, p = .632$; accuracy in the use of regular past tense: $F(3,44) = .379, p = .769$; accuracy in the use of irregular past tense: $F(3,44) = 1.605, p = .202$; global accuracy: $F(3,44) = .233, p = .873$; semantic accuracy: $F(3,44) = .692, p = .562$; morphosyntactic accuracy: $F(3,44) = .698, p = .558$). Based on the results obtained, it could be concluded that participants in the four groups were fairly equivalent in terms of their initial performance. Then, a set of two-way repeated-measures ANOVAs were performed with the participants' linguistic scores at T1/T2 as dependent variables and their group ID as independent variables. Where significant contrasts were found, a set of paired samples *t*-tests were run to determine if there were significant differences between T1 and T2 across different group conditions. The magnitude of effects was calculated and assessed in conjunction with Plonsky and Oswald's (2014) field-specific benchmarks ($d < 0.70$ for small; $0.70 \leq d < 1.00$ for medium; $1.00 \leq d$ for large effects).

4.1 Fluency

4.1.1 Speed Fluency

A two-factor ANOVA yielded a significant main effect of time, $F(1, 44) = 22.50$, $p = .000$, $\eta_p^2 = .338$, as well as a significant interaction between time and group, $F(3, 44) = 8.05$, $p = .000$, $\eta_p^2 = .354$. However, there was no significant main effect of group on articulation rate, $F(1, 44) = .82$, $p = .49$, $\eta_p^2 = .053$. A series of paired samples t -tests compared the participants' performance in T1 versus T2. As summarized in Table 6, the results showed that FE and FE+AE demonstrated significant improvement, but such within-group improvement was not found for Control and AE. The FE group's articulation rate enhanced from 2.68 words per minute at T1 (the first delivery of the first session) to 3.18 words per minute at T2 (the third delivery of the third session), $t(11) = -6.861$, $p < .001$, with a large effect size observed ($d = 1.67$). In the FE +AE group, the increase was from 2.59 at T1 to 3.06 at T2, $t(11) = -3.581$, $p = .004$, with a relatively large effect size ($d = 1.13$). These results suggested that adding fluency enhancement substantially contributed to an increase in articulation rate. Interestingly, the control group (without FE or AE) 's improvement did not reach marginal significance, $p = .175$ although they also repeated the task three times.

Table 6: Group Means, Standard Deviations, Significant Improvement and Effect Size for Articulation Rate

Group	T1		T2		<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Control	2.69	.35	2.85	.39	.175	0.43
FE	2.68	.26	3.18	.34	< .001*	1.67
AE	2.82	.27	2.69	.36	.217	0.41
FE+AE	2.59	.43	3.06	.40	.004*	1.13

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

4.1.2 Breakdown Fluency

The breakdown fluency dimension of the participants' performance was analysed from two different angles: the ratio of pause between clause boundaries versus at the end of clauses. For the mid-clause pause category, a two-way ANOVA showed a significant main effect of time $F(1, 44) = 12.66, p = .001, \eta_p^2 = .223$ and a significant interaction between time and group, $F(3, 44) = 4.18, p = .011, \eta_p^2 = .222$. However, there was no significant main effect of group $F(3, 44) = .657, p = .583, \eta_p^2 = .043$. As follow-up analyses, a set of paired sampled *t*-tests were performed to further explore which group significantly reduced the number of mid-clause pauses over time (T1 to T2). As summarized in Table 7, it was only the FE group that demonstrated significant improvement over time ($p = .002$) with medium effects ($d = 0.87$). In addition, the FE +AE group's improvement reached marginal significance, $p = .017$ with medium effects ($d = 0.86$). These results suggested that fluency

enhancement reduced mid-clause pausing significantly (FE) but engaging in both fluency and accuracy enhancement (FE+AE) at the same time decreased the positive effects to some extent.

Table 7: Group Means, Standard Deviations, Significant Improvement and Effect Size for Mid-clause Pause Ratio

Group	T1		T2		<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Control	.155	.071	.135	.073	.381	0.28
FE	.159	.110	.081	.064	.002*	0.87
AE	.105	.026	.119	.056	.447	0.32
FE+AE	.150	.072	.094	.058	.017†	0.86

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

When it comes to the final-clause pause dimension, ANOVA showed that there was a significant main effect of time $F(1, 44) = 58.5$, $p = .000$, $\eta_p^2 = .571$ and a significant interaction effect of time and group, $F(3, 44) = 4.32$, $p = .009$, $\eta_p^2 = .227$. In contrast, there was no significant main effect of group $F(3, 44) = .374$, $p = .772$, $\eta_p^2 = .025$. According to the results of paired samples *t*-tests (summarized in Table 8), the number of final-pause clauses significantly declined over time among the participants of the FE and FE + AE groups with large effects ($d = 1.0$ - 2.0). In addition, the control group's improvement reached marginal significance, $p = .015$ with large effects ($d = 1.05$). The AE group was the only group that did not show any significant or marginal improvement ($p = .370$). These results suggested that

4/3/2 ACTIVITY REVISITED

138

increasing time pressure (FE and FE+AE) in the 4/3/2 activity significantly contributes to the decrease in final-clause pausing more than task repetition alone (Control).

Table 8: Group Means, Standard Deviations, Significant Improvement and Effect Size for Final-clause Pause Ratio

Group	T1		T2		<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Control	.212	.049	.159	.052	.015†	1.05
FE	.228	.052	.139	.035	< .001*	2.01
AE	.202	.058	.180	.072	.370	0.34
FE+AE	.252	.054	.149	.042	< .001*	2.14

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

4.1.3 Repair Fluency

The results for repair fluency were presented separately as to the ratio of self-repairs and repetitions. With regard to the frequency of self-repairs, a two-way ANOVA did not yield any significant main effects of time $F(1, 44) = 6.56, p = .014, \eta_p^2 = .130$ nor group $F(3, 44) = .358, p = .784, \eta_p^2 = .024$. Additionally, there was no significant interaction effect of time and group $F(3, 44) = 1.37, p = .264, \eta_p^2 = .085$. As summarized in table 9, no significant changes between T1 and T2 were found for all the four groups; notably, however, the FE group demonstrated marginal improvement ($p = .038$) with medium effects ($d = 0.79$), suggesting that fluency enhancement made some tangible contribution to the participants' repair fluency.

Table 9: Group Means, Standard Deviations, Significant Improvement and Effect Size for Repair Ratio

Group	T1		T2		<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Control	.014	.009	.012	.011	.533	0.20
FE	.014	.014	.006	.003	.038†	0.79
AE	.010	.007	.011	.009	.912	0.12
FE+AE	.016	.017	.007	.006	.065	0.70

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

Different results were found for the frequency of repetitions. A two-way ANOVA yielded significant main effects of time $F(1, 44) = 9.92, p = .003, \eta_p^2 = .184$. However, a two-way ANOVA did not generate significant main effects of group $F(3, 44) = .283, p = .838, \eta_p^2 = .019$ nor interaction effects of time and group, $F(3, 44) = 2.09, p = .116, \eta_p^2 = .125$. According to Table 10, a series of paired samples *t*-tests did not detect statistically significant changes over time in any contexts ($p > .012$). Again, the improvement of the FE and FE+AE groups was considered marginally significant ($p < .05$). These results suggested that fluency enhancement tangibly contributed to the participants' repair fluency by reducing the number of self-repetitions.

Table 10: Group Means, Standard Deviations, Significant Improvement and Effect Size for Repetition Ratio

Group	T1		T2		<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Control	.029	.031	.025	.023	.626	0.18
FE	.038	.031	.015	.012	.020†	1.00
AE	.034	.026	.032	.034	.826	0.06
FE+AE	.049	.058	.020	.024	.026†	0.65

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

4.2 Accuracy

4.2.1 Accuracy in the use of the English past tense

Regular past tense. A two-factor ANOVA demonstrated a significant effect of time $F(1, 44) = 7.99, p = .007, \eta_p^2 = .154$. However, there was no significant main effect of group $F(3, 44) = 3.12, p = .036, \eta_p^2 = .175$ or a significant interaction between time and group in terms of targetlike use $F(3, 44) = 2.05, p = .121, \eta_p^2 = .122$. As shown in Table 11, the results of multiple comparisons further revealed that the AE-only group significantly enhanced their targetlike use of regular past tense over time ($p = .011$) with medium effects ($d = 0.91$). According to these results, focusing learners on a feature after the first performance (adding AE) resulted in them being more careful with that feature (regular past tense) on an immediate repetition. However, when AE was used together with FE, the increase in accuracy failed to reach significance.

Table 11: Group Means, Standard Deviations, Significant Improvement and Effect Size for Targetlike Use of Regular Past Tense

Group	T1		T2		p	d
	M	SD	M	SD		
Control	18.86	18.22	28.18	26.59	.376	0.41
FE	23.24	22.74	20.92	16.07	.787	0.12
AE	28.16	33.89	55.99	27.17	.011 *	0.91
FE+AE	18.49	24.52	33.08	13.27301	.065	0.74

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

Irregular past tense. A two-way ANOVA yielded significant main effects of time $F(1, 44) = 33.81, p = .000, \eta_p^2 = .434$ and a significant interaction effect of time and group $F(3, 44) = 14.72, p = .000, \eta_p^2 = .501$. However, there was no significant main effects of group $F(3, 44) = .614, p = .610, \eta_p^2 = .040$. As summarized in table 12, the results of multiples comparisions further revealed that the two groups who received AE (AE-only, FE + AE) significantly enhanced their targetlike use of irregular past tense over time with large effects ($d > 1.0$) suggesting that accuracy enhancement resulted in positive changes in accuracy in the use of irregular past tense.

Table 12: Group Means, Standard Deviations, Significant Improvement and Effect Size for Targetlike Use of Irregular Past Tense

Group	T1		T2		p	d
	M	SD	M	SD		
Control	60.84	23.11	54.66	29.60	.281	0.23
FE	52.21	27.53	54.24	34.17	.804	0.07
AE	44.81	21.51	78.23	20.12	< .001*	1.60
FE+AE	42.40	17.53	86.37	9.80	< .001*	3.10

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

4.2.2 General accuracy

Global accuracy. A two-factor ANOVA revealed a significant main effect of time $F(1, 44) = 7.157, p = .010, \eta_p^2 = .140$. However, there was neither a significant interaction between time and group, $F(3, 44) = 2.03, p = .123, \eta_p^2 = .122$ or significant main effects of group $F(3, 44) = 1.60, p = .204, \eta_p^2 = .098$. According to the results of paired samples t -tests (summarized in Table 13), the participants who received AE (AE-only, FE+AE) showed marginal improvement over time ($p < .05$) with medium effects ($d = 0.71, 0.81$ respectively), which could be attributable to the fact that accuracy enhancement (AE, FE+AE) has positive effects on global accuracy development.

Table 13: Group Means, Standard Deviations, Significant Improvement and Effect Size for Global Accuracy

Group	T1		T2		p	d
	M	SD	M	SD		
Control	6.25	1.29	6.33	.78	.795	0.08
FE	6.41	1.31	6.42	.79	1.000	0.01
AE	6.67	.98	7.25	.62	.027†	0.71
FE+AE	6.42	1.31	7.25	.62	.025†	0.81

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

Semantic accuracy. A two-way ANOVA showed a significant main effect of time $F(1, 44) = 19.07, p = .000, \eta_p^2 = .302$ and a significant interaction between time and group, $F(3, 44) = 5.15, p = .004, \eta_p^2 = .260$. However, there was no significant main effect of group $F(3, 44) = 1.82, p = .159, \eta_p^2 = .110$. According to the results of paired samples t -tests (summarized in Table 14), the FE+ AE group is the only group that showed significant improvement over time ($p < .001$) with large effects ($d = 1.61$), suggesting that engaging in FE and AE simultaneously helps learners improve their semantic accuracy.

Table 14: Group Means, Standard Deviations, Significant Improvement and Effect Size for Semantic Accuracy

Group	T1		T2		p	d
	M	SD	M	SD		
Control	6.50	1.09	6.83	.58	.220	0.38
FE	6.83	1.11	6.75	.87	.586	0.08
AE	7.00	1.04	7.67	.78	.071	0.72
FE+AE	6.50	.90	7.75	.62	< .001*	1.61

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

Morphosyntactic accuracy. A two-way ANOVA yielded significant main effects of both time $F(1, 44) = 27.04, p = .000, \eta_p^2 = .381$ and group $F(3, 44) = 4.33, p = .009, \eta_p^2 = .228$. Additionally, there was a significant interaction effect of time and group $F(3, 44) = 9.30, p = .000, \eta_p^2 = .388$. As summarized in table 15, the participants who received AE (AE-only, FE+AE) showed significant improvement over time ($p < .012$) with large effects ($d > 1.0$). These results indicated that corrective feedback promoted morphosyntactic accuracy development.

Table 15: Group Means, Standard Deviations, Significant Improvement and Effect Size for Morphosyntactic Accuracy

Group	T1		T2		p	d
	M	SD	M	SD		
Control	5.58	.90	5.75	.87	.551	0.19
FE	6.08	.79	5.92	.79	.339	0.21
AE	5.92	1.16	7.25	.87	.006 *	1.30
FE+AE	5.75	.62	7.17	.58	< .001*	2.36

Note. * indicates $p < .012$ (Bonferroni corrected); † indicates $p < .05$

4.3 Cognitive Individual Differences (Aptitude Scores)

The descriptive statistics of aptitude scores by groups were summarized in table 16. A set of one-way ANOVAs was conducted on the language aptitude scores to check whether participants in the four groups were equivalent in terms of language aptitude. The analysis yielded no significant results, (LLAMA-D: $F(3,44) = .696, p = .560$; LLAMA-B: $F(3,44) = .887, p = .455$; LLAMA-E: $F(3,44) = 1.101, p = .359$; LLAMA-F: $F(3,44) = 1.170, p = .332$).

Table 16: Descriptive Statistics of Aptitude Scores by Groups

	LLAMA-D		LLAMA-B		LLAMA-E		LLAMA-F	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
AE	30.42	15.44	34.58	10.76	45.00	29.39	51.67	21.67
FE	22.08	11.57	42.92	13.05	58.75	14.48	62.50	15.00
AE	24.17	15.79	42.92	20.50	60.00	21.74	60.00	12.06
FE + AE	25.83	15.64	41.67	12.67	56.67	22.70	52.92	17.64

Table 17 summarized 48 EFL learners' aptitude scores according to the four subtests (LLAMA-D, B, E and F). As can be observed here, the participants achieved higher scores on the LLAMA E and F ($M > 55$) compared to LLAMA B and LLAMA D ($M = 25-40$).

Table 17: Descriptive Statistics of Aptitude Scores by Subtests

	<i>M</i>	<i>SD</i>	Range	
			Min	Max
<u>Language aptitude</u>				
LLAMA-D (75 points)	25.6250	14.57282	.00	55.00
LLAMA-B (100 points)	40.5208	14.66794	15.00	90.00
LLAMA-E (100 points)	55.1042	22.77292	10.00	100.00
LLAMA-F (100 points)	56.7708	17.05715	.00	90.00

Next, the participants' LLAMA scores were compared via a set of correlation analyses. As shown in Table 18, the results did not yield any significant correlations, suggesting that LLAMA D, LLAMA B, LLAMA E and LLAMA F tapped into four different/independent constructs of L2 aptitude: sound sequence recognition, associative memory, phonemic coding, and language analytic ability.

Table 18: Interrelationships between the Four Subtests (D, B, E, F) in LLAMA

	LLAMA-B		LLAMA-E		LLAMA-F	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
LLAMA-D	.001	.995	.046	.755	-.159	.282
LLAMA-B			.079	.591	.081	.583
LLAMA-E					.245	.094

In order to scrutinize the extent to which the participants' gain scores (fluency accuracy) were related to their aptitude scores (LLAMA-D, LLAMA-B, LLAMA-E, LLAMA-F), a set of partial correlation analyses were performed. Here, the participants' posttest scores were used as dependent variables; and their pretest scores were statistically partialled out. In what follows, the acquisition-aptitude link was reported according to four different task conditions: Control, FE, AE and FE+AE. Given the size of each group ($n = 12$) and in conjunction with Plonsky and Oswald's (2014) strong call for adopting a field-specific benchmark with regard to interpretation of correlation analyses, the analyses below focused only on strong associations ($r > .70$)

4.3.1. Aptitude × Control Group

As illustrated in Table 19, when participants simply engaged in task repetition without any AE or FE, the results suggested that the extent to which they can improve accuracy and fluency could be associated with certain aptitude profiles. When learners showed high-level associative memory (LLAMA-B), they demonstrated marginal development in accuracy (global accuracy). Also, they likely developed their fluency (articulation rate and mid-clause pauses) to certain degrees. However, these contrasts were below .70, suggesting that the effects of aptitude on task repetition may not be clearly observed.

Table 19: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the Control Group

	LLAMA B		LLAMA D		LLAMA E		LLAMA F	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Articulation rate	-.650	.030 †	-.233	.490	.384	.244	-.322	.335
Midclause	.627	.039 †	.124	.717	-.388	.238	-.044	.898
Final Clause	.460	.155	-.088	.797	-.569	.068	.030	.930
Repair	.383	.245	-.312	.350	-.486	.129	.148	.663
Repetition	.048	.888	-.001	.997	-.598	.052	.092	.788
Regular	-.194	.568	-.378	.252	-.418	.201	.191	.575
Irregular	-.521	.100	.036	.916	-.481	.134	.323	.333
Global Accuracy	-.642	.033 †	-.007	.984	.128	.707	.118	.729
Semantic Accuracy	.199	.557	-.223	.510	-.249	.460	.072	.833
Morphosyntactic Accuracy	-.347	.295	-.284	.398	-.580	.061	.224	.509

Notes: T2 scores were used as dependent variables and T1 scores were used as a covariate. * indicates strong correlations ($r > .70$); † indicates marginal significance at $p < .05$.

4.3.2. Aptitude × FE

According to the group analyses reported above, the results showed that when participants completed task repetition under increasing time pressure (4/3/2), their improvement in accuracy was somewhat limited, arguably because FE may use up L2 learners' cognitive resources for the development of L2 fluency (rather than accuracy). However, the results of the partial correlation analyses (summarized in Table 20) demonstrated that certain participants with greater associative memory (LLAMA-B) resulted in greater improvement in repair aspects of fluency (repetition) and semantic accuracy. In addition, learners with high-level language analytic ability (LLAMA-F) demonstrated marginal development in the use of regular past tense. However, these contrasts were below .70. According to Plonsky and Oswald (2014), the strength of the relationship could not be considered strong.

Table 20: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the FE Group

	LLAMA B		LLAMA D		LLAMA E		LLAMA F	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Articulation rate	-.063	.854	.108	.752	-.271	.420	-.163	.633
Midclause	.277	.409	-.275	.413	.042	.901	-.031	.928
Final Clause	.003	.994	-.137	.687	-.223	.510	-.031	.927
Repair	.230	.495	.284	.397	.107	.754	-.387	.240
Repetition	-.628	.039 †	-.338	.310	.002	.995	-.266	.430
Regular	.050	.883	.165	.628	.583	.060	.670	.024 †
Irregular	.177	.602	.060	.860	-.010	.977	.082	.812
Global Accuracy	-.050	.885	-.372	.260	-.045	.897	.282	.401
Semantic Accuracy	-.617	.043 †	-.179	.599	-.388	.239	.182	.592
Morphosyntactic Accuracy	-.256	.447	-.337	.311	-.525	.097	.052	.880

Notes: T2 scores were used as dependent variables and T1 scores were used as a covariate. * indicates strong correlations ($r > .70$); † indicates marginal significance at $p < .05$.

4.3.3. Aptitude × AE

The group analyses earlier pointed out that participants' fluency development could be limited, when they received CF during task repetition activities. According to the results of the partial correlation analyse (summarized in Table 21), it is interesting to note that their aptitude profiles did not facilitate learners' accuracy and fluency development in any contexts ($p > .05$).

Table 21: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the AE Group

	LLAMA B		LLAMA D		LLAMA E		LLAMA F	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Articulation rate	.296	.377	.215	.526	-.048	.889	-.315	.345
Midclause	-.033	.924	-.373	.258	.186	.583	-.257	.446
Final Clause	-.039	.909	-.461	.154	.188	.579	-.060	.860
Repair	-.006	.987	-.509	.110	-.076	.823	-.232	.493
Repetition	.222	.513	-.444	.171	-.157	.646	.082	.809
Regular	.103	.763	.494	.122	-.330	.321	.051	.882
Irregular	-.151	.658	.273	.417	.261	.438	.318	.340
Global Accuracy	.283	.400	.033	.923	.094	.783	.050	.883
Semantic Accuracy	.102	.765	.130	.704	.089	.795	.223	.510
Morphosyntactic Accuracy	.310	.353	.107	.755	-.091	.791	-.184	.588

Notes: T2 scores were used as dependent variables and T1 scores were used as a covariate. * indicates strong correlations ($r > .70$); † indicates marginal significance at $p < .05$.

4.3.4. Aptitude × FE+AE

The results of the group analyses showed that participants generally improved their accuracy and fluency performance when they received both FE and AE. However, their improvement was unclear in regard to certain aspects of fluency and accuracy related to linguistic coding—i.e., mid-clause pauses and regular past tense. As illustrated in Table 22, in terms of fluency development, the results of partial correlation analyses showed that participants with greater phonemic coding (LLAMA-E) improved the relatively difficult aspect of fluency development (mid-clause pauses). Interestingly, language analytic ability (LLAMA-F) appeared to predict what kind of L2 learners could further improve their articulation rate. In addition, participants with greater associative memory (LLAMA-B) resulted in greater improvement in repair aspects of fluency (repair).

In terms of accuracy development, associative memory (LLAMA-B) seemed to play a key role in determining the extent to which participants could show further improvement in the use of irregular past tense. Interestingly, some marginal associations were found between participants' incidental learning aptitude (LLAMA-D) and semantic accuracy development.

In consultation with Plonsky and Oswald's (2014) benchmark, however, the relationship between associative memory (LLAMA-B) and repair fluency could be considered as the only strong correlation ($r = -.705$).

Table 22: Results of Partial Correlations between Aptitude, and Fluency and Accuracy Scores of the FE + AE Group

	LLAMA B		LLAMA D		LLAMA E		LLAMA F	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Articulation rate	.422	.196	.473	.142	.484	.132	.694	.018 †
Midclause	-.511	.108	-.484	.131	-.673	.023 †	-.488	.128
Final Clause	-.128	.708	-.599	.052	-.173	.611	-.346	.297
Repair	-.705 *	.015 †	.101	.767	-.259	.441	-.480	.135
Repetition	-.190	.577	.064	.852	.088	.797	-.077	.822
Regular	.127	.710	-.290	.388	.175	.606	-.287	.392
Irregular	.602	.050 †	-.441	.174	.017	.960	.016	.963
Global Accuracy	.267	.427	-.173	.612	.072	.834	-.017	.960
Semantic Accuracy	.189	.579	-.668	.025 †	.223	.509	-.060	.861
Morphosyntactic Accuracy	.328	.324	-.132	.699	.357	.282	.104	.760

Notes: T2 scores were used as dependent variables and T1 scores were used as a covariate. * indicates strong correlations ($r > .70$); † indicates marginal significance at $p < .05$.

In this chapter, I have presented the results of the analyses in three main sections. The first section was concerned with fluency development (speed, breakdown, and repair). The second section addressed accuracy development including accuracy in the use of the past tense (irregular and regular) and general accuracy (global, semantic and morphosyntactic accuracy). In the last section, the correlation between aptitude, and fluency and accuracy scores were reported. In chapter 5, these results will be discussed.

CHAPTER 5: DISCUSSION

This chapter discusses the results in relation to the predictions and findings from previous studies in four major sections. The first section deals with task repetition. The second section is dedicated to accuracy enhancement. The third section focusses on incorporating both fluency and accuracy enhancement. The final section examines the role of cognitive individual differences in determining the effectiveness of task repetition, fluency enhancement and accuracy enhancement.

To date, previous literature has persuasively shown that adding increasing time pressure to task repetition can greatly facilitate L2 fluency development (i.e., 4/3/2 activity). However, the impact of fluency enhancement on L2 accuracy development remains unclear (e.g., Thai & Boers, 2016). In this study, I attempted to remedy this issue by incorporating the delayed corrective feedback technique (drawing learners' attention for particular linguistic form) as a form of accuracy enhancement (e.g., Li et al., 2016). To examine the extent to which the findings could be also related to participants' cognitive individual differences, their aptitude scores (associative memory, phonemic coding, language analytic ability, and sound sequence recognition) were linked to their longitudinal development of L2 fluency and accuracy. The results are summarized in Table 23.

In what follows, I provide an in-depth discussion on the four crucial components of task-based language learning and teaching in the following order: (a) task repetition; (b) increasing time pressure; (c) corrective feedback and (d) aptitude effects.

Table 23: Summary of Overall Results of Fluency and Accuracy Analyses

	A. Fluency					B. Accuracy				
	Articulation rate	Mid-clause pauses	Final- clause pauses	Repair	Self- repetition	Past tense		Overall		
						Regular	Irregular	Global	Semantic	Morphosyntactic
Control	x	x	Δ	x	x	x	x	x	x	x
FE	○	○	○	Δ	Δ	x	x	x	x	x
AE	x	x	x	x	x	○	○	Δ	x	○
FE+AE	○	Δ	○	x	Δ	x	○	Δ	○	○
Aptitude effects	Δ (LLAMA-B- F)	Δ (LLAMA-B- E)	x	○ (LLAM A-B)	Δ (LLAMA- B-F)	Δ (LLAM A-F)	Δ (LLAMA-B)	Δ (LLAMA -B)	Δ (LLAMA-B- D)	x

Note. ○ indicates significant improvement or strong effects; Δ indicates marginal improvement or medium effects; x indicates insignificant improvement or small effects.

5.1 Task Repetition

Overall, the control group (task repetition only) participated in three different task repetition sessions. In each session, they provided a monologue on one of the three different topics (favorite movie, last summer vacation, 16th birthday) three times under constant time conditions (3, 3, 3 minutes). As summarized in Table 23, the control group marginally enhanced only one aspect of breakdown fluency (final-clause pause ratio) when their speech performance was compared at T1 (the first delivery of Session 1 on topic A) and T2 (the third delivery of Session 3 on topic C). However, the other dimensions of language performance (speed/breakdown fluency and accuracy) remained unchanged over time. The results here are consistent with those of many task repetition studies in that (a) task repetition in the time-constant condition (i.e., 3/3/3) can lead to some tangible improvement in fluency (Ahmadian & Tavakoli, 2010; Boers, 2014, Thai & Boers, 2016); and (b) the transferability of the gains to new topics remains unclear (e.g., De Jong & Perfetti, 2011).

From a theoretical standpoint, the number of final-clause pauses is considered to reflect L2 learners' conceptualization processes (i.e., what to say) (Kormos, 2006). As shown in recent fluency studies, much learning tends to take place in this specific fluency dimension (final-clause pause ratio) especially at the beginning of L2 fluency development in instructed (e.g., Lambert et al., 2016) and naturalistic (e.g., Saito, Ilkan, Magne, Tran, & Suzuki, 2018) settings. In conjunction with the results of the current study, this indicates that when L2 learners engage in task repetition only, their gains could be exclusively limited to the relatively easy aspect of L2 fluency learning (i.e., reduction in the final-clause pauses).

Increasing Time Pressure (Fluency Enhancement)

In the current study, I aimed to investigate how to help L2 learners further enhance fluency during task repetition by testing the effect of increasing time pressure as a form of fluency enhancement (Nation, 1989). According to the results, repeating a task with an increasing amount of time pressure considerably stimulated more fluent speech across all the relevant dimensions (speed, breakdown and repair). However, the participants who received FE only did not improve their accurate use of L2 English. These findings are in line with previous literature on 4/3/2 activities (Boers, 2014; De Jong & Perfetti, 2011; Nation, 1989; Thai & Boers, 2016). These results were also compatible with those of the online-planning studies (Ahmadian & Tavakoli, 2010; Ahmadian, 2012b; Yuan & Ellis, 2003). All in all, I echoed these researchers who have argued that when learners are under time pressure to perform a task more rapidly, they are to prioritize meaning conveyance over linguistic accuracy. This relative focus on meaning rather than form may result in L2 fluency, but not accuracy development—a weakness of the 4/3/2 activity. These results support Levelt (1989)'s and de Bot (1992)'s models. Based on these models, prior to the articulation stage of speech production, learners need time to attend to the formulation and monitoring of their internal speech to formulate accurate structures. When they are under increasing time pressure, they do not have enough time to attend to form which resulted in low accuracy. At the same time, these are also evidences of the trade-off effects. When meaning was prioritized over form, fluency came at the cost of accuracy.

5.2 Corrective Feedback (Accuracy Enhancement)

The current study is one of the first attempts to test the role of corrective feedback in the context of task repetition. In the previous task repetition literature, it has been shown that L2 learners are likely primed to speak fluently rather than accurately during task repetition (Ahmadian, 2012a). Here, I attempted to draw their attention to accuracy via the delayed CF technique, wherein I gave metalinguistic information to participants while they were repeating a task. Compared to the control group, who engaged in task repetition only (demonstrating small gains in fluency but not in accuracy), the AE group significantly improved accuracy in terms of their use of both irregular and regular forms. Also, they significantly improved morphosyntactic accuracy and showed marginal improvement in global accuracy, but not fluency. The effects of CF on L2 grammar learning concurred with the vast amount of the CF literature that CF can help L2 learners attend to linguistic form during meaningful discourse and thus promote the process and product of SLA (Li, 2010).

However, my study also brought to light one disadvantage of adding CF to task repetition. The participants in the AE group exhibited no change in all the dimensions of their fluency performance, despite the fact that they repeated the same task three times. Note that there was no significant improvement even in the AE group's final-clause pause ratio (a relatively easy aspect of L2 fluency). The findings here indicated that providing CF (accuracy enhancement) may offset the benefit of task repetition (i.e., fluency development), arguably because beginner-to-intermediate L2 learners, such as those in the current study (Vietnamese EFL students), have difficulty attending to both fluency and accuracy of L2 speech, while at the same time they are using the target language for meaningful purposes. This could arguably be due to the limited amount of

cognitive resources that these inexperienced, Vietnamese EFL learners could afford to use/handle while using the L2 for communicative purposes (Skehan, 1998, 2014; Skehan & Foster, 2008)—a crucial point in theory and practice that I revisit in the next subsection.

5.3 Incorporating both Fluency and Accuracy Enhancement

To remedy the weaknesses of FE and AE, I finally tested to what degree and how incorporating both FE and AE into task repetition (i.e., providing CF during the 4/3/2 activity) could impact L2 fluency and accuracy development *simultaneously*. The results showed that the FE+AE group significantly enhanced not only all dimensions of L2 fluency (speed, breakdown, repair), but also refine the accurate use of irregular (but not regular) past tense and their overall accuracy (global, semantic and morphosyntactic).

In a broader sense, the findings here lend empirical support to the previous task-based instruction studies which emphasized the pedagogical potential of focus on form before students proceed to any types of task repetition activities (Hawkes, 2011; Van de Guchte et al., 2016). In the current study, the participants in the FE+AE group engaged in the 4/3/2 activities and received delayed CF on English past tense, resulting in fluency and accuracy improvement. On the one hand, the participants were encouraged to enhance all aspects of fluency (speeding up while reducing the number of pauses, repairs and repetitions) thanks to increasing time pressure, as their burden on conceptualization declined through task repetition (Thai & Boers, 2016). On the other hand, delayed CF aided the participants to pay attention to form, even though they concurrently worked on fluency (Li, 2010).

In a narrower sense, however, my discussion here (more is better) needs to be considered tentative and interpreted with caution. Notably, the findings presented in this paper also hinted not only the potentials of combined accuracy and fluency activities, but also the limits of such composite approach. If we closely look at the details of my fluency and accuracy analyses, the results indeed bring to light the trade-off relationship between fluency and accuracy enhancement and development—i.e., fluency enhancement risks accuracy development (Skehan, 1998, 2014; Skehan & Foster, 2008).

When comparing the performance of those who received either FE or AE (FE-/AE-only) versus both FE and AE activities (FE+AE), it is important to remember the following slightly different improvement patterns. The FE-only and AE-only groups greatly developed all the dimensions of fluency (speed, breakdown, repair) and accuracy (accuracy in the use of regular and irregular past tense and general accuracy), respectively. In contrast, the FE+AE group lacked significant improvement in terms of the relatively difficult aspects of L2 fluency (mid-clause pause ratio) and accuracy (regular forms) development (see Table 23).¹

In previous L2 fluency research, this particular dimension of breakdown fluency (mid-clause pause ratio) is thought to relate to L2 learners' linguistic encoding processes (selecting appropriate linguistic forms for the intended message) (Kormos,

¹ In addition, if we compare the AE Group (adding CF to task repetition) vs. the Control Group (task repetition only), we can find another evidence for the trade-off relationship. Whereas the participants who received AE resulted in a great deal of accuracy improvement in both irregular and regular forms, these learners seemed to reveal somewhat limited gains in fluency development at a fine-grained level. As pointed out in the previous subsection, although the control group (task repetition only) showed their marginal improvement in the relatively easy aspect of L2 fluency development (i.e., final-clause pause ratio), the AE group did not show such gains in any dimensions of their fluency performance.

2006). Whereas L2 learners quickly reduce the number of pauses at clause-final positions (relevant to conceptualization), they may need more experience/practice to demonstrate significant, tangible and robust change in their mid-clause pause ratio (e.g., Lambert et al., 2016 for more than three times of repetition; Saito et al., 2018 for more than 5 years of naturalistic exposure). When it comes to the L1 acquisition of English irregular vs. regular forms, it has been argued that different processing systems are used with the former being linked to exemplar-based learning and the latter to rule-based learning (Pinker & Ullman, 2002). In the context of L2 learning, there is some evidence (a) that L2 learners of English acquire irregular prior to regular past tense forms in naturalistic (Bardovi-Harlig & Comajoan, 2008); and instructed settings (Yang & Lyster, 2010).

Following this line of thought, I would like to argue that AE+FE can facilitate L2 learners' fluency and accuracy learning only at a broad level; however, I have yet to know whether FE+AE could be the most optimal option for promoting all dimensions of SLA. When I carefully analyzed the participants' L2 fluency from multiple angles, I did find some evidence of trade-off: adding AE seemed to elicit more attention to linguistic encoding, which could in turn stop learners from reducing mid-clause pauses (thought to relate to linguistic encoding) and enhancing the rule-based learning (the acquisition of regular past tense forms).

5.4 Roles of Cognitive Individual Differences

In line with Plonsky and Oswald's (2014) field-specific benchmark, the results of the partial correlation analyses only identified strong correlations ($r = .70$) in terms of associative memory (LLAMA-B) and the reduction of self-repairs, when learners

participated in 4/3/2 activity and received CF (FE+AE). Besides, improvement in accuracy and fluency was also associated with certain aptitude profiles in other cases although these contrasts were below .70. For example, (a) associative memory (LLAMA-B) and articulation rate, mid-clause pausing and global accuracy when learners engaged in task repetition only; (b) associative memory (LLAMA-B) and self-repetitions and semantic accuracy as well as language analytic ability (LLAMA-F) and accuracy in the use of regular past tense when learners completed the task under increasing time pressure (FE only); and (c) associative memory (LLAMA-B) and the acquisition of irregular past tense; language analytic ability (LLAMA-F) and articulation rate; phonemic coding (LLAMA-E) and the reduction of mid-clause pauses; sound sequence recognition ability (LLAMA-D) and semantic accuracy when learners engaged in both 4/3/2 and delayed CF treatment (FE+AE).

Taken together, my tentative argument is that three different types of aptitude (associative memory, phonemic coding, and language analytic ability) may differentially help promote L2 fluency and accuracy development. On the one hand, certain aspects of participants' L2 fluency development (mid-clause ratio, and self-repairs) may require high-level aptitude in associative memory. This is the case especially when task conditions involve both fluency and accuracy enhancement—the most cognitively demanding task condition for beginner-to-intermediate Vietnamese EFL learners. When receiving L2 input, L2 learners with greater phonemic coding and associative memory are hypothesized to have the better capacity to engage in linguistic encoding processes (directly relevant to mid-clause pause ratio, and self-repair ratio in L2 fluency) in a more effective and efficient manner.

On the other hand, the other two aptitude constructs—associative memory and language analytic ability—may play a key role in boosting the effectiveness of the treatment when it involves fluency enhancement (4/3/2). Language analytic ability could impact on L2 accuracy development (regular past tense), as such cognitive ability is to help L2 learners make careful morphosyntactic analyses. Finally, associative memory seems to promote both fluency (self-repetition frequency) and accuracy (semantic) development when learners participate in FE only (4/3/2).

My claim here (the different effects of aptitude on L2 acquisition) is in line with the recent view on the interaction between different dimensions of aptitude and acquisition, proposed by a number of scholars (Li, 2013, 2015, 2016; Saito, 2017; Saito et al., 2018; Skehan, 2015, 2016). According to this position, different constructs of aptitude are uniquely tied to different stages of acquisition. For example, Skehan (2016) claimed that phonemic coding is specifically related to input processing; language analytic ability to pattern identification; and associative memory to automatization. Li (2013) pointed out that aptitude is crucial especially when L2 learners attempt to acquire relatively difficult, complex and perceptually non-salient linguistic features. In contrast, L2 learners can acquire relatively easy, simple and perceptually salient linguistic features as long as they practice them sufficiently.

As shown in the current study, L2 fluency and accuracy development is mainly driven by types of tasks (task repetition with or without accuracy and fluency enhancement). However, L2 learners can make the most of such experience resulting in more gains according to their cognitive profiles (phonemic coding for fluency;

associative memory for both fluency and accuracy; language analytic ability for both fluency and accuracy; and sound sequence rescognition ability for semantic accuracy).

In this chapter, I have interpreted the results and discussed how they are related to the existing body of research in the field in four main sections. The first section was concerned with the roles of task repetition and increasing time pressure. The second section covered the facilitative effects of corrective feedback. The third section focused on the combined effects of fluency and accuracy enhancement on L2 development. The last section delved into the roles of cognitive individual differences. In the last chapter, I will summarize the key findings, address limitations of the present study and offer recommendations for future research.

CHAPTER 6: CONCLUSION

This chapter includes three main sections. The first section provides a summary of the key findings. The second section discusses limitations of the present studies and addresses recommendations for future research. The last section is dedicated to theoretical and pedagogical implications.

6.1 SUMMARY

Over the past several decades, L2 researchers and practitioners alike have paid a great amount of attention to task repetition together with some forms of fluency enhancement (i.e., increasing time pressure) (e.g., De Jong & Perfetti, 2011); however, its positive influence on L2 accuracy development remains unclear (e.g., Thai & Boers, 2016). In addition, the role of corrective feedback in L2 grammar learning has been one of the most extensively researched areas in the field of SLA (e.g., Li, 2010). By interfacing the somewhat independently evolving topics (task repetition and corrective feedback) in an interdisciplinary manner, the current study scrutinized the complex relationship between task repetition, fluency enhancement (4/3/2 activity) and accuracy enhancement (delayed CF). Furthermore, I investigated the extent to which such learner gains could be also ascribed to participants' different cognitive profiles, operationalized as four different kinds of language aptitude—i.e., associative memory, phonemic coding, sound sequence recognition and language analytic ability (Meara, 2005).

Overall, based on the five research questions, the study generated five conclusions. The first research question asked about the degree to which task repetition impacts L2 fluency and accuracy development. The results of the present study

revealed that merely engaging in task repetition only marginally promoted one very easy aspect of breakdown fluency (final-clause pause ratio). Besides, task repetition did not result in any significant accuracy changes.

The second research question asked about the degree to which task repetition + FE impacts L2 fluency and accuracy development. According to the results of the statistical analyses, the 4/3/2 activity (repeating a monologue task with increasing time pressure) can lead to robust fluency development. Specifically, articulation rate increased while ratios of mid-clause and final-clause pauses decreased significantly. Increasing time pressure also contributed to the marginal reduction in the ratio of self-repairs and self-repetitions. However, no significant changes in accuracy were found for this condition suggesting that when learners devoted a lot of attention to fluency, they could not attend to accuracy. This was an evidence of trade-off effects.

The third research question pertained to the degree to which task repetition + AE impacts L2 fluency and accuracy development. According to the results of the analysis, providing corrective feedback to students before they repeated the tasks resulted in large gains in accuracy. These students showed significant improvement not only in the use of the English past tense (both regular and irregular forms) but also in general accuracy (morphosyntactic accuracy). However, this task condition did not lead to any significant changes in any aspects of L2 fluency.

The fourth research question concerned the degree to which task repetition + FE+ AE impact L2 fluency and accuracy development. Results revealed that drawing learners' attention to form through integrating delayed CF into the 4/3/2 activity can help improve both fluency and accuracy at a broad level, suggesting that offering students

opportunities to focus on form between the iterations of the talks somewhat facilitated accuracy development. However, I have yet to know the extent to which a combined approach (fluency and accuracy enhancements) is most effective in helping improve all dimensions of language. When AE is introduced to elicit L2 learners' focus on form, certain aspects of their fluency and accuracy development, especially those related to linguistic encoding (reduction in pauses between clauses and regular past tense), did not change significantly. Again, these results evidenced the trade-off relationship between fluency and accuracy.

Finally, the fifth research question was designed to scrutinize the degree to which learners' improvement patterns are related to their cognitive individual differences in associative memory, phonemic coding, language analytic ability and sound sequence recognition. The results of the present study demonstrated that aptitude is predictive of the extent to which learners benefit from various task activities (task repetition, increasing time pressure, delayed CF). When their treatment cancels out the benefit of FE to some degree (i.e., adding AE to FE), learners with greater associative memory tend to demonstrate larger gains in both fluency (i.e., self-repairs) and accuracy (in the use of irregular past tense) aspects of L2 development. Those with greater phonemic coding show gains in mid-clause pausing. Those with higher language analytic ability show gains in articulation rate. Those with sound sequence recognition ability demonstrate development in semantic accuracy. When the treatment incorporates fluency enhancement (FE only), those with greater associative memory appear to better enhance both their accuracy (i.e., semantic accuracy) and fluency (i.e., self-repetitions) while those greater language analytic ability demonstrate improvement in the use of

regular past tense. Language aptitude therefore, boosts the effectiveness of the treatment.

6.2 Limitations and Future Directions

To close, I would like to emphasize here again the exploratory nature of the current study. With a view of future replication studies, I acknowledge a number of crucial limitations in methodology. First of all, this study examined whether CF promoted students' accurate use of the past tense form, a structure that students have already known. In the CF literature, it has been shown that the presence/absence of L2 learners' explicit knowledge for target structures plays a key role in determining the degree of CF effectiveness (e.g., Ammar & Spada, 2006). According to Ellis and Sheen (2006), CF is more effective if it targets a structure that learners have partial knowledge rather than an entirely new structure. It would be interesting if future studies replicate the findings of the current study but focus on the acquisition of new grammatical features of which learners have no prior knowledge.

Second, in the present study, only one type of CF—i.e., metalinguistic explanation—was used as a form of accuracy enhancement treatment. What remains unanswered here concerns whether students benefit from different types of CF when they are incorporated into the 4/3/2 sequences. Future studies could compare the differential effects of various types of CF (i.e., input providing vs. input prompting or implicit vs. explicit) on the acquisition of the target structures. As evidenced in a considerable number of experimental studies, different types of CF affect L2 development differently (e.g., Ammar & Spada, 2006; Lyster, 2004; Sheen, 2007a, 2010).

Another methodological issue relates to the fact that the current study assessed the students' accuracy based on their targetlike/nontargetlike use of English past tense. The question emerging from this study is whether using global accuracy measures would yield different results (for a list of global accuracy measures, see Yuan & Ellis, 2003). Additionally, although the participants in the current study joined three different sessions of task repetition activities, future studies could adopt a longitudinal design to track L2 learners' performance over multiple repetition sessions for an extensive period of time (e.g., Lambert et al., 2016). In this regard, it would be interesting if future studies can highlight both short- and long-term effectiveness of 4/3/2 activities by adopting both immediate and delayed post-tests. Such studies will shed light on whether the combined effects of task repetition, fluency enhancement and accuracy enhancement will be sustained in the long run. With respect to CF effectiveness and duration of exposure, Long (2007b) claimed that short-term studies of feedback tend to show that explicit types of feedback were more beneficial than implicit CF. In contrast, Long predicted that longer-term treatments would be more likely to favour implicit CF. A further study could also be carried out to gather further evidence to support this claim (see Saito & Akiyama, 2017, 2018).

Another limitation of this study which could have affected the measurements of CF effectiveness was that only accuracy scores in pre/posttests were taken into consideration as outcome measures. In previous CF studies, the impact of CF on various dimensions of SLA can be determined by examining learner uptake (Choi & Li, 2012; Panova & Lyster, 2002; Sheen, 2004, 2006), noticing (Goo, 2012; Mackey, 2006; Philp, 2003) or accuracy scores (Tavakoli et al., 2016; Yuan & Ellis, 2003). In some

studies, uptake was found to be proof of learners' noticing of feedback (Egi, 2010; Lyster & Ranta, 1997). Moreover, it has been suggested in the literature that both successful uptake (i.e., Loewen, 2005) and noticing (i.e., Mackey, 2006) predicted accuracy test scores. Therefore, future research could capture learners' uptake or noticing together with pre/posttest scores to get a full-fledged insight into the efficacy of CF.

Furthermore, I need to acknowledge that all the task repetition and 4/3/2 activities in the current study were operationalized under laboratory conditions with the researcher as a conversational partner, instead of their peers as operationalized in other 4/3/2 studies (Boers, 2014; Thai & Boers, 2016). This decision was made because I wanted to control the nature and amount of delayed CF treatment. However, the generalizability of my results should be tested in classroom settings in the future. Such studies may provide more practical implications for language pedagogy in real classrooms. Moreover, in future studies, students could be trained to be feedback providers. Previous peer CF studies (Sato & Lyster, 2007, 2012; Adams, 2007; Adams et al., 2011; Fujii et al., 2016) have found that peer CF has beneficial effects on L2 development.

Also, it should be noted that the current study was conducted with participants whose L2 proficiency levels were considered pre-intermediate and intermediate. The proficiency level of participants may also have affected their development patterns. Therefore, future studies might wish to examine more advanced learners or include different levels of proficiency and treat learner proficiency as an individual learner factor (cf. Ammar & Spada, 2006; Lambert et al., 2016; Lyster, 2004).

In addition to proficiency levels, working memory also emerged as important learner variables affecting language development in previous task repetition (i.e., Ahmadian, 2013) and CF (Mackey et al., 2002; Mackey et al., 2010; Miyake & Friedman, 1998; Sagarra, 2007 Yilmaz, 2013). These studies have also shown that learners with greater working memory capacity demonstrated more improvement in fluency and accuracy. Therefore, future studies should address working memory as an individual difference variable, as such studies can provide further insights into the effectiveness of CF, task repetition, and 4/3/2 activity.

It is also important to acknowledge that in the present study, students repeated the tasks three times. Future research should compare the effects of task repetition under different repetition conditions in which the number of repetition could be manipulated. In fact, Lambert et al. (2016) found that certain dimensions of fluency (i.e. mid-clause pause, self-repairs) can only be enhanced after the third or fourth performance. In addition, in the present study, the effects of repetition have been measured immediately without adopting any delayed posttest measures. According to Thai and Boers (2016), immediate repetition of the task encourages verbatim duplication which in turn affects accuracy negatively as students carried their errors from one delivery to the next. Future research could compare the effects of immediate repetition and repetition after days or weeks.

Another issue that was not addressed in this study was whether familiarity with the topic could be one of the affecting variables especially with regard to the development of fluency. All the three topics used are familiar topics. Previous literature on content familiarity has shown that familiar topics encouraged students to speak more

fluently (Bui, 2014; Bui & Huang, 2016; Qiu & Lo, 2017). Besides, Révész and Han (2016) figured out that students who received recasts while performing the tasks with familiar content exhibited significant improvement on their L2 oral accuracy. Therefore, further research could be conducted to examine the efficacy of task repetition, fluency and accuracy enhancement on L2 development using both familiar and unfamiliar topics to corroborate these findings.

The findings of the present study could be considered limited in scope, as they were exclusively tied to only fluency and accuracy measures but not complexity measures (but see extensive literature on CAF and TBLT, see Skehan, 2002). Besides, the tasks that were used were complex along resource dispersing dimensions. According to Robinson (2001a, 2003, 2015), tasks that are complex along resource directing dimensions result in more complex L2 performance. A further study is suggested to focus more on increasing cognitive demands along both dimensions.

Finally, this study was limited by the absence of qualitative methods. Lambert et al. (2016) used post-performance questionnaire in their study and obtained useful insights into the value of task repetition. In Sato's (2013) study, learners' beliefs about corrective feedback were investigated. Further research could yield fruitful findings if eliciting detailed information about learners' preferences for task conditions (i.e., constant time and increasing time pressured), types of corrective feedback, and the numbers of repetition by using questionnaires or interviews could be used. Moreover, it would be also interesting to find out learners' perceptions of the beneficial effects of task repetition and corrective feedback. In addition, every effort was made to make sure that students in each group used all the time they had. They were explicitly asked to do so

before engaging in the tasks. According to casual observation, it was clear that students in all conditions enjoyed speaking and filled all the available time with speech. However, at the same time, it was impossible to examine what participants were actually doing in their mind. Therefore, it would be very useful to include individual interviews or think-aloud protocols in future studies to figure out whether students experience much time pressure while carrying out the tasks.

6.3 Implications

Based on the findings of the present study, some implications for both teachers and researchers could be given. Firstly, results from previous studies have shown that the second performance of the task assists conceptualization (Lambert et al., 2016). As a result, final clause are reduced. In addition, future repetition resulted in the reduction of mid-clause pauses (Lambert et al., 2016). However, the results of the present study suggested that adding time pressure resulted in the reduction of mid-clause pausing (encoding) as well as final clause pausing (conceptualization). In other words, the absence of online planning promoted fluency so 4/3/2 activity is really recommendable for foster students' fluency. Secondly, the simultaneous use of fluency enhancement and accuracy enhancement led to increase in both fluency and accuracy but in a broad level only. Thirdly, the effectiveness of the treatment depends largely on learners' foreign language aptitude. Therefore, if teachers collect students' aptitude scores before applying any instructional methods, they could decide which methods are more beneficial for which groups of students.

References

- Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near native second language acquisition. *Studies in Second Language Acquisition*, 30(4), 481–509. doi: 10.1017/S027226310808073X
- Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 29–51). Oxford: Oxford University Press
- Adams, R., Nuevo, A. M., & Egi, T. (2011). Explicit and implicit feedback, modified output, and SLA: Does explicit and implicit feedback promote learning and learner–learner interactions? *The Modern Language Journal*, 95(1), 42–63. doi: 10.1111/j.1540-4781.2011.01242.x
- Ahmadian, M. J. (2011). The effect of ‘massed’ task repetitions on complexity, accuracy and fluency: does it transfer to a new task? *The Language Learning Journal*, 39(3), 269–280. doi: 10.1080/09571736.2010.545239
- Ahmadian, M. J. (2012a). Task repetition in ELT. *ELT Journal*, 66(3), 380–382. doi: 10.1093/elt/ccs020
- Ahmadian, M. J. (2012b). The effects of guided careful online planning on complexity, accuracy and fluency in intermediate EFL learners’ oral production: The case of English articles. *Language Teaching Research*, 16(1), 129–149. doi: 10.1177/1362168811425433
- Ahmadian, M. J. (2013). Working memory and task repetition in second language oral production. *Asian Journal of English Language Teaching*, 23(1), 37–55.

- Ahmadian, M. J., & Tavakoli, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, 15(1), 35–59. doi: 10.1177/1362168810383329
- Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research*, 12(2), 183–210. doi: 10.1177/1362168807086287
- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28(4), 543–574. doi: 10.1017/S0272263106060268
- Arevart, S., & Nation, P. (1991). Fluency improvement in a second language. *RELC Journal*, 22(1), 84–94. doi: 10.1177/003368829102200106
- 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(4), 689–725. doi: 10.1017/S0272263113000429
- Bardovi-Harlig, K., & Comajoan, L. (2008). Order of acquisition and developmental readiness. In B. Spolsky, & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 383–397). Malden, MA: Blackwell. doi: 10.1.1.452.5648
- Benson, S., & DeKeyser, R. (2018). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research*. 1–25. doi: 10.1177/1362168818770921
- Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal*, 45(3), 221–235. doi: 10.1177/0033688214546964

- Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & De Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 3(2), 159–175. doi: 10.1177/0265532212455394
- Braidi, S. M. (2002). Reexamining the role of recasts in native-speaker/nonnative-speaker interactions. *Language Learning*, 52(1), 1–42. doi: 10.1111/1467-9922.00176
- Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research*, 20(4), 436–458. doi: 10.1177/1362168814563200
- Bui, G. H. Y. (2014). Task readiness: Theoretical framework and empirical evidence from topic familiarity, strategic planning, and proficiency level. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 63–94). Philadelphia, PA: John Benjamins. doi: 10.1075/tblt.5.03gav
- Bui, G., & Huang, Z. (2016). L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy. *Language Teaching Research*, 20(1), 1–21. doi: 10.1177/1362168816656650
- Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis, & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–46). London, UK: Heinemann.
- Bygate, M. (1999). Task as the context for the framing, re-framing and unframing of language. *System*, 27(1), 33–48. doi: 10.1016/S0346-251X(98)00048-7
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic*

- tasks, second language learning, teaching and testing* (pp. 23–48). Harlow: Longman.
- Bygate M. (2006). Areas of research that influence L2 speaking instruction. In E. Uso-Juan, & A. Martinez-Flor (Eds.), *Current trends in the development and teaching of the four language skills* (pp. 159–186). Berlin: Mouton De Gruyter. doi: 10.1515/9783110197778.3.159
- Bygate, M. (2009). Effects of task repetition on the structure and control of oral language. In K. Van den Branden, M. Bygate, & J. Norris (Eds.), *Task-based language teaching: A reader* (pp. 249–274). Philadelphia, PA: John Benjamins. doi: 10.1075/tblt.1.15eff
- Bygate, M. & Samuda, V. (2005). Integrative planning through the use of task repetition. In R. Ellis (Ed.), *Planning and task performance in second language* (pp. 37–74). Amsterdam: John Benjamins. doi: 10.1075/llt.11.05byg
- Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2012). Does first language maintenance hamper nativelikeness in a second language? *Studies in Second Language Acquisition*, 34(2), 215–241. doi: 10.1017/S0272263112000034
- Carpenter, H., Jeon, K. S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, 28(2), 209–236. doi: 10.1017/S0272263106060104
- Carroll, J. B. (1965). The contributions of psychological theory and educational research to the teaching of foreign languages. *The Modern Language Journal*, 49(5), 273–281.

- Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *International Journal of Psycholinguistics*, 2, 5–14
- Carroll, J. (1981). Twenty-five years of research on foreign language aptitude. In K. Diller (Ed.), *Individual Differences and Universals in Language Learning Aptitude* (pp. 83–118). Newbury House.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Carroll, J. B., & Sapon, S. M. (1959). *The Modern Language Aptitude Test*. Washington, DC: Second Language Testing Incorporated.
- Carroll, J. B., & Sapon, S. (2002). *Manual for the MLAT*. Bethesda, MD: Second Language Testing.
- Carroll, S., Swain, M., & Roberge, Y. (1992). The role of feedback in adult second language acquisition: Error correction and morphological generalizations. *Applied Psycholinguistics*, 13(2), 173–198. doi: 10.1017/S0142716400005555
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15(3), 357–386. doi: 10.1017/S0272263100012158
- Choi, S. Y., & Li, S. (2012). Corrective feedback and learner uptake in a child ESOL classroom. *RELC Journal*, 43(3), 331-351. doi: 10.1177/0033688212463274
- Cochran, J., McCallum, R. & Bell, S. (2010). Three A's: How do attributions, attitudes, and aptitude contribute to foreign language learning? *Foreign Language Annals*, 43(4), 566–582. doi: 10.1111/j.1944-9720.2010.01102.x

De Bot, K. (1992). A Bilingual Production Model Levelt's 'Speaking' Model Adapted.

Applied Linguistics, 13, 1-24. doi: 10.1093/applin/13.1.1

De Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19(2), 249–275. doi: 10.1017/s0272263197002064

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, 61(2), 533–568. doi: 10.1111/j.1467-9922.2010.00620.x

De Jong, N. (2012). Does time pressure help or hinder oral fluency? In N. deJong, K. Juffermans, M. Keijzer, & L. Rasier (Eds.), *Papers of the An_ELA 2012 applied linguistics conference* (pp. 43–52). Delft, the Netherlands: Eburon.

De la Fuente, M. (2002). Negotiation and oral acquisition of L2 vocabulary: The role of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition* 24(1), 81–112.

DeKeyser, R. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *The Modern Language Journal*, 77(4), 501–514. doi: 10.1111/j.1540-4781.1993.tb01999.x

DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17(3), 379–410.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499–533. Doi: 10.13140/RG.2.1.3958.4486

- DeKeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, 62(2), 189–200. doi: 10.1111/j.1467-9922.2012.00712.x
- DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-Linguistic Evidence for the Nature of Age Effects in Second Language Acquisition. *Applied Psycholinguistics*, 31(3), 413–438. doi: 10.1017/S0142716410000056
- DeKeyser, R., & Koeth, J. (2011). Cognitive aptitudes for second language learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 395–406). New York, NY: Routledge.
- Derwing, M. and J. Rossiter. 2003. 'The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech,' *Applied Language Learning* 13, 1–18.
- Dilans, G. (2010). Corrective feedback and L2 vocabulary development: Prompts and recasts in the adult ESL classroom. *Canadian Modern Language Review*, 66(6), 787–816. doi: 10.3138/cmlr.66.6.787
- Dornyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell. doi: 10.1111/b.9781405132817.2005.00001.x
- Doughty, C., & Varela, E. (1998). Communicative focus *on* form. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). Cambridge: Cambridge University Press.

- Egi, T. (2007a). Interpreting recasts as linguistic evidence: The role of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, 29(4), 511–537. doi: 10.1017/S0272263107070416
- Egi, T. (2007b). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 249–267). Oxford: Oxford University Press.
- Egi, T. (2010). Uptake, modified output, and learner perceptions of recasts: Learner responses as language awareness. *Modern Language Journal*, 94(1), 1–21. doi: 10.1111/j.1540-4781.2009.00980.x
- Ehrman, M. E., & Oxford, R. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal*, 79(1), 67–89. 10.1111/j.1540-4781.1995.tb05417.x
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: OUP.
- Ellis, R. (2005). *Planning and task performance in a second language* (Vol. 11). John Benjamins Publishing. doi: 10.1075/llt.11
- Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review*, 19, 18–41. doi: 10.1075/aila.19.04ell
- Ellis R. (2007). The differential effects of corrective feedback on two grammatical structures. In Mackey A. (Ed.), *Conversational interaction in second language acquisition* (pp. 339–360). New York: Oxford University Press.

- Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, 1(1), 3–18.
- Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research*, 20(3), 405–428. doi: 10.1177/1362168816628627
- Ellis, R., & X. He (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition* 21(2), 285–301.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51(2), 281–318. doi: 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(2), 339–368. doi: 10.1017/S0272263106060141
- Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in second language acquisition*, 28(4), 575–600. doi: 10.1017/S027226310606027X
- Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, 9(2), 147–171. doi: 10.1191/1362168805lr161oa
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3), 299–323. doi: 10.1017/S0272263100015047

- Freed, B., N. Segalowitz, and D. Dewey. 2004. Context of learning and second language fluency in French. *Studies in Second Language Acquisition*, 26(2), 275–301.
- Fu, T., & Nassaji, H. (2016). Corrective feedback, learner uptake, and feedback perception in a Chinese as a foreign language classroom. *Studies in Second Language Learning and Teaching*, 6(1), 161–183. doi: 10.14746/sslit.2016.6.1.8
- Fujii, A., Ziegler, N., & Mackey, A. J. (2016). Peer interaction and metacognitive instruction in the EFL classroom. In M. Sato, & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 63-89). (Language Learning and Language Teaching; Vol. 45). John Benjamins. doi: 10.1075/llt.45.03fuj
- Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, 20(3), 321–340. doi: 10.1177/1362168815570142
- Gardner, R., & MacIntyre, P. (1992). A student's contributions to second language learning. Part 1: Cognitive variables. *Language Teaching*, 25(4), 211–220. doi: 10.1017/S026144480000700X
- Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49(4), 549–581. doi: 10.1111/0023-8333.00102
- Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, 34(3), 445–474. doi: 10.1017/S0272263112000149

- Goo, J. (2016). Corrective feedback and working memory capacity: A replication. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition*. (pp. 279–302). Amsterdam: John Benjamins.
- Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, 35(1), 127–165. 10.1017/S0272263112000708
- Gooch, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English /r/ to Korean adult EFL learners. *System*, 60, 117–127. doi: 10.1016/j.system.2016.06.007
- Granena, G. (2013a). Cognitive attitudes for second language learning and the LLAMA Language Aptitude Test. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude and ultimate attainment* (pp. 3–41). Amsterdam: John Benjamins. doi: 10.1075/llt.35.04gra
- Granena, G. (2013b). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63(4), 1–39. doi: 10.1111/lang.12018
- Granena, G. (2014). Language aptitude and long-term achievement in early childhood L2 learners. *Applied Linguistics*, 35(4), 483–503. doi: 10.1093/applin/amu013
- Granena, G., & Long, M. H. (2013a). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311–343. doi: 10.1177/0267658312461497
- Granena, G., & Long, M. (2013b). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: John Benjamins Publishing Company.

- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36(4), 543–572. doi: 10.2307/3588240
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19(3), 379–400. doi: 10.1017/S0272263197003045
- Harmer, J. (2007). *The practice of English language teaching* (4th Ed.). London: Longman.
- Harrington, M., & Sawyer, M. (1992). L2 Working Memory Capacity and L2 Reading Skill. *Studies in Second Language Acquisition*, 14(1), 25-38.
doi:10.1017/S0272263100010457
- Hawkes, M. L. (2011). Using task repetition to direct learner attention and focus on form. *ELT Journal*, 66(3), 327–336. doi: 10.1093/elt/ccr059
- Hawkes, L., & Nassaji, H. (2016). The role of extensive recasts in error detection and correction by adult ESL students. *Studies in Second Language Learning and Teaching*, 6(1), 19–41. doi: 10.14746/ssllt.2016.6.1.2
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473. doi: 10.1093/applin/amp048
- Housen, A., Kuiken, F. & Vedder, I. 2012, Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken & I. Vedder (eds), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and*

- Fluency in SLA*. Language Learning & Language Teaching, vol. 32, John Benjamins Publishing Company, pp. 1-20.
- Hummel, K. (2009). Aptitude, phonological memory, and second language proficiency in novice adult learners. *Applied Psycholinguistics*, 30(2), 225-249. doi: 10.1017/S0142716409090109
- Hwu, F., Pan, W., & Sun, S. (2014). Aptitude-treatment interaction effects on explicit rule learning: A latent growth curve analysis. *Language Teaching Research*, 18(3), 294–319. doi: 10.1177/1362168813510381
- Hwu, F., & Sun, S. (2012). The aptitude-treatment interaction effects on the learning of grammar rules. *System*, 40(4), 505–521. doi: 10.1016/j.system.2012.10.009
- Ishida, M. (2004). Effects of Recasts on the Acquisition of the Aspectual Form-te i-(ru) by Learners of Japanese as a Foreign Language. *Language Learning*, 54(2), 311–394. doi: 10.1111/j.1467-9922.2004.00257.x
- Iwashita, N. (2001). The effect of learner proficiency on interactional moves and modified output in nonnative–nonnative interaction in Japanese as a foreign language. *System*, 29(2), 267–287. Doi: 10.1016/S0346-251X(01)00015-X
- Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, 25(1), 1–36. doi: 10.1017/S0272263103000019
- Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44(2), 137-166. doi:10.1017/S0261444810000509

- Kartchava, E., & Ammar, A. (2014). The noticeability and effectiveness of corrective feedback in relation to target type. *Language Teaching Research*, 18(4), 428–452. doi: 10.1177/1362168813519373
- Kim, Y. (2013). Promoting attention to form through task repetition in a Korean EFL context. In K. McDonough, & A. Mackey (Eds.), *Second language interaction in diverse educational settings* (pp. 3–24). Philadelphia, PA: John Benjamins. doi: 10.1075/llt.34.04ch1
- Kim, J. H., & Han, Z. H. (2007). Recasts in communicative EFL classes: Do teacher intent and learner interpretation overlap? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 269–297). Oxford: Oxford University Press.
- Kim, Y., Payant, C. (2014). A pedagogical proposal for task sequencing: An exploration of task repetition and task complexity on learning opportunities. In: Baralt, M., Gilabert, M.R., & Robinson, P. (Eds.), *Task sequencing and instructed second language learning* (pp. 151–177). London: Bloomsbury Academic.
- Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, 41(3), 829–840. doi: 10.1016/j.system.2013.08.005
- Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Kormos, J. (2013). New conceptualizations of language aptitude in second language attainment. In G. Granena, & M. H. Long (Eds.), *Sensitive periods, language*

- aptitude, and ultimate L2 attainment. Language Learning and Language Teaching* (pp. 131–152). Amsterdam: John Benjamins.
- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and cognition*, 11(2), 261-271. doi: 10.1017/S1366728908003416
- Lambert, C., Kormos, J., & Minn, D. (2016). Task repetition and second language speech processing. *Studies in Second Language Acquisition*, 39(1), 167–196. doi: 10.1017/S0272263116000085
- Lee, A. H., & Lyster, R. (2016). The effects of corrective feedback on instructed L2 speech perception. *Studies in Second Language Acquisition*, 38(1), 35-64. doi: 10.1017/S0272263115000194
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25(1), 37–63. Doi: 10.1017/S0272263103000020
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1993). Language use in normal speakers and its disorders. In G. Blanken, J. Dittmann, H. Grimm, J. C. Marshall, & C. W. Wallesch, (Eds), *Linguistic disorders and pathologies* (pp. 1–15). Berlins: Walter de Gruyter.
- Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 83–122). Oxford: Oxford University Press.

- Levkina, M., & Gilabert, R. (2012). The effects of cognitive task complexity on L2 oral production. In A. Housen, I. Vedder & F. Kiuken (Eds.), *Dimensions of L2 performance and proficiency investigating complexity, accuracy, and fluency in SLA* (pp. 171–198). Amsterdam: John Benjamins.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–365. doi: 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. *The Modern Language Journal*, 97(3), 634–654. doi: 10.1111/j.1540-4781.2013.12030.x
- Li, S. (2015a). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, 36(3), 385–408. doi: 10.1093/applin/amu054
- Li, S. (2015b). Working memory, language analytical ability and L2 recasts. In Z. Wen, M. Borges, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 139–159). Bristol: Multilingual Matters.
- Li, S. (2016). The Construct Validity of Language Aptitude: A Meta-Analysis. *Studies in Second Language Acquisition*, 38(4), 801–842. doi: 10.1017/S027226311500042X
- Li, S., Zhu, Y., & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *The Modern Language Journal*, 100(1), 276–295. doi: 10.1111/modl.12315

- Linck, J., Hughes, M., Campbell, S., Silbert, N., Tare, M., Jackson, S., Smith, B., Bunting, M., & Doughty, C. (2013). Hi-LAB: A new measure of aptitude for highlevel language proficiency. *Language Learning*, 63(3), 530–566. doi: 10.1111/lang.12011
- Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27(3), 361–386. doi: 10.1017/S0272263105050163
- Loewen, S. (2012). The role of feedback. In S. Gass, & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 42–58). Routledge.
- Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, 19(1), 1–14. doi: 10.1080/09588220600803311
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 361–377). New York: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in the adult L2 classroom: characteristics, explicitness and effectiveness. *Modern Language Journal*, 90(4), 536–556. doi: 10.1111/j.1540-4781.2006.00465.x
- Long, M. H. (2007a). Recasts in SLA: The story so far. In M. H. Long (Ed.), *Problems in SLA* (pp. 75–118). Mahwah, NJ: Lawrence Erlbaum Associates.
- Long, M. H. (2007b). *Problems in SLA*. Lawrence Erlbaum.
- Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*,

- 82(3), 357–371. doi /10.1111/j.1540-4781.1998.tb01213.x
- Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4(3), 221–250. doi: 10.1177/136216880000400303
- Lynch, T., & Maclean, J. (2001). A case of exercising: Effects of immediate task repetition on learners' performance. *Researching pedagogic tasks: Second language learning, teaching and testing*, (pp. 141–162). Addison Wesley Longman.
- Lyster, R. (1998a). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20(1), 51–81. doi: 10.1017/S027226319800103X
- Lyster, R. (1998b). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48(2), 183–218. doi: 10.1111/j.1467-1770.2001.tb00019.x
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26(3), 399–432. doi: 10.1017/S0272263104263021
- Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59(2), 453–498. doi: 10.1111/j.1467-9922.2009.00512.x
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28(2), 269–300. doi: 10.1017/S0272263106060128

- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19*(1), 37–66. doi: 10.1017/S0272263197001034
- Lyster, R., & Saito, K. (2010a). Corrective feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition, 32*(2), 265–302. doi: 10.1017/S0272263109990520
- Lyster, R., & Saito, K. (2010b). Interactional feedback as instructional input: A synthesis of classroom SLA research. *Language, Interaction and Acquisition, 1*(2), 276–297. doi: 10.1075/lia.1.2.07lys
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching, 46*(1), 1–40. doi: 10.1017/S0261444812000365
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics, 27*(3), 405–430. doi: 10.1093/applin/ami051
- Mackey, A. (2012). *Input, interaction and corrective feedback in L2 classrooms*. Oxford: Oxford University Press.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies*. Oxford: Oxford University Press.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal, 82*(3), 338–356. doi: 10.1111/j.1540-4781.1998.tb01211.x

- Mackey, A., Gass, S. M., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22(4), 471–497. doi: 10.1017/S0272263100004010
- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In R. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209), Philadelphia, PA: John Benjamins. doi: 10.1075/llt.2.12mac
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60(3), 501–533. Doi: 10.1111/j.1467-9922.2010.00565.x
- Maurice, K. (1983). The fluency workshop. *TESOL Newsletter*, 17(4), 29.
- McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, 21 (3), 395–423. doi: 10.1017/S0142716400003064
- McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, 56(4), 693–720. doi: 10.1111/j.1467-9922.2006.00393.x
- Meara, P. (2005). *LLAMA language aptitude tests: The manual*. Lognostics: University of Swansea.
- Michel, M. C., F. Kuiken, and I. Vedder. 2007. The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics* 45(13), 241–59.

- Miyake, A., & Friedman, N., (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Erlbaum.
- Muñoz, C. (ed.). 2006. Age and the Rate of Foreign Language Learning. *Multilingual Matters*.
- Nassaji, H. (2007). Reactive focus on form through negotiation on learners' written errors. In S. Fotos, & H. Nassaji (Eds.), *Form focused instruction and teacher education: Studies in honour of Rod Ellis* (pp. 117–129). Oxford: Oxford University Press.
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59(2), 411–452. doi: 10.1111/j.1467-9922.2009.00511.x
- Nassaji, H. (2011). Correcting students' written grammatical errors: The effects of negotiated versus non-negotiated feedback. *Studies in Second Language Learning and Teaching*, 1(3), 315–334. doi: 10.14746/sslit.2011.1.3.2
- Nassaji, H. (2016). Anniversary article interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, 20(4), 535–562. doi: 10.1177/1362168816644940
- Nakata, T. (2014). Effects of feedback timing on second language vocabulary learning: Does delaying feedback increase learning? *Language Teaching Research*, 19(4), 416–434. doi: 10.1177/1362168814541721
- Nation, P. (1989). Improving speaking fluency. *System*, 17(3), 377–384.

- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51(4), 719–758. doi: 10.1111/0023-8333.00172
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47 (3), 203–210. doi: 10.1093/elt/47.3.203
- O'Brien, M. G. (2014). L2 learners' assessments of accentedness, fluency, and comprehensibility of native and nonnative German speech. *Language Learning*, 64(4), 715–748. doi: 10.1111/lang.12082
- Otto, L. (2002). Magyar Egysesges Nyelverzekmero Teszt. Unpublished material.
- Panova, I., & Lyster, R. (2002). Patterns of feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36(4), 573–595. 10.2307/3588241
- Philp, J. (2003). Constraints on “noticing the gap”: non-native speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 25(1), 99–126. doi: 10.1017/S0272263103000044
- Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery (form S)*. New York: Harcourt, Brace & World.
- Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6(11), 456–463. doi: 10.1016/S1364-6613(02)01990-3
- Pinter, A. (2005). Task repetition with 10-year old children. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 113–126). New York, NY: Palgrave Macmillan.
- Pinter, A. (2007a). Some benefits of peer-peer interaction: 10-year-old children practising with a communication task. *Language Teaching Research*, 11(2), 189–207. doi: 10.1177/1362168807074604

- Pinter, A. (2007b). What children say: Benefits of task repetition. In K. Van den Branden, K. Van Gorp, & M. Verhelst (Eds.), *Tasks in action: Task-based language education from a classroom-based perspective* (pp. 131–158). Newcastle, UK: Cambridge Scholars Publishing.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. doi: 10.1111/lang.12079
- Qiu, X., & Lo, Y. Y. (2017). Content familiarity, task repetition and Chinese EFL learners’ engagement in second language use. *Language Teaching Research*, 21(6), 681–698. doi: 10.1177/1362168816684368
- Ranta, L. & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The awareness–practice–feedback sequence, in R. DeKeyser (Ed.) *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511667275.009
- Révész, A. (2002). How to avoid the comparative fallacy during data analysis: A Review of Doughty and Varela (1998) and Mackey (1999). *TESOL & Applied Linguistics*, 4(1), 1–4.
- Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, 31(3), 437–470. doi: 10.1017/S0272263109090366
- Révész, A. (2012). Working Memory and the Observed Effectiveness of Recasts on Different L2 Outcome Measures. *Language Learning*, 62(1), 93–132. doi: 10.1111/j.1467-9922.2011.00690.x

- Révész, A., & Han, Z. (2006). Task content familiarity, task type and efficacy of recasts. *Language Awareness, 15*(3), 160–179. doi: 10.2167/la401.0
- Révész, A., Sachs, R., & Mackey, A. (2011). Task complexity, uptake of recasts, and L2 development. In Second language task complexity. In P. Robinson (Ed.), *Researching the cognition hypothesis of language learning and performance*. (pp. 203–235). Amsterdam: John Benjamins.
- Robinson, P. (1995). Aptitude, awareness and the fundamental similarity of implicit and explicit foreign language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 303–357). Honolulu, HI: University of Hawai'i Press,
- Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult SLA. *Language Learning, 47*(1), 45–99. doi: 10.1111/0023-8333.21997002
- Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.): *Cognition and second language instruction* (pp. 287–318). Cambridge University Press. doi: 10.1017/CBO9781139524780.012
- Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied linguistics, 22*(1), 27-57. doi: 10.1093/applin/22.1.27
- Robinson, P. (2003). The Cognition Hypothesis, task design and adult task-based language learning. *Second Language Studies, 21*(2), 45–107.

- Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 46–73. doi: 10.1017/S0267190505000036
- Robinson, P. (2011). *Second language task complexity: researching the cognition hypothesis of language learning and performance* (Vol. 2). Amsterdam: John Benjamins
- Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In *Domains and directions in the development of TBLT* (pp. 87–122). John Benjamins Amsterdam. doi: 10.1075/tblt.8.04rob
- Rolin-lanziti, J. (2010). The organization of delayed second language correction. *Language Teaching Research*, 14(2), 183–206. doi: 10.1177/1362168809353874
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 Grammar: A meta-analysis of the research. In: J. Norris, & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Amsterdam: John Benjamins. doi: 10.1075/llt.13.09val
- Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *IRAL-International Review of Applied Linguistics in Language Teaching*, 46(2), 113–136. doi: 10.1515/IRAL.2008.005
- Sample, E., & M. Michel. (2014). “An Exploratory Study into Trade-Off Effects of Complexity, Accuracy, and Fluency on Young Learners' Oral Task Repetition.” *TESL Canada Journal* 31(8), 23–46. doi: 10.18806/tesl.v31i0.1185

- Sagarra, N. (2007). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 229–248). New York: Oxford University Press.
- Saito, K. (2013). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English /r/ to adult Japanese learners. *Language Learning*, 63(3), 499–529. doi: 10.1111/lang.12015
- Saito, K. (2015). Variables affecting the effects of recasts on L2 pronunciation development. *Language Teaching Research*, 19, 276–300.
- Saito, K. (2017). Effects of sound, vocabulary and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning*, 67(3), 665–693. doi: 10.1111/lang.12244
- Saito, K. (2018). To What Extent Does Long-Term Foreign Language Education Help Improve Spoken Second Language Lexical Proficiency?. *TESOL Quarterly*, 53(1), 82-107. doi: 10.1002/tesq.468
- Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language Learning*, 67, 43-74.
- Saito, K., & Akiyama, Y. (2018). Effects of video-based interaction on the development of second language listening comprehension ability: A longitudinal study. *TESOL Quarterly*, 52, 163-176.
- Saito, K., & Lyster, R. (2012a). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of

- English. *Language Learning*, 62(2), 595–633. doi: 10.1111/j.1467-9922.2011.00639.x
- Saito, K. & Lyster, R. (2012b). Investigating the pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly*, 46(2), 385–396. doi: 10.1002/tesq.25
- Saito, K., Ilkan, M., Magne, V., Tran, M., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low, mid and high-level second language fluency. *Applied Psycholinguistics*. 39(3), 593–617. doi:10.1017/S0142716417000571
- Saito, K., Suzukida, Y., & Sun, H. (2018). Aptitude, experience and second language pronunciation proficiency development in classroom settings: a longitudinal study. *Studies in Second Language Acquisition*. doi: 10.1017/S0272263117000432
- Sato, M. (2013), Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *Modern Language Journal*, 97(3), 611-633. doi:10.1111/j.1540-4781.2013.12035.x
- Sato, M. & Lyster, R. (2007). Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 123–142). Oxford: Oxford University Press.
- Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, 34(4), 591–626. doi: 10.1017/S0272263112000356

- Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning and Technology* 13(1), 96–120. doi: 10.125/44170
- Sawyer, M. & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 319–354.
- Scrivener, J. (2005). *Learning teaching: A guidebook for English language teachers* (2nd Ed.). Oxford: Macmillan.
- Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. Malden, MA: Blackwell.
- Segalowitz, N. (2012). Second language fluency. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 240–244). New York: Routledge.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research* 8(3), 263–300. doi: 10.1191/1362168804lr146oa
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10(4), 361–392. doi: 10.1191/1362168806lr203oa
- Sheen, Y. (2007a). 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 series of empirical studies* (pp. 301–322). Oxford: Oxford University Press.

- Sheen, Y. (2007b). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283. doi: 10.1002/j.1545-7249.2007.tb00059.x
- Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32(2), 203–234. doi: 10.1017/S0272263109990507
- Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 593–607). New York: Routledge.
- Shintani, N., & Ellis, R. (2015). Does language analytical ability mediate the effect of written feedback on grammatical accuracy in second language writing? *System*, 49, 110–119. doi: 10.1016/j.system.2015.01.006
- Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing*, 3(2), 188–221.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2001). Tasks and language performance. In Bygate M., P. Skehan and M. Swain. (eds.): *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 167–185). London: Longman.
- Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–95). Philadelphia, PA: John Benjamins.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics*, 30(4), 510-532.

Skehan, P. (2012). Language aptitude. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp. 381–395). New York, NY: Routledge.

Skehan, P. (2014). Limited attentional capacity, second language performance and task-based pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–262). Amsterdam: Benjamins.

Skehan, P. (2015). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics*, 36(3), 367–384. doi: 10.1093/applin/amu072

Skehan, P. (2015). Working memory and second language performance: A commentary. In Z. E. Wen, M. B. Mota, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 189–204). Bristol: Multilingual Matters.

Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in L2 processing and acquisition* (pp. 15–38). Amsterdam, the Netherlands: John Benjamins. doi: 10.1075/bpa.3.02ske

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1(3), 185–211. doi: 10.1177/136216889700100302

- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49(1), 93–120. doi: 10.1111/1467-9922.00071
- Skehan, P. & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, & I. Vedder (Eds.). *Complexity, Accuracy, and Fluency in Second Language Use, Learning, & Teaching* (pp. 207–226). Brussels: Contactforum.
- Sparks, R., & Ganschow, L. (1991). Foreign language learning difficulties: Affective or native language aptitude differences? *Modern Language Journal*, 75(1), 3–16. doi: 10.1111/j.1540-4781.1991.tb01076.x
- Sparks, R., & Ganschow, L. (1995). A strong inference approach to causal factors in foreign language learning: A response to MacIntyre. *Modern Language Journal*, 79(2), 235–244. doi: 10.1111/j.1540-4781.1995.tb05436.x
- Sparks, R., Patton, J., Ganschow, L. & Humbach, N. (2011). Subcomponents of L2 aptitude and L2 proficiency. *The Modern Language Journal*, 95(2), 253–273. doi: 10.1111/j.1540-4781.2011.01176.x
- Suzuki, Y., & Dekeyser, R. (2017). Exploratory research on second language practice distribution: An Aptitudex Treatment interaction. *Applied Psycholinguistics*, 38(1), 27–56. doi: 10.1017/S0142716416000084
- Takashima, H., & Ellis, R. (1999). Output enhancement and the acquisition of the past tense. *Learning a second language through interaction* (pp. 173–188). Amsterdam, the Netherlands: John Benjamins. doi: 10.1075/sibil.17.12tak

- Tavakoli, P. & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.): *Planning and task performance in a second language* (pp. 239–277). Amsterdam: Benjamins. doi: 10.1075/llt.11.15tav
- Tavakoli, P., Campbell, C., & McCormack, J. (2016). Development of speech fluency over a short period of time: Effects of pedagogic intervention. *TESOL Quarterly*, 50 (2), 447–471. doi: 10.1002/tesq.244
- Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly*, 50(2), 369–393. doi: 10.1002/tesq.232
- Thomas, K. E. (2018). Comparing Explicit Exemplar-Based and Rule-Based Corrective Feedback: Introducing Analogy-Based Corrective Feedback. *The Modern Language Journal*. 102(2), 371–391. doi: 10.1111/modl.12470
- 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 series of empirical studies* (pp. 171–195). Oxford: Oxford University Press.
- Van De Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2015). Learning New Grammatical Structures in Task-Based Language Learning: The Effects of Recasts and Prompts. *Modern Language Journal*, 99(2), 246–262. doi: 10.1111/modl.12211
- Van de Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2016). Focus on form through task repetition in TBLT. *Language Teaching Research*, 20(3), 300–320. doi: 10.1177/1362168815609616

- Van Patten, B. & Borst, S. (2012). The roles of explicit information and grammatical sensitivity in processing instruction: Nominativeaccusative case marking and word order in German L2. *Foreign Language Annals*, 45(1), 92–109. doi: 10.1111/j.1944-9720.2012.01169.x
- Van Patten, B., Collopy, E., Price, J. E., Borst, S., & Qualin, A. (2013). Explicit Information, Grammatical Sensitivity, and the First-Noun Principle: A Cross-Linguistic Study in Processing Instruction. *Modern Language Journal*, 97(2), 506–527. doi: 10.1111/j.1540-4781.2013.12007.x
- Vercellotti, M. L. (2017), The Development of Complexity, Accuracy, and Fluency in Second Language Performance: A Longitudinal Study. *Applied Linguistics*. 38(1), 90–111, doi: 10.1093/applin/amv002
- Wang, Z. (2014). L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). Amsterdam: John Benjamins.
- Wen, Z. E., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, 50(1), 1–31. doi: 10.1017/S0261444816000276
- Wood, D. 2010. *Formulaic language and second language speech fluency: Background, evidence, and classroom applications*. London/New York: Continuum.
- Yalçın, Ş., & Spada, N. (2016). Language aptitude and grammatical difficulty. *Studies in Second Language Acquisition*, 38(2), 239–263. doi: 10.1017/S0272263115000509

- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32(2), 235–263. doi: 10.1017/S0272263109990519
- Yilmaz, Y. (2012). The relative effects of explicit correction and recasts on two target structures via two communication modes. *Language Learning*, 62(4), 1134–1169. doi: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(3), 344–368. doi: 10.1093/applin/ams044
- Yilmaz, Y., & Granena, 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(1), 147–161. Doi: 10.1017/S136672891400090X
- Yilmaz, Y., Granena, G., & Meyer, Z. S. (2016). The role of explicit language aptitude in implicit, explicit, and mixed feedback conditions. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition*. (pp. 327–349). Amsterdam: John Benjamins.
- Yilmaz, Y., & Koylu, Y. (2016). The interaction between feedback exposure condition and phonetic coding ability. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition*. (pp. 303–326). Amsterdam: John Benjamins.

- Yilmaz, Y., & Yuksel, D. (2011). Effects of communication mode and salience on recasts: A first exposure study. *Language Teaching Research, 15*(4), 457–477. doi: 10.1177/1362168811412873
- Yoshida, R. (2008). Teachers' choice and learners' preference of corrective feedback types. *Language Awareness, 17*(1), 78–93. doi: 10.2167/la429.0
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*(1), 1–27. doi: 10.1093/applin/24.1.1
- Zyzik, E., & Polio, C. (2008). Incidental focus on form in university Spanish literature courses. *Modern Language Journal, 92*(1), 53–70. 10.1111/j.1540-4781.2008.00686.x

Appendices

Appendix A: Information Sheet and Consent Form for EFL Students

Department of Applied Linguistics and Communication

BIRKBECK

University of London

Malet Street,

London WC1E 7HX

020 7631 6000

Title of your study: Effects of task repetition and corrective feedback on fluency and accuracy in EFL learners' oral production

Name of researcher: Mai Ngoc Tran

The study is being done as part of my PhD Applied Linguistics degree in the Department of Applied Linguistics and Communication, Birkbeck, University of London. The study has received ethical approval.

This study wants to investigate how conversation activities (e.g., task repetition and corrective feedback) help students improve their L2 fluency and accuracy.

If you agree to participate you will agree a convenient time and place to respond to a questionnaire, take part in three training sessions which will take place on three different days (about 20 minutes per day per student). In every session, you will talk about one easy topic three times. Your speeches will be audio – recorded during the training sessions. You are free to withdraw at any time.

Your data will be kept anonymous by myself and will be stored privately.

The analysis of your participation in this study will be written up in a report for my degree. You will not be identifiable in the write up or any publication which might ensue.

The study is supervised by Dr Saito who may be contacted at the above address and telephone number.

1. Who is conducting the study?

I will be conducting the study as your teacher.

2. Will this affect my learning and how will the study be done?

You will participate in the study in your free time which will not affect your learning. The obtained results will have some implications which allow your teachers to select more effective activities to help you improve your speech performance. The study also provides you with guidance about how you should practice to be more competent speakers of English.

3. Where will the information be stored and is it confidential?

The information will be stored in my own laptop and one USB. All information regarding who you are will be removed and kept anonymous.

4. Can I see the final study?

The final study will be made available to the school, but ultimately destroyed after a set period of time deemed suitable by the university.

5. How will this study benefit me?

Task repetition and corrective feedback technique are being used around the world to help students improve fluency and accuracy in their oral production. Hopefully, you can also benefit when they are applied in your language classroom.

6. Is participation compulsory?

Participants are asked to volunteer in the research and can withdraw at any time.

7. If I have questions?

If you have any questions regarding this study please contact me either at work, face to face or via email: tranngocmai100@gmail.com

Name of researcher: Mai Ngoc Tran

I have been informed about the nature of this study and willingly consent to take part in it.

I understand that the content of the data collected will be kept confidential.

I understand that I may withdraw from the study at any time.

I am over 16 years of age.

Name _____

Signed _____

Date _____

There should be two signed copies, one for participant, one for researcher.

Appendix B: Background Questionnaire for Students

This survey questionnaire is used to elicit background information of the participants for the study “Effects of task repetition and corrective feedback on fluency and accuracy in EFL learners’ oral production”. Your assistance in completing the following questions is greatly appreciated.

Part 1: Background Information

1. Gender: Male/ Female

2. Age: ____

3. Where are you from? _____

4. What is your native language? 1) Vietnamese 2) English 3) Other: _____

5. What language(s) do you speak at home?

1) Vietnamese 2) Other _____

5a. If more than one, with whom do you speak each of these languages?

6. In what language(s) did you receive the majority of your *precollege* education?

1) Vietnamese 2) Other _____

6a. If more than one, please give the approximate number of years for each language. _____

7. Have you ever been to an English-speaking region *for the purpose of studying English?*

Circle one: Yes/ No

7a. If yes, when? _____ 7b. Where? _____

7c. For how long?

(1) 1 semester or less (2) 2 semesters (3) more than 2 semesters

8. In the boxes below, rate your English ability. Use the following ratings:

0) Poor 1) Good 2) Very good 3) Native/nativelike

Listening

Speaking

Reading

Writing

9. How old did you first start learning English? _____

How many hours do you take English classes in the following schools per week?

a) Elementary school

b) Secondary school

c) High school

d) University/college

10. What year are you in? (Circle one):

Freshman

Sophomore

Junior

Senior

11. What is your major? _____

Part 2: Your Use of English

12. Do you *communicate* with native or fluent speakers of English *in English* currently? _____

12a. If yes, with whom do you talk to? _____

12b. How many hours do you communicate with native or fluent speakers of English *in English* per week? _____

13. Are you taking any English courses this semester? _____

13a. If yes, what are they? _____

13b. How many hours do you take these courses per week?

14. How often do you practice English outside the classroom?

14a. How do you practice English?

14b. How many hours do you practice English per week?

15. Please provide the scores of the below tests:

- IELTS: _____

- TOEIC: _____

- TOEFL: _____

*Appendix C: Speaking Tasks***Shrinking time condition****Topic 1: The last favourite movie you watched**

You will have to talk about this topic THREE times. First, you will talk about it for 4 minutes. Then repeat the talk in 3 minutes, and finally deliver the talk again in just 2 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. *What was it called?*
2. *What kind of movie was it?*
3. *When and where did you watch it?*
4. *Who were the main characters?*
5. *What happened in the movie?*
6. *Why did you like it?*

Begin your talk with the first sentence given

The favourite movie I recently watched was...

Topic 2: Your last summer vacation

You will have to talk about this topic THREE times. First, you will talk about it for 4 minutes. Then repeat the talk in 3 minutes, and finally deliver the talk again in just 2 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. *Did you enjoy your last summer vacation? If so, what did you enjoy about it? If not, why did you encounter any problems?*
2. *Where did you go?*
3. *Who did you go with?*
4. *How did you get there?*
5. *How long did you stay there?*
6. *What did you do?*

Begin your talk with the first sentence given

For my last summer vacation, I went to....

Topic 3: Your 16th birthday celebration

You will have to talk about this topic THREE times. First, you will talk about it for 4 minutes. Then repeat the talk in 3 minutes, and finally deliver the talk again in just 2 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. *How did you celebrate your 16th birthday?*
2. *Where did it take place?*
3. *Who did you invite?*

4. *What kinds of presents did you receive?*
5. *What did you eat?*
6. *Did you enjoy your party? Why/ Why not?*

Begin your talk with the first sentence given

I celebrated my 16th birthday...

Constant- time condition

Topic 1: The last favourite movie you watched

You will have to talk about this topic THREE times, each in 3 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. *What was it called?*
2. *What kind of movie was it?*
3. *When and where did you watch it?*
4. *Who were the main characters?*
5. *What happened in the movie?*
6. *Why did you like it?*

Begin your talk with the first sentence given

The favourite movie I recently watched was...

Topic 2: Your last summer vacation

You will have to talk about the topic THREE times, each in 3 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. *Did you enjoy your last summer vacation? If so, what did you enjoy about it? If not, why did you encounter any problems?*
2. *Where did you go?*
3. *Who did you go with?*
4. *How did you get there?*
5. *How long did you stay there?*
6. *What did you do?*

Begin your talk with the first sentence given

For my last summer vacation, I went to...

Topic 3: Your 16th birthday celebration

You will have to talk about the topic THREE times, each in 3 minutes.

You have 3 minutes to think about what you are going to say.

You can make notes if you wish.

You should talk about:

1. *How did you celebrate your 16th birthday?*
2. *Where did it take place?*
3. *Who did you invite?*
4. *What kinds of presents did you receive?*

5. *What did you eat?*

6. *Did you enjoy your party? Why/ Why not?*

Begin your talk with the first sentence given

I celebrated my 16th birthday...

*Appendix D: Examples of Repeated Monologue in the Shrinking Time Condition (4/3/2)***First delivery**

For my last summer vacation, I went to Dalat with my parents. I enjoy it very much. And I went went to, when I went to Dalat, it had flower festival. And it so the weather is cold and I can away from the sum summer heat. But unfortunately, the weather was bad. It is raining all day and the market is crowded and noisy. I go to Dalat by car and it took us four hours to get there. And I stay it in four days. In the first day, I went to the flower festival festival and and we we took a picture together in there. And then I we went to Truc Lam pagoda and the air is fresh and the view beautiful, in the second day, we take take a double-bike and ride around the Xuan Huong lake. And drink some hot coffee in the morning. In the third day, we visit the mansion of Bao Dai King. Then we get get to the museum, with a lot of the animal in there. In the the last day, we went to the Langbiang mountain with the jerk car and the, you know the view is very amazing with a lot of trees, fresh air and good food. Although I have travelled to a lot of place like Vung Tau, Nha Trang, Mui Ne but it is the trip that I am most impressed. After this trip, we feel comfortable and relax. We can see new and and famous place. We could learn about the history of these place. We also know about the traditional and the customs of the people in Dalat. I really like there is a special feature about the students in Dalat. They usually wear a blue sweater when they go to school. And they usually use a bike to go to school. If they have time. If I have time I will come back to Dalat in three years. I like travelling to different place. Travel make me feel happy, relaxed and enjoyable. I go there by car and it took us four hours to get there. Uh in the way we go, I had some troubles and with the traffic jam. And the road is dangerous. We had to go up and down

to the hills and I was so scared. And on the way we go, we see some restaurant and stop station. And we get so tired so we come to a stop station and take a rest. After about ten minutes, we keep going and in the way we go we see some bus, travelling bus and lot of people. During the trip I can't sleep because I was very excited. My seat is next to the car window so I can enjoy beautiful scene, such as pine forests, waterfalls, many kinds of flowers, hills and valleys. They are so amazing. I take some wonderful photos and I keep them in my album.

Second delivery

For my last vacation, I went to Dalat. I enjoy it very much. When I went to Dalat, it's got it has flower vacation. And it's such cool weather so I can avoid avoid summer heat. It's got a lot of flowers there. We take a photo together but unfortunately, the weather was bad. It is raining all day and the market is crowded and noisy. I go with my parents. I go by car and it took four hours to get there. But in the way we go, we got some troubles with the traffic jam. The road is dangerous. We go up and down the hills. On the way we go we see some restaurant and stop station. We get tired so we came to a stop station and take a rest. After about ten minutes we keep going. During the trip I can't sleep because I was very excited. My seat is next to the car window so I can enjoy beautiful scene, such as pine forests, waterfalls, many kinds of flowers, hills and valleys. They are amazing and I take some wonderful photos and I keep them in my album. In the first day, we enjoy the flower festival with lots of flowers in there and then we go to Truc Lam pagoda. The atmosphere in there is fresh and it has beautiful scene. In the second day, we take around the Xuan Huong Lake with double-bike and drink some hot coffee in Trung Nguyen. In the third day, we visit mansion of King Bao Dai and animal

museum. And it has a lot of the models of the animal in there. And in the the last day, we get to the Langbiang mountain by the jerk car and have a good food in there.

Although I have travelled to a lot of place like Vung Tau, Nha Trang, Mui Ne but it is the trip that I am most impressed. After this trip, we feel comfortable and relax. We can see new and and famous place. We could learn about the history of these place. We also know about the traditions and the customs of the people in Dalat. I really like there is a special feature of the students in Dalat. They usually wear sweater when they go to school. If I have time I will come back to Dalat in three years. I like travelling to different place. Travel make me feel happy, relaxed and enjoyable.

Third delivery

For my last summer vacation, I went to Dalat with my parents. I enjoy it very much.

When I went to Dalat, it was having flower festival. A lot and a lot of flowers in there I take a, we take a photo together but unfortunately, the weather was bad. It is raining all day and the market is crowded and noisy. We go it by car. We travel it by car and it took us four hours to get there but in the the way we go, we get trouble with the traffic jam.

The road is dangerous. We go up and down the hills on the way we go we see some restaurants and stop stations. We get tired and we come to the stop station take a rest.

After about ten minutes, we keep going. During the trip, I cannot sleep because I am very excited. My seat is next to the car window so I can enjoy the beautiful scenes such as pine forests, waterfalls, many kinds of flowers hills and valleys. They are amazing. I take some waterfall photos and I keep them in my album. I stay four days in Dalat. In the first day, we enjoy the flower festival and then we go to Truc Lam pagoda. In the second day, we we taked we ride around, we ride a double-bike around the Xuan

Huong lake and drink hot coffee in the morning. And the third day we visit the mansion of King Bao Dai and animal Museum and and the last day, we go to the Langbiang mountain by the jeep car and enjoy the food in there. Although I have travelled to a lot of place like Vung Tau, Nha Trang, Mui Ne but it is the trip that I am most impressed. After this trip, we feel comfortable and relax. We.

*Appendix E: Examples of Repeated Monologue in the Constant Time Condition (3/3/3)***First delivery**

For my last summer vacation, I went to Dalat with my friend and it was a nice trip so I really like the the weather here the weather there. The wea the weather is very cool although it is the summer. So the first day I went to Dalat, I it is a little bit tired so I just stay alone stay at the hotel and sleep, just sleep. And after that I go to Dalat market with my friend to buy some scarf, some sweater or something like that and it is not take a long time for me at the market. But after that it rained a lot so I just I must find the umbrella to buy and come back my hotel and after that I just come back my hotel and sleep, take a bath and sleep. And in the in the evening my friend and I go around Xuan Huong Lake and take some photographs. It is a very beautiful place and and eat some fresh food. The food there is very nice. I like to eat the food there and and the second day I just I just go to flow I just went to flower market. There is a lot of flowers that I like like rose, roses, sunflowers, or something like that or lyly. I like lyly. And or something like that. And we take lots of photographs. We really like to take the photograph and my friends too. And after that we go to elephant waterfall. It takes a long time to go elephant waterfall but it's really really beautiful. And lots of people go to, went to elephant waterfall with us too. And they are very nice. They usually say hi when they see me. And some foreigners they went to elephant fall, waterfall, too. And after that we come back, we come back my hotel and eat some food and sleep. After that the third day, the third day we go to pine forest, it is a, it is very, it is a nice place and the atmosphere is very fresh. And and furthermore, it's very. It's lots of beautiful villas. I just I just look the villas and take some photograph with my friends and after that we we share it on

facebook to tell my friends that they are villas. And my friends say that oh no. And after that my friends and I go to, went to Lam Dong to visit my uncles. They are very nice. They cook for us lots of food. And they took they took us to some pagodas like Duoc Su Pagoda, Linh Phuoc pagoda, Linh Son Pagoda, Truc Lam Pagoda and lots of place that we really like so.

Second delivery

For my last summer vacation, I went to Dalat with my friend. It's a nice trip. The first thing I like the weather here. The weather is very cool and although it is the summer, it was summer. The first day I went to Dalat, I feel a little bit tired so I just stay at the hotel and sleep. After that we go to, we went to market to buy some scarves, some sweater or something like that. And but after that the weather is rained outside so we just find some place to buy the umbrella and come back my hotel and we just sleep. After that in the morning, no in the evening we come to Xuan Huong Lake. We go around Xuan Huong Lake and take some photographs and eat some food, some fresh food there. And the second day, we went to visit flower market. I like lots of flowers there like lyly, roses, sunflowers. And we take lots of, lots of photos there. And my friends also like to take some photos too. And after that we went to elephant waterfall. And there are lots of people went to elephant waterfall too. And there are lots of foreigners too. They they really, and they also to. They also like take some photographs there. And after that we went we come back my hotel and eat some food and sleep. And the third day, we went to the pine forest. The atmosphere there is very fresh. And lots of beautiful villas there. And we we just take some we just take some photos and after that we share it on facebook. And we say that its' our villas. And my friends say oh no. And and after that

we went to Lam Dong to visit my uncles. They are very nice. They cook for us some food, some nice food. And after that they took us to pagodas, to Duoc Su Pagoda, Linh Son Pagoda, or Truc Lam Pagoda. And it's a nice trip. I really like the trip. But I think I think my experience is.

Third delivery

For my last summer vacation, I went to Dalat with my friend. It is a nice trip. The first thing I like of Dalat is the weather. It is very cool although it was summer and the first day I went to Dalat I feel so tired so we just stay at the hotel and sleep. After that, we go to we went to market to buy some scarf some sweater. And after that the rain is it rain outside a lot so we just find some place to buy the umbrella to come back the hotel buy and we just take a bath and sleep. And in the evening we go around Xuan Huong Lake to take some photos and to eat some fresh food there. The food there is very nice. I love the food there and the second day we went to flow flower market. There is a lot of lowers that I like like lyly roses, or sunflowers. And I love lyly. And we take the we take lots of photographs. We took lots of photographs there. And after that we went to elephant waterfall. There is lots of people who visit elephant waterfall and lots of foreigners too and they also took some photographs like like us. And after that we just come back my hotel and eat some food and sleep. After that the third day we went to pine forest and the atmosphere there is very fresh. And there is lots of villas. There's lots of beautiful villas. And we take lots of photographs there. And we share it on facebook. And after that we went to Lam Dong to visit my uncles. They are very nice. And they cook lots of food for us. And after they took us to go to some pagodas like Duoc Su pagodas, Truc Lam Pagodas, Linh Son pagodas. And from the from the trip, I

have some experience like if you if you went to if you went to, go to Dalat or somewhere you should take Phuong Trang coach or Thanh Bui coach but not take outside coach because it take, it's so small. It's not comfortable for you and it have no food for you on the trip.

Appendix F: Obligatory Contexts for Past Tense Use

		Time 1		Time 2	
Groups	ID	Regular	Irregular	Regular	Irregular
Control	101	6	13	6	22
	102	6	14	7	21
	103	5	19	5	14
	104	9	15	8	21
	105	11	9	5	21
	106	11	7	12	13
	107	10	7	12	14
	108	11	20	13	17
	109	16	15	13	11
	110	9	10	11	14
	111	7	16	9	15
	112	7	12	9	6
FE	201	10	14	8	17
	202	8	9	10	19
	203	5	19	8	19
	204	3	20	14	22
	205	3	26	10	26
	206	8	16	5	20
	207	5	20	10	24

	208	4	16	7	21
	209	9	22	14	21
	210	2	9	8	20
	211	7	15	8	15
	212	11	14	6	19
AE	301	8	10	9	12
	302	9	15	7	17
	303	9	11	2	13
	304	3	9	11	21
	305	6	17	10	13
	306	3	25	3	27
	307	6	18	10	16
	308	7	18	4	14
	309	6	14	6	16
	310	3	6	8	6
	311	5	12	5	12
	312	6	12	8	15
FE+AE	401	8	18	8	23
	402	8	26	6	22
	403	8	15	10	25
	404	5	18	7	18
	405	7	12	9	13

	406	11	14	12	11
	407	5	13	9	16
	408	6	15	9	21
	409	9	14	8	17
	410	10	14	7	24
	411	7	12	11	21
	412	8	18	12	20

Appendix G: Intercoder Reliability Analysis

Coders	ID	Fluency					Accuracy				
		Speed	Breakdown		Repair		Regular		Irregular		
		Articulatio	Mid-	Final-	Repair	Repetitio	Regular	Irregular	Global	Semanti	Morphos
		n rate	clause	clause		n				c	yntactic
First	101	2.336	0.1188	0.2574	0	0	16.67	69.23	5.00	8.00	6.00
coder	102	2.655	0.1389	0.125	0	0.0486	33.33	78.57	7.00	6.00	7.00
	103	2.443	0.0833	0.2833	0.0167	0.0167	.00	31.58	7.00	5.00	4.00
	104	2.79	0.1181	0.2677	0.0472	0.1260	.00	86.67	5.00	5.00	5.00
	105	2.086	0.2	0.3	0.01	0.05	18.18	33.33	6.00	6.00	5.00
	106	2.375	0.2046	0.1932	0.0227	0.0568	54.55	100.00	7.00	7.00	7.00
	107	2.414	0.2571	0.2476	0	0.0095	30.00	28.57	4.00	6.00	5.00
	108	3.351	0.0671	0.1275	0	0	18.18	60.00	7.00	7.00	6.00
	109	3.021	0.1452	0.1694	0.0081	0.0323	12.50	53.33	5.00	6.00	5.00
	110	3.016	0.1168	0.1387	0.0146	0.0073	.00	70.00	8.00	8.00	6.00
	111	2.147	0.2	0.2316	0.0316	0.1053	16.67	72.27	6.00	7.00	7.00

4/3/2 ACTIVITY REVISITED

237

	112	3.027	0.1026	0.0961	0.0192	0.0256	14.29	76.19	7.00	7.00	6.00
	113	2.859	0.0916	0.1409	0.0070	0.0070	.00	21.43	6.00	6.00	5.00
	114	3.025	0.125	0.1579	0.0329	0.0526	100.00	95.24	7.00	7.00	7.00
	115	3.25	0.1410	0.1154	0.0064	0.0128	40.00	19.05	7.00	7.00	5.00
	116	2.725	0.1429	0.1203	0.0075	0.0075	33.33	84.62	6.00	7.00	7.00
	117	2.988	0.0655	0.0774	0.0060	0.0060	.00	.00	5.00	6.00	5.00
	118	3.658	0.0108	0.1081	0.0054	0.0162	38.46	58.82	7.00	7.00	6.00
	119	3.081	0.0938	0.1797	0.0078	0.0078	30.77	69.23	6.00	7.00	6.00
	120	3.281	0.1208	0.1073	0.0067	0	9.09	35.71	7.00	7.00	5.00
Secon	201	2.398	0.1386	0.2772	0	0	16.67	69.23	5.00	7.00	5.00
d	202	2.173	0.1389	0.1736	0.0486	0.0069	33.33	78.57	6.00	6.00	7.00
coder	203	2.476	0.0833	0.2833	0.0167	0.0167	12.50	31.58	6.00	4.00	4.00
	204	2.8068	0.1339	0.2598	0.1260	0.0472	.00	86.67	5.00	5.00	6.00
	205	2.087	0.2	0.32	0.05	0.02	18.18	33.33	6.00	6.00	5.00
	206	2.295	0.2045	0.1932	0.0682	0.0227	54.55	100.00	7.00	7.00	7.00
	207	2.365	0.2667	0.2476	0.0095	0	30.00	28.57	3.00	6.00	4.00

4/3/2 ACTIVITY REVISITED

238

208	3.346	0.0604	0.1342	0.0134	0	18.18	70.00	6.00	7.00	6.00
209	3.081	0.1452	0.1694	0.0323	0.0080	12.50	53.33	5.00	7.00	5.00
210	3.05	0.1168	0.1387	0.0073	0.0146	12.50	70.00	7.00	8.00	7.00
211	2.214	0.2105	0.2526	0.1158	0.0421	16.67	72.27	7.00	6.00	7.00
212	3.053	0.1090	0.1090	0.0256	0.0192	14.29	76.19	7.00	7.00	5.00
213	2.856	0.0915	0.1408	0.0070	0.0070	.00	21.43	5.00	6.00	5.00
214	2.971	0.125	0.1579	0.0526	0.0329	100.00	95.24	6.00	7.00	7.00
215	3.251	0.1410	0.1154	0.0128	0.0064	40.00	19.05	7.00	7.00	5.00
216	2.673	0.1278	0.1128	0.0075	0.0075	33.33	84.62	7.00	6.00	7.00
217	3.007	0.0595	0.0893	0.0060	0.0060	.00	12.50	4.00	6.00	5.00
218	3.675	0.0108	0.1135	0.0216	0.0054	38.46	58.82	7.00	7.00	6.00
219	3.054	0.0938	0.1797	0.0078	0.0078	30.77	69.23	6.00	7.00	6.00
220	2.995	0.1208	0.1074	0	0.0067	9.09	35.71	7.00	7.00	5.00
