TASK COMPLEXITY, TASK TYPE, AND LEARNER-LEARNER INTERACTION:
A REPLICATION STUDY WITH ADULT ESL LEARNERS

by

Daniel Riccardi

A thesis submitted in conformity with the requirements
for the degree of Master of Arts
Graduate Department of Curriculum, Teaching and Learning
Ontario Institute for Studies in Education
University of Toronto

© Copyright by Daniel Riccardi (2014)
Motivated by Robinson’s Cognition Hypothesis (2001), research has investigated whether increasing task complexity leads to a higher number of interaction-driven learning opportunities, including language-related episodes (Swain & Lapkin, 1998), which may lead to L2 learning (Kim, 2009; Robinson & Gilabert, 2007; Nuevo, 2006). Kim (2009a) reports that both task type and proficiency play a role in mediating the relationship between task complexity and the occurrence of such opportunities. However, replication studies are needed to improve the generalizability of results particularly in studies with small sample sizes (Porte, 2012). This study is a conceptual replication of Kim (2009a). Sixteen pairs of adult ESL learners received pre-task modeling and completed four English language tasks differing in their complexity. The results partially confirm Kim’s (2009a) findings while also highlighting the potential mediating effects of pre-task modelling on the relationship between task complexity and learner-learner interaction.
Acknowledgments

The process of completing this thesis was both challenging and rewarding. First and foremost, I would like to thank my thesis supervisor, Dr. Nina Spada. Her immediate and ongoing support over the past two years has allowed my thesis to develop from a proposal to its final form. Her clear feedback and insight was indispensable in helping this thesis take shape, and I was also fortunate to receive support from a very friendly and genuine person who made the process both productive and enjoyable, despite the challenges that inevitably arise when conducting research. I am also very grateful for Dr. Alister Cumming’s insightful comments and feedback, which allowed me to improve the thesis greatly in its final stages. I also benefitted greatly from the courses I completed with Nina and Alister, which both inspired me to choose my thesis topic as well as provided me with the requisite skills to complete this project.

I would also like to thank my good friend and colleague Shereen Seoudi who made a tremendous contribution to my research. Shereen’s unwavering patience, attention to detail, and determination when transcribing a large part of the data allowed me to move through the analysis stage of the thesis efficiently. Shereen also completed the inter-coder reliability checks, a task that required a great deal of patience, focus, and time. Though I will forever wonder how I might repay her, I am extremely grateful for her contributions, and this thesis would not have been completed in a timely manner without her help.

I also extend my thanks to Monique Herbert, who assisted with the statistical analyses in the thesis. Her patience, along with her ability to guide me through a process that was quite new to me was indispensable in ensuring my success. I am also very grateful for the financial support I received from the Social Science and Humanities Research Council (SSHRC) in the form of a Joseph-Armand Bombardier Canada Graduate Master’s Scholarship, which allowed me to prioritize my studies. Finally, these acknowledgements would not be complete without thanking my parents, who have supported me throughout my post-secondary education. Their love and support throughout the years ensured that I would complete my graduate work.
# Table of Contents

Chapter 1: Introduction and Literature Review ................................................................. 1

Literature Review ........................................................................................................... 4
  Defining Task Complexity .......................................................................................... 4
  Sequencing criteria. ...................................................................................................... 5
  Task complexity and linguistic performance .............................................................. 6

Task Complexity and Learner-Learner Interaction ......................................................... 7

Measuring Interaction .................................................................................................... 9

Pre-task modelling ......................................................................................................... 15

Mediating Factors .......................................................................................................... 16

The Original Study: Kim (2009a) .................................................................................. 19

Research Questions ....................................................................................................... 21

Hypotheses .................................................................................................................... 22

Chapter 2: Methodology ............................................................................................... 24

Participants and Context ............................................................................................... 24

Materials ....................................................................................................................... 25
  Picture difference tasks............................................................................................... 25
  Picture narration tasks............................................................................................... 25

Pilot Studies .................................................................................................................. 27
  Results of pilot study 1. ............................................................................................ 27
  Results of pilot study 2. ........................................................................................... 28

The Main Study .............................................................................................................. 29

Participants and proficiency. ......................................................................................... 29

Data collection procedures. ......................................................................................... 29

Data analysis procedures. ............................................................................................ 31

Development of the coding scheme. ............................................................................. 34

Analysis of LREs by task. ............................................................................................ 34

Interrater reliability ...................................................................................................... 35
Chapter 3: Results .................................................................................................................... 36

Original Research Question ........................................................................................................ 36

Picture narration tasks .............................................................................................................. 36
Picture difference tasks ............................................................................................................. 37

Results for Additional Research Questions .............................................................................. 39

Lexical vs. grammatical LREs .................................................................................................... 39
Separate occurrence of interactive LREs and self-corrections .................................................... 40

Picture narration tasks .............................................................................................................. 40
Picture difference tasks ............................................................................................................ 41

Proportion of interactive LREs and self-corrections according to lexical or grammatical
features of English ....................................................................................................................... 42

Picture narration tasks .............................................................................................................. 43
Picture difference tasks ............................................................................................................ 44

Chapter 4: Discussion ................................................................................................................ 46

Original Research Question ........................................................................................................ 46

Picture narration tasks .............................................................................................................. 46
Picture difference tasks ............................................................................................................ 48

Additional Research Questions .................................................................................................. 50

Research question two ............................................................................................................... 50
Research question three ............................................................................................................. 52
Research question four ................................................................................................................. 53

Limitations ................................................................................................................................... 54

Directions for Future Research .................................................................................................. 55

Pedagogical Implications and Conclusion .................................................................................. 57

References ..................................................................................................................................... 59

Appendix A: Sample Materials ..................................................................................................... 64

Appendix B: Transcribing Conventions ....................................................................................... 66

Appendix C: Coding Scheme for Transcribing Conventions ....................................................... 67

Appendix D: Supplemental Statistical Tables .............................................................................. 70
List of Tables

**Table 1** Studies Investigating the Effects of Task Complexity on Learner-learner Interaction
.................................................................................................................................................14

**Table 2** Occurrence of LREs by Proficiency for Picture Narration and Picture Difference Tasks
.................................................................................................................................................20

**Table 3** Differences Between the Original Study (Kim, 2009a) and the Present Replication Study
.....................................................................................................................................................21

**Table 4** Complexity Dimensions Used in the Replication Study and the Original Study
.....................................................................................................................................................29

**Table 5** Latin Square Design Adopted in the Replication Study
.....................................................................................................................................................31

**Table 6** Average Number of LREs Produced per 100 Words by Proficiency During Picture Narration Task
.....................................................................................................................................................37

**Table 7** Average Number of LREs Produced per 100 Words by Proficiency During Picture Difference Task
.....................................................................................................................................................38

**Table 8** Summary of Wilcoxon Signed Ranks Test Results for All Tasks
.....................................................................................................................................................38

**Table 9** Summary of Results by Proficiency, Complexity and Task Type
.....................................................................................................................................................38

**Table 10** Relative Occurrence of Interactive LREs and Self-corrections for Picture Narration Tasks
.....................................................................................................................................................40

**Table 11** Relative Occurrence of Interactive LREs and Self-corrections for Picture Difference Tasks
.....................................................................................................................................................40

**Table 12** Summary of Wilcoxon Signed Ranks Tests Results for all Tasks Analyzed by the Separate Occurrence of Interactive LREs and Self-corrections
.....................................................................................................................................................42

**Table 13** Proportions of the Occurrence of LREs Organized by Type of Interaction and Language Focus for All Tasks
.....................................................................................................................................................43

**Table D1** Average LREs produced per 100 words during the picture narration tasks organized by type of interaction and language focus
.....................................................................................................................................................71

**Table D2** Average LREs produced per 100 words during the picture difference tasks organized by type of interaction and language focus
.....................................................................................................................................................72
List of Figures

**Figure 1** Occurrences of the Four types of LREs by Task and Complexity among the Low Proficiency Learners by Percentage..........................................................45

**Figure 2** Occurrences of the Four Types of LREs by Task and Complexity among the High Proficiency Learners by Percentage.............................................................45
Chapter 1: Introduction and Literature Review

Tasks have long been used by teachers within the second language (L2) classroom to help promote more meaning-oriented, student-centered learning while also providing learners with opportunities to focus on language form. While it is difficult if not impossible to predict precisely how different learners will engage with particular tasks (Ellis, 2000; Van den Branden, 2013), the selection and sequencing of tasks within a syllabus requires a principled and coherent approach based on research (Common European Framework of Reference for Languages [CEFR], 2001; Robinson, 2001). In particular, as L2 pedagogical research becomes increasingly concerned with student-centered classrooms characterized by collaborative learning (Kim, 2009a), a growing number of researchers have investigated how specific task design features may impact interaction among L2 learners (Gilabert, Barón, & Llanes, 2009; Kim, 2012; Révész, 2011; Robinson, 2001, 2007).

Task complexity has emerged as an important variable that may have effects on learner-learner interaction, as well as the learning opportunities that may be afforded by a task (Gilabert et al., 2009; Kim, 2009a; Robinson, 2001). Based on Robinson’s (2001) theoretical framework, researchers have investigated how increasing the complexity of a task may impact interaction outcomes. Robinson’s Cognition Hypothesis predicts that increases in task complexity lead to greater learner-learner interaction, and several studies have found that more complex tasks result in a higher amount of meaning negotiation in the form of interaction moves such as comprehension checks and clarification requests, as well as broader categories of interaction-driven learning opportunities in the form of language-related episodes (LREs) (Kim, 2009a, 2009b, 2012; Gilabert et al., 2009; Robinson & Gilabert, 2007; Nuevo, 2006; Révész, 2011; Robinson, 2001, 2007).

Furthermore, the extent to which task complexity increases such learning opportunities appears to be mediated by both learner proficiency and task type. Kim (2009a), for example, found that more complex tasks actually led to less interaction-driven learning opportunities (LREs) among low proficiency learners, and furthermore, that more complex tasks led to a greater amount of LREs among high proficiency learners for a picture narration task, but not for a picture difference task. Such findings suggest that proficiency and task type are important variables to consider when manipulating the complexity of a task in order to generate increased interaction during task-based learner-learner interaction. However, as in much second language
acquisition (SLA) research, Kim’s study sampled a relatively small number of participants from one English program (n=34), and thus results cannot be generalized to other educational settings with a high degree of confidence (Kim, 2009a). This reflects recognition in the wider field of SLA of the need to replicate research in order to substantiate theoretical and empirical claims made within the literature (Cumming, 2012; Mackey, 2012; Mackey & Gass, 2005; Nassaji, 2012; Norris & Ortega, 2006; Plonsky, 2012; Polio, 2012; Polio & Gass, 1997; Language Teaching Review Panel [LTRP], 2008; Porte, 2012). In the effort to address these concerns, this thesis is a conceptual replication of Kim’s (2009a) study entitled “The effects of task complexity on learner-learner interaction”. Both the original study and the present replication study pursue empirical experiments guided by the theoretical framework of Robinson’s Cognition Hypothesis.

The recognition of the need for more replication studies within the field of SLA has become particularly important given the “increasingly diverse scope and investigation of topics” and the fact that “most of the published studies base generalizations on relatively few subjects” (LTRP, 2008, p. 11). Furthermore, replication studies are particularly critical when the results of studies contradict the results of previous studies or the arguments proposed by theories (Porte & Richards, 2012). For example, as discussed above, Kim (2009a) found that increased task complexity resulted in less interaction-driven learning opportunities among low proficiency learners during performance of a picture narration task, a result that runs contrary to Robinson’s (2001) Cognition Hypothesis.

There are different types of replication research. For example, exact replication studies involve the same participants, tasks, and physical setting as the original study. However, exact replication is not always feasible in the field of SLA given the difficulties associated with including the exact same elements in a replication study (Polio & Gass, 1997; Porte, 2013). Conceptual replications have been proposed as viable alternatives in which a similar research question is investigated as in the original study (Porte, 2012). However, conceptual replications use a new research design to verify the findings of the original study in order to “assess the generalizability of a study or to test the reliability of previous operational definitions by, for example, using a different treatment variable or a different measurement” (Porte, 2012, p. 9). At the same time, conceptual replications typically repeat most elements of the original study as closely as possible so that the two remain comparable. The present study may be defined as a
conceptual replication of Kim (2009a) because it includes an additional treatment variable while also extending the measures used to analyze learner-learner interaction in the original study.

The rationale for conducting this conceptual replication study is threefold. First, this research will either confirm or reject Kim’s findings either partially or fully, thereby allowing researchers to understand more clearly if and how complexity affects interaction, and specifically, how manipulating complexity can lead to more or less interaction among learners of different proficiency learners, which has been shown to be an important aspect of L2 learning (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006; Mackey & Goo, 2007). Second, this research may assist ESL teachers working within universities in making more informed decisions about what tasks to use for which learners at what stage of their L2 development if they hope to encourage interaction among learners. This is particularly pedagogically relevant in larger classrooms with many students where learner-learner interaction may be the main means of student oral production practice. Finally, the answer to this question may help material, course book, and curriculum designers make more informed choices when developing and sequencing tasks designed to promote interaction among L2 learners.

This thesis is organized into four chapters. Chapter 1 provides a comprehensive and detailed review of the research conducted to date regarding task complexity and learner-learner interaction and concludes with a statement of the research questions and hypotheses of the present study. Chapter 2 describes the methodology of the study, including detailed information on the similarities and differences between the original study (Kim, 2009a) and the current replication study. Chapter 3 presents the results of the present analysis, while Chapter 4 serves as a discussion of the results in relation to the existing research and the original study. The final chapter also discusses the limitations and pedagogical implications of the present study and potential directions for future research on the effects of task complexity on learner-learner interaction.
Literature Review

This section provides an overview of the research conducted to date that has investigated the effects of task complexity on learner-learner interaction. The literature review is organized into the following sections: (a) Robinson’s (2001) definition of task complexity and its relation to a task-based syllabus; (b) review of research on task complexity and its effect on learner-learner interaction; (c) measures that have been used to analyze learner-learner interaction; (d) the use of pre-task modelling to encourage collaborative patterns of learner-learner interaction; (e) factors that mediate the relationship between task complexity and learner-learner interaction; (f) review of the original study (Kim, 2009a) that is replicated in this thesis; (g) research questions and hypotheses.

Defining Task Complexity

Among a number of design variables, researchers have argued that task complexity has effects on learner behaviour and production while they engage in task-based interaction (Robinson, 2001; Skehan, 1998). Task complexity is defined by Robinson as “the result of the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner” (2001, p. 29). That is, tasks may be more or less cognitively complex for learners depending on their design, and this affects the way in which learners process information while completing a task.

Within the task complexity aspect of his framework (labeled as the Triadic Componential Framework), Robinson’s (2001) Cognition Hypothesis distinguishes between resource-directing and resource-dispersing (or depleting) dimensions. The presence or absence of a dimension is identified by a +/- component. Robinson argues that tasks manipulated along resource-directing dimensions encourage learners to place a greater focus on linguistic form by making greater demands on working memory and attention, where learners may meet these demands through the use of specific linguistic features relevant to the task. For example, simple narration tasks manipulated along the + here-and-now resource-directing dimension require learners to narrate a set of pictures in the present tense while looking at them, whereas during more complex narrative tasks manipulated by the + there-and-then condition, learners must narrate them in the past tense from memory without the aid of the pictures. Thus, the more complex version requires learners to use more complex verb forms (past tense instead of present tense) along with different determiners (‘that’ and ‘there’ instead of ‘this’ and ‘here’) (Robinson, 2005). Robinson and
others (Cromer, 1974) argue that such increases in complexity resemble the conceptual and linguistic development that occurs in childhood L1 acquisition, where children initially refer to the here-and-now before the there-and-then, and a similar process has also been documented in L2 linguistic development (Behrens, 2001; Meisel, 1987; Von Stutterheim & Klein, 1987). Another example of a resource-directing dimension is +/- few elements. For example, when learners are asked to narrate a story based on a set of pictures, an increase in the number of elements (objects) in the picture renders the task more complex. Finally, the +/- reasoning demand dimension of complexity could refer to picture-based narration tasks in which learners need to decide on the correct order of a sequence of pictures as opposed to having the order provided for them, thus making it more complex.

Resource-dispersing dimensions also make greater processing demands on learners, but they do not direct learners’ attention to particular linguistic forms (Robinson, 2005). Examples include a lack of planning time ( - planning time), performing more than one task at the same time ( - single task), and having previous knowledge of something involved in the task, such as the characters or setting involved in a picture narration task ( + prior knowledge). Robinson contends that as opposed to resource-directing dimensions, increasing complexity along resource-dispersing dimensions does not mirror L1 or L2 conceptual and linguistic development as discussed earlier. Instead, increasing complexity along resource-dispersing dimensions simply challenges a learner to improve their ability to “access and deploy knowledge during the performance of a complex skill” (Robinson, 2005, p.7). Increasing complexity along resource-dispersing dimensions serves an important pedagogical function, as manipulating these dimensions, for example, by removing planning time or the presence of prior knowledge, allow learners to engage in tasks that more closely approximate real world tasks such as answering unanticipated job interview questions in real time without planning (Robinson, 2001). However, these dimensions, as discussed, do not direct learners’ attention to particular linguistic forms.

**Sequencing criteria.** In the broader effort to propose sequencing criteria for tasks used within a task-based-syllabus, Robinson proposed the Triadic Componential Framework discussed above that allows researchers and teachers to sequence tasks according to their relative complexity while researching the effects of specific variables on performance and interaction. Robinson’s ultimate goal is to provide a framework that would help inform the way in which tasks might be sequenced within a syllabus.
Given that only resource-directing dimensions promote learners’ attention to specific
features of linguistic code, Robinson argues that making tasks simpler first along resource-
depleting dimensions and more complex along resource-directing dimensions, “allows [for]
optimum resource allocation to satisfy the linguistic demands of the task” (Robinson, 2001,
p.31). By increasing complexity along resource-directing dimensions, instructors would promote
learner engagement with more conceptually and linguistically demanding tasks (e.g., narrating in
the past tense instead of the present tense). At the same time, by decreasing complexity along
resource-dispersing dimensions (e.g., providing planning time), learners would have more
cognitive resources available to fulfill the linguistic demands of tasks. This would be particularly
important in the earlier stages of a syllabus where instructors may want to avoid overwhelming
learners by increasing complexity along resource-dispersing dimensions (e.g., taking away
planning time). In this manner, tasks might be sequenced within a syllabus in order of their
relative complexity, which would create the conditions for L2 development in a principled
manner. Ultimately, simpler tasks would gradually be made increasingly complex within a
syllabus along both resource-directing and dispersing dimensions in the effort to approximate
real world tasks.

**Task complexity and linguistic performance.** In regards to linguistic performance,
Robinson (2001, 2005) predicts that tasks made more complex along resource-directing
dimensions will result in speech that is characterized by both greater complexity and accuracy,
but decreased fluency. He maintains that learners are pushed to produce more complex language
(e.g., past tense instead of present tense) to meet the more complex linguistic demands of the
task. In terms of accuracy, Robinson (2005) suggests that more complex tasks will induce
learners to pay more careful attention to their own speech production, and in particular, the way
in which the grammar of a learner’s L1 and L2 differ. The cognitive demands made by resource-
directing dimensions in more complex tasks thus may push learners to produce more accurate
speech, particularly with language that may be grammatically relevant to a task (Robinson, 2001)
(e.g., use of the past tense when a task is made more complex along the + there-and-then
resource-directing dimension).

Skehan (1998, 2007) has proposed an alternative hypothesis to Robinson’s Cognition
Hypothesis, arguing that learners have limited attentional resources, and thus more complex
tasks may lead them to prioritize either meaning or form. Skehan’s “Trade-off Hypothesis” thus
predicts that more complex tasks will not necessarily lead to increases in the complexity of learner speech (Skehan & Foster, 2001). Furthermore, since complexity and accuracy utilize the same attentional resources, in most cases, learners may prioritize only one at a time, and increased linguistic complexity normally occurs at the expense of accuracy. These trade-off effects do not depend simply on task complexity but on task conditions such as the inclusion of planning time, or the degree of task structure, such as a greater or lesser amount of narrative structure provided to learners during picture narration tasks (Skehan & Foster, 2012).

In addition to these two competing claims for the effects of task complexity on speech production, Robinson (2001) also hypothesized that increased complexity would result in a greater amount of interaction between learners in the form of the negotiation of meaning, which is described in more detail below. Although the effects of task complexity on a speaker’s oral production has been researched extensively in terms of the degree of complexity, accuracy, fluency and lexical variety used during L2 learner speech production, this review will focus attention on the effects of task complexity on learner-learner interaction given its direct relevance to the present study.

**Task Complexity and Learner-Learner Interaction**

Influenced by Long’s Interaction Hypothesis (Long, 1983), many SLA researchers sought to understand the relationship between task design variables and their effects on interaction, which was mainly defined in reference to the negotiation of meaning. This negotiation was generally measured through interactional moves such as clarification requests, comprehension checks, confirmation checks, and recasts (Mackey, 2007).

In regards to task complexity, Robinson hypothesized that more complex tasks would lead to more communication breakdowns, leading the listener to produce more clarification requests and confirmation checks during task performance. In Robinson’s (2001) first empirical study that addressed this question, he investigated the effects of two different task complexity dimensions on learners’ production of clarification requests and confirmation checks, the +/- few elements resource-directing dimension and the +/- prior knowledge resource-dispersing dimension. In Robinson’s (2001) study, pairs of learners completed two versions of a map task in which the speakers dictated a route for the listeners to draw on their map. In the simple version, learners were provided with a map of a small area (+ few elements) of which they had prior knowledge,
their university campus, whereas the complex version included a map of a larger area of Tokyo (- few elements) that was likely unfamiliar to participants.

Results indicated that the more complex tasks resulted in significantly greater amounts of confirmation checks, with a trend towards the greater use of clarification requests. Thus, Robinson demonstrated that more complex tasks resulted in a greater number of one type of interaction, confirmation checks, but only a trend towards a greater number of clarification requests. However, Robinson manipulated the tasks used in this study by both resource-directing (- few elements) and dispersing dimensions (- prior knowledge) simultaneously, making it difficult to ascertain how the two dimensions may have resulted in the increased interaction.

In a later study, Robinson (2007) isolated the effects of resource-directing and resource-dispersing dimensions by manipulating a picture narration task solely by the +/- reasoning demands dimension. In this study, Robinson (2007) used three interactive narrative tasks arranged on a continuum from low, middle to high complexity (simple, +complex, ++complex), The simple version involved a single character who wishes to build a house, whereas the more progressively complex versions involved a character who must design a house based on his or her perceptions of other characters’ opinions, a condition which placed increased reasoning demands on the learners. During this task, the speaker/storyteller first decided on a sequence for the pictures provided and then narrated a complete story to the listener/sequencer, who was asked to sequence the pictures in the order described.

In addition to determining the effects of task complexity on speech production in terms of lexical and syntactic complexity, fluency and accuracy, Robinson measured interaction in terms of confirmation checks and clarification requests, while also expanding these measures to incorporate the number of turns taken and the amount of learner uptake of linguistic forms relevant to the cognitive/conceptual demands of the task made salient in the input. Furthermore, Robinson measured the mediating effect of anxiety as an individual difference variable. Learners completed a questionnaire that measured the input, output and processing anxiety they might experience during task production, in which input and processing anxiety were related to listening and decoding L2 input while output anxiety referred to writing or speaking in the L2. Robinson found that increased complexity clearly led to significantly more interaction, with a greater number of turns, as well as confirmation checks, clarification requests, and uptake. Furthermore, this increased progressively as learners completed simple, +complex, and
++complex tasks. Finally, when participants were split into low and high anxiety groups, Robinson found that the low anxiety groups produced more complex speech.

Thus, the two studies conducted by Robinson (2001, 2007) generally lend support to the claim that increasing task complexity results in greater amounts of interaction in terms of increased confirmation checks and clarification requests as well as the number of turns and uptake. However, only the second study found a significant result for all interaction measures, and the non-significant findings for clarification requests in the first study seems to suggest that effects may differ according to the type of interaction measures employed within the study or perhaps the type of task in which learners engage. An obvious limitation of these studies is that they employed only two task types in their design, an interactive one-way map task (2001) and a picture narration task (2007), leaving it unclear how task type may mediate this relationship. Furthermore, while ID variables (e.g., anxiety, proficiency, etc.) also appear to influence learner speech production (Robinson, 2007), their effects on interaction were not analyzed in either study. Nonetheless the finding that more complex tasks led to increased uptake is encouraging as it suggests that increased complexity might heighten learners’ attention to forms made salient in the input. Other research has extended this inquiry to examine how task complexity affects the occurrence of learner-learner interaction that is primarily concerned with linguistic form (Kim, 2009a, 2009b, 2012; Nuevo, 2006; Révész, 2011).

In summary, Robinson’s two studies identified three key issues of interest to the present replication study: 1) attention should be paid to the types of interaction measures employed within this area of research, 2) increasing task complexity may increase learner-learner interaction concerned with linguistic form, and 3) task type and perhaps individual difference variables appear to mediate the relationship between task complexity and learner-learner interaction. In the following section, I will elaborate on each of these three issues in greater detail.

**Measuring Interaction**

With the exception of the analysis of learner uptake in Robinson (2007), the research discussed above has measured interaction mainly in terms of the negotiation of meaning (i.e., clarification requests, comprehension checks, confirmation checks, and recasts). However, other research that has investigated the effects of task complexity on learner-learner interaction has
extended this inquiry to examine how learners attend to L2 forms during task-based interaction (Kim, 2009a, 2012; Nuevo, 2006; Révész, 2011).

Nuevo (2006) investigated both the occurrence of interaction-driven L2 learning opportunities as well as learner development of English past tense and locative prepositions in relation to task complexity. In addition to measuring interaction in terms of the negotiation of meaning, Nuevo employed a total of nine interaction measures—including metalinguistic talk, hypothesis formulation, recasts and uptake, other repetitions and self-repair—to examine interaction related to form as well. Nuevo argues that metalinguistic talk approximates the more commonly used term LRE (Swain & Lapkin, 1998), which is described in detail in the following section. Metalinguistic talk refers to utterances in which learners discuss or ask questions about the language forms they use during interaction. Hypothesis formulation refers to moments in which learners attempt to guess correct forms during task production. Nuevo defines recasts as “more target-like reformulations of a previous interlocutor’s utterance” and the uptake of recasts is defined as “a learners’ [sic] exact repetition of the recast” (2006, p. 89). Other repetitions simply refer to the exact repetition of an interlocutor’s utterance, whereas self-repairs are defined as “learner self-correction of non-target like speech to more target like speech in the same turn or across turns” (2006, pp. 92-93) without any feedback from an interlocutor.

The results indicated that different task complexity conditions appeared to encourage different kinds of L2 learning opportunities. Whereas low complexity tasks resulted, for example, in higher numbers of recasts and greater uptake, along with comprehension checks and other repetitions, high complexity tasks resulted in more hypothesis formulation. Nuevo speculates that low complexity tasks may free up attentional resources allowing learners to provide each other with feedback (i.e., recasts), whereas high complexity tasks lead to more communication breakdowns, which often trigger confirmation checks. However, no clear patterns are established in terms of task complexity and the amount of interaction, as these distinctions appear to differ among the interaction measures employed as well as the different narrative and decision-making task types administered to learners. Furthermore, no significant differences in L2 learner development of the two target features were found between the groups that performed low and high complexity tasks.

Nuevo’s study outlines the importance of investigating moments in which L2 learners pay attention to form through their analysis of the occurrence of metalinguistic talk, hypothesis
formulation, corrective feedback, and self-repair during learner-learner interaction. Other research has addressed this issue directly by analyzing interaction in terms of LREs.

LREs have been widely used to define potential learning opportunities during interaction in which learners pay attention to linguistic form through talking about, questioning or correcting the language they are producing (Swain & Lapkin, 1998; Williams, 1999). Swain and Lapkin (1998) argue that LREs that occur during collaborative dialogue may provide occasions for learners to generate and test hypotheses related to the target language while also allowing them to revise or extend their existing knowledge of the target L2. LREs refer to a large number of discourse moves such as learner-learner meta-talk, requests for assistance, and feedback such as recasts (Kim, 2009a; Williams, 1999).

Swain and Lapkin (1998) also suggest that LREs may provide occasions for L2 learning, a claim that is supported by later research on the effects of task complexity on learner-learner interaction (Kim, 2009b, Kim, 2012). These studies have also demonstrated that increasing task complexity may increase the occurrence of LREs during collaborative dialogue, which in turn may result in L2 learning for specific grammatical structures. Given that L2 learning was not investigated within the present study, this review will focus attention on task complexity and its effects on learner production of LREs during task-based learner-learner interaction.

Four studies have found that increasing task complexity leads learners to produce higher numbers of LREs during task-based learner-learner interaction (Gilabert et al., 2009; Kim, 2009b, 2012; Révész, 2011). Kim (2009b) investigated the effects of increasing task complexity on the occurrence of LREs related to the formation of past tense and question forms. She demonstrated that more complex tasks encouraged learners to produce more LREs in relation to the two target structures. A similar result was obtained in her later study (Kim, 2012), which demonstrated that increased task complexity along the + reasoning demands and – few elements dimensions led to a higher occurrence of LREs related to the development of English question formation. In addition, Gilabert et al. (2009) found that a complex picture narration task, manipulated by the + there-and-then dimension, as well as a complex map task, manipulated by the – few elements dimension, led learners to produce significantly more LREs. Finally, Révész (2011) also demonstrated a significant increase in the number of LREs when small groups of 3-4 participants engaged in a more complex argumentative task manipulated by the + reasoning demands and – few elements dimensions.
Thus, these three studies suggest that in general when interaction is measured in terms of LREs, it appears to occur at higher rates in more complex tasks. However, as described earlier, research also suggests that task type and proficiency, both of which are described below, may mediate the number of LREs produced during interaction. Before investigating these mediating effects, the occurrence of self-corrections (or self-repair) during learner-learner task-based interaction will be discussed first in an effort to illustrate how existing research has operationalized self-corrections both as an independent interaction measure and as an aspect of LREs. This is followed by a discussion of pre-task modelling, a type of task demonstration, which has been used by researchers (Kim, 2013; Kim & McDonough, 2011) to encourage learner production of LREs and increased collaboration during task-based pair work.

Following Swain (1998), a few studies have included self-corrections as LREs when investigating the effects of task complexity on learner-learner interaction (Kim, 2009a, 2009b, 2012; Révész, 2011). Other studies have analyzed self-corrections separately (Nuevo, 2006). Gilabert et al. (2009) analyzed self-corrections as LREs but also analyzed them separately, partly due to the results of a pilot study that demonstrated that pairs of learners produced high numbers of self-corrections but very few instances of other types of LREs. In that study, the findings indicated that both self-corrections and LREs (with self-corrections included) occurred at higher rates during the complex versions of two task types, a picture narration task and map task. However while self-corrections occurred at higher rates during a third task type, a decision-making task, LREs (again, with self-corrections included) did not.

As described above, Nuevo (2006) used nine different interaction measures, including self-repair and metalinguistic talk. She argues that metalinguistic talk approximates the term LRE, even though it occurs “at the level of the utterance”, whereas an LRE may occur over several turns and encompass several different interaction moves within one episode (e.g., comprehension checks, recasts, etc.). Significant differences for metalinguistic talk were not found between simple and complex versions of a picture narration task nor on a decision making task. Self-repairs did not reach significance for either task type. Thus, while learners produced more self-corrections during complex tasks in Gilabert et al. (2009), they did not in Nuevo’s study. Due to these conflicting results, research on the relative occurrence of self-corrections as a separate interaction measure is warranted in addition to measuring LREs with self-corrections both included and excluded.
In an attempt to provide clarity on LREs with and without self-corrections, Fortune and Thorp (2001) investigated small group interaction during the completion of dictogloss activities, making a clear distinction between LREs that occur collaboratively and self-corrections, which by definition, occur individually. They argue that text reconstruction is a collaborative activity in which learners must build off of one another’s contributions in order to complete a task successfully. Thus self-corrections are not included in their definition of LREs. Given these considerations, as well as the fact that Swain and Lapkin’s (1998) definition of LREs, which includes self-corrections, was used to analyze LREs in the Kim (2009a), the present study analyzes LREs with self-corrections included, but also conducts additional analyses that distinguish between LREs that occur collaboratively between two learners (with self-corrections excluded) and LREs that occur individually as self-corrections. While self-corrections are still considered instances of LREs in the present study, they will be referred to simply as self-corrections in the remainder of this thesis.

To summarize the studies discussed in this chapter, Table 1 provides an overview of the complexity variables, interaction measures, and the populations that have been examined in these studies investigating the relationship between task complexity and learner-learner interaction.
<table>
<thead>
<tr>
<th>Study</th>
<th>Complexity variable(s)</th>
<th>Interaction measure(s)</th>
<th>Participants, L1(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilabert et al.</td>
<td>+/- here-and-now</td>
<td>Confirmation checks</td>
<td>60, Spanish</td>
</tr>
<tr>
<td>(2009)</td>
<td>+/- few elements</td>
<td>Clarification requests, Comprehension checks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+/- reasoning demands</td>
<td>Recasts, LREs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Repairs</td>
<td></td>
</tr>
<tr>
<td>Kim (2009a)</td>
<td>+/- reasoning demands</td>
<td>LREs</td>
<td>191, Korean</td>
</tr>
<tr>
<td>Kim (2009b)</td>
<td>+/- reasoning demands</td>
<td>LREs</td>
<td>34, ESL(^b)</td>
</tr>
<tr>
<td></td>
<td>+/- few elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim (2012)</td>
<td>+/- reasoning demands</td>
<td>LREs</td>
<td>191, Korean</td>
</tr>
<tr>
<td></td>
<td>+/- few elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuevo (2006)</td>
<td>+/- reasoning demands</td>
<td>Recasts and uptake</td>
<td>113, ESL (predominantly Spanish)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clarification requests, Comprehension checks, Confirmation checks, Hypothesis formulation, Metalinguistic talk, Self-repair, Noticing a linguistic deficiency, Other repetitions (imitation)</td>
<td></td>
</tr>
<tr>
<td>Révész (2011)</td>
<td>+/- reasoning demands</td>
<td>LREs</td>
<td>43, ESL</td>
</tr>
<tr>
<td></td>
<td>+/- few elements</td>
<td>Confirmation checks</td>
<td></td>
</tr>
<tr>
<td>Robinson (2001)</td>
<td>+/- prior knowledge</td>
<td>Confirmation checks</td>
<td>44, Japanese</td>
</tr>
<tr>
<td></td>
<td>+/- few elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robinson (2007)</td>
<td>+/- reasoning demands</td>
<td>Confirmation checks</td>
<td>42, Japanese</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of turns, Uptake</td>
<td></td>
</tr>
</tbody>
</table>

---

\(^a\) All studies were conducted with L2 English learners

\(^b\) “ESL” refers to learners of diverse L1 backgrounds
Pre-task modelling

Pre-task modelling refers to a type of demonstration that models how pairs or small groups can work collaboratively during task-based interaction (Kim, 2013). Such modelling highlights how learners may work collaboratively while completing tasks through strategies such as negotiating the meaning of words, asking for clarification, or correcting each other’s errors. These demonstrations also provide examples of how learners may attend to L2 forms, vocabulary use, or pronunciation during learner-learner interaction, for example, by asking questions about verb tenses, the spelling of words, or self-correcting their own use of L2 forms, lexis or pronunciation. A small number of studies have employed pre-task modelling in order to encourage the production of LREs during learner-learner interaction (Kim, 2013; Kim & McDonough, 2011; Swain & Lapkin, 1998). While Swain and Lapkin (1998, 2001) did not investigate such effects directly, all participants in both studies viewed videos that served as demonstrations of collaborative patterns of interaction prior to completing the tasks. These videos portrayed learners having discussions about lexical items and grammatical forms, as well as asking each other questions and providing feedback. Along with brief lessons on relevant grammatical forms and practice sessions, they served as examples of how tasks could be completed while also demonstrating how learners could attend to linguistic forms collaboratively during task production.

Kim and McDonough (2011) addressed the effects of pre-task modelling on learners’ production of LREs directly by splitting participants into two groups, one that received pre-task modelling in the form of short demonstration videos as well as explicit instructions on how to work collaboratively, and one control group that did not view the videos or receive any instructions. Overall, “the pre-task modelling videos demonstrated how to carry out the tasks collaboratively through idea-sharing, asking and answering questions, and providing feedback, and how to successfully resolve linguistic questions and problems during interactive tasks” (p. 188). They found that learners that received pre-task modelling and instructions produced significantly more grammatical and lexical LREs than the control group. Furthermore, based on Storch’s framework of collaborative pair dynamics (2002), results indicated that the treatment group worked more collaboratively than the control group by providing each other with consistent feedback, engaging with each other’s ideas, encouraging partners to participate, and making similar contributions to task completion.
In a later study, Kim (2013) investigated the effects of pre-task modelling on the occurrence of LREs both during planning time (analyzed through think aloud protocols) and during performance of the task. In her analysis, she also measured learners’ question development based on Pienemann and Johnston’s (1987) developmental stages, and divided measures into LREs involving questions, LREs involving other grammatical forms, and lexical LREs. The videos emphasized instances of LREs involving question formation given the focus of the study. Results from the think aloud protocols demonstrated that learners produced higher numbers of all three types of LREs during the planning stages than the control group did. However, during task performance, only LREs involving question forms reached statistical significance. Furthermore, the treatment group significantly outperformed the control group on oral production tests measuring question development. Thus, despite the small amount of research conducted, there is some initial evidence that pre-task modelling may increase the occurrence of LREs during learner-learner task-based interaction, though it is less clear if this modelling results in a higher occurrence of LREs that do not involve specific forms targeted within a study (Kim, 2013).

**Mediating Factors**

While the research discussed above generally suggests that increasing task complexity may lead to higher numbers of interaction-driven learning opportunities, a number of these studies have also indicated that the relationship between task complexity and learner-learner interaction appears to be mediated by task type and learner proficiency (Gilabert et al., 2009; Kim, 2009a, 2009b; Nuevo, 2006; Robinson, 2001, 2007).

Research conducted to date suggests that the effects of task complexity on the occurrence of different interaction measures may differ according to task type. As discussed above, Robinson (2001) demonstrated a significant increase in confirmation checks for an instruction-giving map task, but not for clarification requests. However, results from Robinson (2007) involving a picture narration task indicated a significant increase both in the number of confirmation checks and clarification requests. Nuevo (2006) reported that while less complex tasks led to a higher incidence of particular interaction measures for two task types, a more complex picture narration task led to more hypothesis formulations while a more complex decision-making task led to more confirmation checks.

Kim (2009b) analyzed the occurrence of LREs related to question forms and past tense
using a total of 12 tasks manipulated along a continuum with three levels of complexity: simple, + complex, and ++ complex. The simple tasks did not require any reasoning demands, and were categorized as information gap tasks in which learners had to exchange information to complete profiles for different characters. The + complex tasks were manipulated by the + reasoning demands dimension by requiring learners to make a decision. For example, learners had to decide what part-time job would be appropriate for the characters based on their completed profiles. The ++ complex tasks were made more complex by adding the – few elements dimension, requiring learners to make these decisions based on a higher number of criteria. Thus, simple tasks were categorized as information gap tasks and the more complex ones as decision-making tasks. As a result, task type was directly dependent on the degree of task complexity.

Although results indicated that more complex tasks elicited higher numbers of LREs related to question forms (but not past tense) as complexity progressed from simple to ++ complex tasks, the conflation of task type and task complexity makes it impossible to separate the relative effects of either variable on the occurrence of LREs.

In Gilabert et al. (2009) the importance of task type was addressed more explicitly by measuring the effects of task complexity on six interaction measures between pairs in terms of three different task types: an instruction-giving map task, a picture narration task, and a decision-making task. Interaction was measured in terms of the negotiation of meaning (i.e., confirmation checks, comprehensions checks, and clarification requests), recasts, LREs, and self-corrections. Results indicated that increased complexity resulted in greater amounts of interaction, but this differed for different interaction measures as well as across task types. For example, the complex picture narration task led to significantly greater numbers of clarification requests and a strong trend towards more confirmation checks but significantly lower numbers of comprehension checks. In fact, comprehension checks were exclusively associated with the instruction-giving map task. As explained earlier, LREs (with self-corrections included) and self-corrections (when analyzed separately) occurred at higher rates only during the complex narration and map tasks, while recasts occurred at similar rates in both simple and complex versions of all three tasks. Finally, task complexity had no impact on any of the interaction measures in the decision-making task aside from the number of self-corrections.

In summary, though the mediating effects of task type on the relationship between task complexity and learner-learner interaction is not entirely clear in Kim’s study (2009b), other
research (Gilabert et al., 2009; Nuevo, 2006; Robinson, 2001, 2007) suggests that researchers must consider task type as an important variable when predicting the effects task complexity might have on learner-learner interaction.

In addition to task type, Robinson (2005) stresses the importance of investigating proficiency as a mediating variable in the relationship between task complexity and learner-learner interaction. He argues that while his predictions for task complexity should hold for learners of any proficiency, empirical research is needed to investigate this claim. A small number of studies have addressed the mediating effect of proficiency directly (Kim, 2009a, 2009b)

In the study replicated in this thesis, Kim (2009a) divided learners into high and low proficiency groups based on their placement within a university program as well as their paper-based Test of English as a Foreign Language (TOEFL PBT) (Educational Testing Service [ETS], 2014) scores. She found that while high proficiency learners produced more LREs during a complex picture narration task, low proficiency learners produced more during the simple version. In addition, low proficiency learners produced more LREs during a complex picture difference task while high proficiency learners produced similar numbers of LREs during both simple and complex versions. Based on these results, Kim concludes that proficiency plays an important mediating role on the effects of task complexity on learner-learner interaction.

In another study by Kim (2009b) discussed above, the effects of task complexity on learner-learner interaction also differed according to proficiency level. In that study, learners were divided into high and low based on the results of Test of English for International Communication (TOEIC) Bridge scores (ETS, 2008). Results indicated that complex tasks encouraged learners to produce more LREs in relation to two target structures (i.e. past tense and question forms), and this led to greater development, particularly of developmentally advanced question forms. In addition, proficiency pair-groupings had statistically significant effects, with high-high pairs producing the greatest number of LREs related to the structures in question, while low-low pairs tended to produce LREs focused on lexical items. In terms of development, high-high pairs and high-low pairs benefitted most, with low-low pairs scoring low on the post-tests. These two studies (Kim, 2009a, 2009b) provide additional evidence for the argument that proficiency plays an important role in determining the effects of task complexity on learner-learner interaction and L2 development. The mediating effects of task type and proficiency on
the relationship between task complexity and learner-learner interaction reported in Kim (2009a) are of particular interest given that the research reported here is a conceptual replication of that study.

The Original Study: Kim (2009a)

Kim (2009a) investigated the effects of task complexity on the occurrence of interaction-driven learning opportunities among learners of different proficiency levels. In her research, 34 international students from an intensive English program at a US university participated in simple and complex versions of a picture narration task and a picture difference task. Interaction-driven learning opportunities were operationalized as LREs and were mainly classified as grammatical or lexical in nature. As discussed above, participants were divided into two proficiency levels (low and high) based on their enrolment status in their program and their TOEFL PBT scores (ETS, 2014). The study was motivated by the following research question: “Does task complexity impact the occurrence of LREs during task-based interaction between ESL learners of different proficiency levels?” (Kim, 2009a, p. 257)

As indicated above, Kim (2009a) found that when students were split into two proficiency groups and completed both simple and complex versions of the two types of task, the predictions of the Cognition Hypothesis (CH) - that increasing task complexity leads to increased interaction in terms of the negotiation of meaning and LREs - were only partially supported. For the picture narration task, high proficiency learners produced more LREs in the complex version (in line with the CH), while low proficiency learners produced more in the simple version (counter to the CH). For the picture difference task, the low proficiency learners produced more LREs in the complex version (in line with the CH) while the high proficiency learners produced a similar number of LREs in the simple and complex versions (counter to the CH). Table 2 summarizes these results. Furthermore, low proficiency learners produced more lexical LREs during the simple and complex picture difference tasks as well as the simple picture narration task, while high proficiency learners produced more lexical LREs in the simple and complex picture difference tasks. In summary, Kim’s study indicates that task type and proficiency may mediate the relationship between task complexity and the amount of learner-learner interaction in terms of the occurrence of LREs.
Table 2 Occurrence of LREs by Proficiency for Picture Narration and Picture Difference Tasks (Kim, 2009a)

<table>
<thead>
<tr>
<th>Task type</th>
<th>Picture narration</th>
<th>Picture difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (9 pairs)</td>
<td>Significantly more LREs during simple task(^a)</td>
<td>Significantly more LREs during complex task.</td>
</tr>
<tr>
<td>High (8 pairs)</td>
<td>Significantly more LREs during complex task</td>
<td>No difference between LREs in simple and complex tasks(^a)</td>
</tr>
</tbody>
</table>

\(^a\)Finding runs counter to the Cognition Hypothesis (Robinson, 2001)

As Kim and others have suggested, studies that have investigated the relationship between task complexity and learner-learner interaction are comparatively rare in relation to studies focused on learner production in terms of the complexity, accuracy, and fluency of L2 learners’ speech (Kim, 2009a; Robinson & Gilabert, 2007). Studies that investigate the potential interaction effects between learner individual difference variables such as proficiency and task variables such as complexity remain even more rare. Kim’s study takes an important step in attempting to clarify how proficiency and task type may mediate the relationship between task complexity and interaction.

However, the generalizability of Kim’s study remains limited “as a relatively small number of participants was sampled from one Intensive English Program in the USA” (Kim, 2009a, p. 264). This present replication study should help to further clarify how proficiency plays a mediating role in the effects of task complexity on learner-learner interaction. The number of LREs produced by both high and low proficiency learners on both the simple and complex picture narration and picture difference tasks in this study will be compared with Kim’s results.

Before the research questions and hypotheses are stated, it is important to note how the research design of the present conceptual replication study differs from the original study, as these modifications have implications for the comparability of the two studies, as well as for the wider literature on task complexity and learner-learner interaction. While these modifications are described in detail in Chapter 2, Table 3 provides a summary of these changes in order to highlight these differences between the studies.
Table 3 Differences between the Original Study (Kim, 2009a) and the Present Replication Study

<table>
<thead>
<tr>
<th></th>
<th>Kim (2009a)</th>
<th>Conceptual Replication Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Design</td>
<td>No pre-task modelling</td>
<td>Pre-task modelling</td>
</tr>
<tr>
<td></td>
<td>Counterbalancing of tasks either absent or unreported</td>
<td>Tasks counterbalanced</td>
</tr>
<tr>
<td>Materials</td>
<td>Four tasks (Kim, 2009a):</td>
<td>Four tasks:</td>
</tr>
<tr>
<td></td>
<td>One simple and one complex picture narration task</td>
<td>One simple and one complex picture narration task (both adapted from</td>
</tr>
<tr>
<td></td>
<td>One simple and one complex picture difference task</td>
<td>Gilabert et al., 2009)</td>
</tr>
<tr>
<td>Coding Scheme</td>
<td>LREs</td>
<td>LREs (analyzed as a whole as in Kim)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interactive LREs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-corrections</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>Raw numbers of LREs</td>
<td>Average numbers of LREs produced per 100 words (resulting in a ratio)</td>
</tr>
</tbody>
</table>

Research Questions

As discussed above, the research question in Kim was worded as follows: “Does task complexity impact the occurrence of LREs during task-based interaction between ESL learners of different proficiency levels?” (2009a, p. 257). Kim used two different task types with simple and complex versions to investigate this question: a picture narration task and a picture difference task. As discussed above, research has indicated that task type plays an important mediating role on the relationship between task complexity and learner-learner interaction (Gilabert et al., 2009; Kim, 2009a; Nuevo, 2006). Since Kim analyzed her data and discussed her findings both in relation to the effects of task complexity and task type, this replication study has reformulated her research question to account for this design. Thus, the present study is motivated by the following research question: (1) Do task complexity and task type impact the occurrence of LREs during task-based interaction between ESL learners from different proficiency levels?

In addition to the main research question explored in this conceptual replication study, additional research questions arose during the data analysis stage of the research based on other
relevant studies described above (Fortune & Thorp, 2001; Gilabert et al., 2009; Nuevo, 2006). First, while Kim analyzed the relative occurrence of lexical and grammatical LREs descriptively, the present study analyzed these data through statistical tests in order to determine if either lexical or grammatical LREs occurred more frequently at statistically significant rates during either simple or complex tasks that vary according to task type (i.e., picture narration tasks or picture difference tasks). Thus, in the effort to extend Kim’s findings, this study also investigates the following research question: (2) Do learners at different proficiency levels produce more lexical or grammatical LREs depending on task complexity and task type?

Furthermore, the present study also extends Kim’s original research question in order to investigate the effect of task complexity and task type on LREs involving both participants in a pair (labeled as interactive LREs in the present study) and its separate effect on self-corrections as an independent interaction measure. While LREs were identified and analyzed in the same manner as they were in Kim in order to render the two studies comparable, the present replication further coded and analyzed LREs as interactive LREs or self-corrections. This represents a conceptual addition to the research design of the original study, which will demonstrate whether results similar to Kim’s are found when LREs are divided into interactive LREs that occur between learners or individually as self-corrections. Finally, as a follow up to the second and third research questions, the relative occurrence of lexical and grammatical LREs are examined in terms of whether they are interactive or individual in nature leading to the final two research questions: (3) Do task complexity and task type effect the occurrence of LREs produced interactively or individually among L2 learners of different proficiency levels? (4) Do task complexity and task type effect the proportional occurrence of lexical or grammatical LREs that are produced interactively or individually among L2 learners of different proficiency levels?

**Hypotheses**

(1) In line with Kim’s findings, I predict that both task type and proficiency will play important roles in mediating the relationship between task complexity and interaction-driven L2 learning opportunities. Based on Kim’s results, I propose the following four hypotheses for the two proficiency groups and two task types. For the picture narration tasks, the low proficiency learners will produce significantly more LREs in the simple version while the high proficiency

---

^c As discussed above, despite the fact that self-corrections are analyzed and labeled separately, because Swain and Lapkin’s (1998) definition of LREs includes self-corrections, they are still considered LREs in the present study.
learners will produce more in the complex version. For the picture difference tasks, the low proficiency learners will produce significantly more LREs in the complex version than in the simple version while the high proficiency learners will produce similar numbers of LREs in the simple and complex versions.

(2) Based on Kim’s (2009a) descriptive results, I hypothesize that when the data are submitted to statistical testing, both the low and high proficiency learners will produce more lexical LREs than grammatical LREs during the simple and complex picture difference tasks while the low proficiency learners will also produce more lexical LREs in the simple picture narration task.

(3) Given that no previous studies investigating the effects of task complexity on learner-learner interaction have analyzed LREs with self-corrections excluded (coded in this study as interactive LREs) and categorized as a separate interaction measure (coded in this study as self-corrections), I predict that there will be no differences between simple and complex tasks in either proficiency group when interactive LREs and self-corrections are analyzed independently.

(4) In regards to lexical and grammatical LREs and their proportional occurrence as either interactive LREs or self-corrections, given the lack of research conducted to date on such a division, I also predict that there will be no differences for the occurrence of each type of LRE between simple and complex tasks (lexical and grammatical interactive LREs, as well as lexical and grammatical self-corrections).
Chapter 2: Method

This chapter outlines the research methods used in the study. First, general information regarding the program and the participants involved in the pilot and main studies are outlined. I then provide information on the materials used. This is followed by detailed information on the two pilot studies and their results. Finally a description of the data collection and data analysis procedures used in the main study is provided along with detailed information about the learners who participated in the main study.

Participants and Context

This conceptual replication study was conducted with a different sample of participants from a similar population of learners who participated in Kim’s study (2009a), namely, international ESL students enrolled in an intensive university English program in Toronto, Canada, as opposed to a similar program in a US city. Though the program emphasizes the teaching of English for Academic Purposes (EAP), it is communicative in its orientation and students participate in interactive tasks involving pair and small group work on a daily basis. The students within the program come from a wide variety of countries and spend anywhere from two weeks to a few years improving their English proficiency within the program. Students have a variety of aims such as fulfilling English language requirements for Canadian and other English-medium universities or improving their job prospects either in Canada or elsewhere in the world where English proficiency may be advantageous or required. The pilot study participants were recruited from the same program and in the same manner as the main study participants to help ensure comparability between the participants across the studies. I visited classrooms during regularly scheduled instruction and explained the goals of the research as well as the commitment involved in as simple language as possible (a handout with this information was also provided). Students who were interested provided their names on a form, which was kept strictly confidential from their classroom instructor.

In order to facilitate future replications and to help readers determine its generalizability (Mackey & Gass, 2005), a biodata questionnaire was administered to all participants. This questionnaire collected the major demographic characteristics of participants such as gender, age, L1 background as well as the results of any standardized language proficiency tests they may have taken in the past.
Materials

Four oral interactive tasks were used: two picture narration tasks and two picture difference tasks. Learners completed two versions of each task type, one simple and one complex. Thus, the materials administered to learners were as follows: one simple picture narration task, one complex picture narration task, one simple picture difference task, and one complex picture difference task.

Picture difference tasks. In Kim (2009a) each pair of learners was provided with two pictures labeled either as Picture A or B. The two pictures were highly similar and depicted a typical beach scene but differed according to the presence or absence of ten different elements. For example, while Picture A included a sandcastle, Picture B did not. In terms of task complexity, the pictures were manipulated by the resource-directing +/- few elements dimension, meaning the complex version contained more characters and objects than the simple version did.

The picture difference tasks used in Kim (2009a) were made partially available to me by You Jin Kim as the full versions were unavailable. That is to say, I was provided with a single picture (Picture A) for both the simple and complex versions of the task. As a result, a research assistant and I developed a second set of pictures (Picture B) for both the simple and complex tasks that closely resembled the existing pictures provided to me. As in Kim (2009a), the second set of pictures were designed to include ten differences from the first set to promote interaction as pairs of participants identified the differences between their pictures during both the simple and complex tasks.

Picture narration tasks. In the original study (Kim, 2009a) several pictures that made up a single narrative story were divided between participants for both simple and complex versions of the picture narration tasks. In the simple version, each picture was numbered, and thus participants had knowledge of how the narrative should be sequenced and were instructed to tell the story accordingly. In the complex version, task complexity was manipulated by the + reasoning demands dimension, meaning the pictures did not include numbers and participants had to decide on an order for the story in addition to narrating it.

As the original pictures for the narration task used in Kim’s study were unavailable, the present study used two picture narration tasks - one simple and one complex - that were closely adapted from Gilabert et al. (2009) (see Appendix A), who also investigated the effects of task complexity on learner-learner interaction. As indicated in Chapter 1, Gilabert et al. found that
their complex picture narration task had a significant effect on the occurrence of LREs with self-corrections included as LREs compared to a simple picture narration task. This is consistent with Kim’s findings for the high proficiency learners but not the low proficiency learners, who produced more LREs during a simple picture narration task. In addition, Gilabert et al. analyzed self-corrections separately, which also occurred at higher rates during the complex picture narration task. Thus, these tasks seemed particularly appropriate for use in this replication study. In both Gilabert et al. and the present study, the pictures were divided equally between each pair of participants and they were instructed to create a storyline collaboratively through oral interaction with one another while also being asked to order the story in the more complex version.

It is important to note that Gilabert et al. also manipulated the complexity condition in which these picture narration tasks were completed along +/- here-and-now/there-and-then dimension. As indicated in Chapter 1, a narration task which includes the + here-and-now dimension refers to a simple task complexity condition in which learners are asked to look at a set of pictures and then prompted by the researcher to narrate them in the present tense while looking at them. The opposite and more complex task condition, + there-and-then, refers to a situation in which learners are given a short period of time to view the set of pictures, and then asked to narrate them in the past tense without the visual aid of the pictures in front of them. Thus, in Gilabert et al. the pictures were left in front of learners throughout the duration of the simple narration task and they were asked to narrate a story in the present tense. Alternatively, during the complex narrative task, learners were given 50 seconds to study the pictures, at which point they were removed from their sight and they were asked to narrate a story in the past tense. Fifty seconds has been operationalized as a period of time in several studies as an unplanned task condition (Foster & Skehan, 1996; Gilabert, 2007; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997). This is important because the inclusion of planning time would involve manipulating the task by an additional resource-dispersing complexity dimension, +/- planning time, which could also mediate the role the + there-and-then dimension may play on participant interactions. In the present study, however, the picture narration task adapted from Gilabert et al. was not manipulated by the +/- here-and-now/there-and-then dimension due to the results of the pilot study, which are described below along with other revisions made to these tasks.
Pilot Studies

In order to ensure that the oral tasks elicited LREs during learner-learner interaction at both low and high proficiency levels as well as possessed face value for English language learners, two pilot studies were conducted with four low proficiency learners and four high proficiency learners of English.

In the first pilot study, two low and two high proficiency learners were organized into two pairs and completed the bio data questionnaire described above. Afterwards they viewed two brief pre-task modeling videos that depicted the primary researcher and an assistant completing a picture narration and picture difference task in order to provide instructions on how to complete the tasks as well as to demonstrate collaborative patterns of interaction during task performance. These videos were designed and modeled on research discussed in Chapter 1 indicating that pre-task modeling may lead learners to engage in more collaborative patterns of interaction (Kim, 2013; Kim & McDonough, 2011) while also encouraging learners to engage with linguistic forms through the production of LREs. The tasks performed in the videos used in the present study, although of the same type, were different from the ones learners completed themselves. As discussed in Chapter 1, it is important to note that pre-task modeling was not included in Kim’s (2009a) original study, and thus its inclusion represents a modification within this conceptual replication study. More detailed information regarding these videos is provided in the main study section below.

Finally, each pair was asked to complete the four tasks described above: a simple and complex version of a picture difference task and a picture narration task. They were given fifteen minutes to complete each task and were audio recorded. Once learners had completed all tasks the pilot data collection session ended.

Results of pilot study 1. The results of the first pilot study demonstrated that learners at both levels of proficiency had no difficulty in completing simple and complex versions of the picture difference tasks, and both tasks elicited LREs successfully. The simple picture narration task was similarly successful. However, few LREs were produced during the complex picture narration task, particularly among the low proficiency learners. This appeared to be due to the difficulty level of the task. In the first pilot study, I had initially attempted to manipulate the complex picture narration task by the + there-and-then dimension as Gilabert et al. (2009) had done by allowing learners to view the pictures for 50 seconds, at which point the pictures were
removed from their sight and they were asked to order the pictures and narrate a story together. However, the low proficiency learners were unable to proceed and found this overwhelmingly difficult. The high proficiency learners also experienced frustration and found it difficult to recall the pictures. Due to these difficulties, I decided to restart the task for both proficiency pairs of learners, and I allowed them to view the pictures while ordering and narrating them throughout the task. Thus, the complex picture narration task was manipulated solely by the + reasoning demands dimension, as opposed to both the + reasoning demand and the + there-and-then dimension as in Gilabert et al. In the end this revision worked well because in Kim’s study the complex picture narration tasks were manipulated solely by the + reasoning demands dimension, not the + there-and-then dimension.

Despite these changes to the condition in which learners completed the task few LREs were produced during the interaction. Feedback following the first pilot study from the four learners made it apparent that they found a number of the pictures included in the complex picture narration task highly similar, and thus difficult to differentiate and order. All four participants reported that it was difficult to negotiate language with their partners as they felt they were often describing the same pictures repeatedly. Given that in previous research more complex picture narration tasks have been shown to elicit higher numbers of LREs than simple picture narration tasks (Gilabert et al., 2009), particularly among higher proficiency learners (Kim, 2009a), additional revisions were made to the complex narrative task in order to make it more feasible to differentiate between the pictures and elicit more successful interaction among pairs.

Results of pilot study 2. In the second pilot study, one low and one high proficiency pair participated in the same process with the revised complex picture narration task. Learners found this revised task more feasible to complete, and it also elicited LREs more consistently. Overall, all tasks elicited LREs successfully and possessed excellent face value for learners. Based on these results, I decided to proceed with the main study. Table 4 summarizes the four tasks used in the present study by complexity and task type, which are identical to those used in Kim’s original study (2009a).
Table 4 ComplexityDimensions Used in the Replication Study and the Original Study (Kim, 2009a)

<table>
<thead>
<tr>
<th>Task type</th>
<th>Picture narration 1</th>
<th>Picture narration 2</th>
<th>Picture difference 1</th>
<th>Picture difference 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity Dimension</td>
<td>- reasoning demands</td>
<td>+ reasoning demands</td>
<td>+ few elements</td>
<td>- few elements</td>
</tr>
<tr>
<td>Complexity Level</td>
<td>Simple</td>
<td>Complex</td>
<td>Simple</td>
<td>Complex</td>
</tr>
</tbody>
</table>

The Main Study

Participants and proficiency. After obtaining formal consent from the program administration and asking for volunteers to participate in the study, 34 male and female ESL students gave their consent. However, due to a recording malfunction, data were not successfully collected for two participants. This resulted in a total of 32 students (18 females and 14 males) from a variety of countries and L1 backgrounds (Japanese, Portuguese, Chinese, Spanish, and Arabic) participating in the study. They ranged from 18-58 years of age with an average age of 25 ($SD = 8.95$).

Since proficiency level was a variable under investigation, 16 low proficiency learners and 18 high proficiency learners were recruited. Proficiency in English was determined by the enrolment status of participants in the school’s program. Low proficiency learners were solicited from beginner or lower intermediate levels of the program, whereas high proficiency learners all came from classes at high intermediate or advanced levels of English.

Data collection procedures. The independent variable in this study was task complexity, which was manipulated in each of the tasks as described above. The dependent variable was the occurrence of learning opportunities during task-based interaction, identified as LREs. The data collection procedures for both the low and high proficiency learners were the same and conducted in two separate sessions within the same week and are described below. Eight pairs of low proficiency learners ($n=16$) and nine pairs of high proficiency learners ($n=18$) participated in the main study.

Due to the possibility of error when administering the tasks with participants, I and a research assistant administered the tasks. This included standardized instructions in completing the tasks. The research assistant was trained prior to data collection and helped administer the tasks, monitor students, as well as provide brief instructions as participants moved from task to task. These instructions were read from a handout distributed to learners to ensure that all
participants received the same instructions. The research assistant was a university level English language instructor and possessed both a Master’s of Education degree in Second Language Education as well as a certificate in Teaching English to Speakers of Other Languages (TESOL).

Learners were paired using simple random sampling procedures (Mackey & Gass, 2005). A list of the participants was created prior to data collection and upon arrival to the data collection site, learners were instructed to draw a piece of paper from an opaque bag that included pairs of letters for each proficiency group: from A-G for the low proficiency or A-H for the high proficiency group (8 pairs of letters for a total of 16 participants in the low group and 9 pairs of letters for a total of 18 participants in the high group). Once all learners drew one letter they were instructed to find their partner with the corresponding letter.

When learners were seated with their partners they were formally welcomed and provided with a brief overview of the sequence and plan for data collection. First, students were provided with consent forms and given detailed information about them in as simple language as possible. Each student was asked if they understood and agreed to the consent form, at which point they were given an additional 10 minutes to read and sign them. All participants had been provided with information letters one week before the day of data collection that included the same information on the consent forms. Shortly afterwards, students were administered the biodata questionnaire, which they were given 10 minutes to complete.

The learners were then told that they would be watching two videos that would demonstrate a collaborative pattern of pair work. As indicated above, these brief pre-task modeling videos demonstrated how pairs could work collaboratively and help one another with language difficulties and misunderstanding while completing the tasks. They also served as an example of how the tasks could be conducted. These videos were created and included in my study following personal communication with the author of the original study being replicated (Y. Kim, personal communication, January 22, 2014) whose subsequent research indicates that task-modeling may lead to more collaborative patterns of pair work, and to the occurrence of higher numbers of LREs (Kim, 2013; Kim & McDonough, 2011). The videos were created following the same procedures described in detail in Kim & McDonough (2011) and Kim (2013). Both videos included myself and an assistant participating in the two different types of tasks (picture difference and picture narration) for 2-3 minutes each, and they included moments of misunderstanding as well as self- and other-correction. It is important to recall that the
incorporation of pre-task modeling represents a departure from the methodology of the original study.

Once learners had viewed the videos they were asked to read the instructions in the form of a handout placed at each pair’s table. Once they had confirmed that they had read and understood the instructions the audio recorders distributed to each pair were turned on and learners were instructed to begin the first task. Learners completed all four tasks with the same partner in succession in approximately 70 minutes. They were given 15 minutes to complete each task and were instructed to perform them in succession for a total of 60 minutes. The remaining 10 minutes were used between the four tasks to give brief instructions to pairs as they began their next task.

In order to avoid an ordering effect, a Latin square design was employed in this study (Gilabert, 2005; Gilabert et al., 2009). Thus, each pair of learners performed the four activities in a different order. A sample of this design is displayed below in Table 5.

<table>
<thead>
<tr>
<th>Pair 1</th>
<th>Pair 2</th>
<th>Pair 3</th>
<th>Pair 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>Task 2</td>
<td>Task 3</td>
<td>Task 4</td>
</tr>
<tr>
<td>Task 3</td>
<td>Task 4</td>
<td>Task 2</td>
<td>Task 1</td>
</tr>
<tr>
<td>Task 2</td>
<td>Task 1</td>
<td>Task 4</td>
<td>Task 3</td>
</tr>
<tr>
<td>Task 4</td>
<td>Task 3</td>
<td>Task 1</td>
<td>Task 2</td>
</tr>
</tbody>
</table>

Note. Task 1 – Simple Picture Difference
Task 3 – Simple Picture Narration
Task 2 – Complex Picture Difference
Task 4 – Complex Picture Narration

**Data analysis procedures.** I and a research assistant transcribed audio recordings of the learner interactions. The research assistant was provided with instructions as well as a document that I created that gave details on appropriate transcribing conventions for the purposes of the present study (see Appendix B).

The transcriptions were then coded for the occurrence of LREs, defined by Swain and Lapkin (1998) as “any part of a dialogue in which students talk about the language that they are producing, question their language use, or correct themselves or others” (p. 70). As indicated in Chapter 1, LREs were chosen because they are considered to demonstrate potential for language learning in progress by providing learners with the opportunity to make hypotheses about L2 forms and to restructure their existing L2 knowledge.
As in Kim’s (2009a) study, LREs were coded as either grammatical or lexical in order to determine whether learners at different levels of English proficiency performing different task types pay attention to different aspects of linguistic form. Following Swain and Lapkin (1998), Kim (2009a) defines grammatical LREs as “LREs in which learners discussed grammatical features of English, including verb tenses, word order, preposition use and so on” (p. 258). Though Kim does not specify other grammatical features aside from those listed in the above quotation, according to Swain and Lapkin (1998; 2002) grammatical LREs also involve attention to articles and pronouns, and were thus included in the definition for the present study. The following example (1) illustrates a typical grammatical LRE within the transcript.

1) Grammatical LRE

1. Learner A: And the dog run.
2. Learner B: Dog is running in the picture
3. Learner A: Is running yes is running

Lexical LREs refer to instances where “learners discuss the meaning, pronunciation or spelling of lexical items” (Kim, 2009a, p. 258). While lexical items are not defined more explicitly in Kim (2009a), Swain and Lapkin (2002) specify that they focus on adverbs, adjectives, nouns and verbs. Although LREs involving the spelling of a word originally fall under the category of form (grammatical) in Swain and Lapkin (1998; 2002), they were categorized by Kim (2009a) as lexical. In an effort to follow the same methods as the study being replicated, spelling was also coded as lexical in the present study. A typical lexical LRE is illustrated below (2).

2) Lexical LRE

1. Learner A: There is a dog (pronounced as “dug”)
2. Learner B: Dog? (“dug”)
3. Learner A: Dog (“dug”) an animal dog (“dug”), bow wow wow.
4. Learner B: Oh a dog (correctly pronounced as “dog”!)

In this example, Learner A mispronounces the word ‘dog’, which causes Learner B to question his or her language use. This results in Learner A attempting to define the word, followed by Learner B’s comprehension, which finally results in the word being uttered with correct pronunciation.
In addition to these two types of LREs and in contrast with Kim’s study, LREs were further coded as lexical or grammatical self-corrections. Self-corrections were contrasted with interactive LREs. Interactive LREs involve both learners in a pair, whereas self-corrections only involve one learner. Thus, LREs were categorized in four different ways: as Lexical Interactive LREs (INT LEX), Grammatical Interactive LREs (INT GRAM), Lexical Self-Corrections (SELF LEX), and Grammatical Self-Corrections (SELF GRAM). The decision to further subdivide LREs to differentiate between interactive LREs and self-corrections was based on research that has found self-corrections to occur at a much higher rate than interactive LREs (Gilabert et al., 2009). As discussed in Chapter 1, some researchers have argued that self-corrections are less collaborative than those made in interaction with another learner (Fortune & Thorp, 2001). Thus, it was decided to distinguish between the occurrences of self-corrections and interactive LREs in this study.

The following examples (3) (4) illustrate typical self-corrections identified in the transcript with self-corrections bolded.

3) Grammatical Self-Correction Examples
1. Learner A: There are castle. There is, there is, castle (subject-verb agreement self-correction)
2. Learner B: There are tent in the, in the, in, on the beach. (preposition self-correction)

4) Lexical Self-Correction Examples
1. Learner A: How much bee, bird do you have? Four yes? (lexical self-correction)
2. Learner B: Your colour (pronounced as “cooler”)? Colour (pronounced correctly as “colour”)? Colour? (pronunciation self-correction)

Kim also analyzed the resolution of LREs based on Swain (1998) and Leeeser’s (2004) categorization, which includes correctly resolved (Type 1), unresolved (Type 2), and incorrectly resolved LREs (Type 3). No major differences were found across tasks and proficiency levels, so I decided not to pursue an analysis of LRE resolution in this study.

---

4 Presumably, this was motivated by Robinson’s argument (2007), which contends that more complex tasks lead to a greater amount of uptake than less complex tasks. In Kim’s study, uptake appears to refer to LREs that are correctly resolved. However, Kim found that learners correctly resolved similar amounts of LREs in both simple and
Development of the coding scheme. The development of the coding scheme was to ensure that coding conventions were applied consistently both by me and the assistant researcher, who coded a randomly selected sample of the transcriptions for inter-coder reliability. The first step in developing this coding scheme was to include all the available coding information in the original study (Kim, 2009a). Also included was Swain and Lapkin’s (1998) definition of LREs, upon which Kim’s analysis was based. This definition was extended and clarified in Swain and Lapkin (2002) and was used to further clarify the distinction between lexical and grammatical LREs. I also conferred with a PhD student with several years of experience coding LREs in various research projects supervised by Swain and Lapkin. These discussions led to further refinement of the coding scheme used for the present study. The coding scheme was first used on the transcriptions from the pilot study in practice sessions in order to determine if and where any problems or ambiguities remained. Resolving these led to continued refinement of the coding scheme in order to minimize the need for interpretation and inference when coding. A detailed set of rules were compiled and organized in the form of a document that was used to code the transcriptions for the main study. Next, I coded eight (half) of the transcriptions and then subsequently recoded four of them to measure intra-coder reliability. This led to further revisions and clarifications. Following this, the document was finalized and provided to the research assistant who was trained to assist with the inter-coder reliability check (see Appendix C).

Analysis of LREs by task. All LREs were identified in the 16 transcripts used for the main study and categorized in the manner discussed above (i.e., INT LEX, INT GRAM, SELF LEX, SELF GRAM). Due to differences in transcription length, particularly between the simple and complex narrative tasks, the average number of LREs produced per 100 words was calculated for each activity. Though it is unclear whether or not Kim (2009a) corrected for transcription length, research suggests that this may have an impact on results obtained from transcriptions that are based on oral dialogue (Gilabert et al., 2009). In order to render simple and complex versions of both the picture narration and picture difference task with the exception of the high learners’ performance during the picture narration task. In this task, learners correctly resolved 10% more LREs in the complex task as opposed to the simple task.
complex tasks more comparable, a calculation based on Gilabert (2007) was used to calculate the ratio of LREs in simple and complex tasks. The ratio was calculated as follows:

\[
\frac{\text{Number of LREs}}{\text{Total Number of Words}} \times 100 = \text{Average Number of LREs per 100 words}
\]

**Interrater reliability.** To ensure reliability, a randomly selected subset of the data (25%) was coded by the same volunteer researcher assistant who had assisted with the data collection. The volunteer coder was given detailed instructions and training to help ensure strong inter-coder reliability. A percentage agreement was calculated by comparing the identification of LREs in my coding and the research assistant’s coding. Results indicated that there was 92.86% agreement for the occurrence of LREs.
Chapter 3: Results

In this chapter, the results of the statistical analyses are presented. Non-parametric tests were used because that the sample sizes for both the low proficiency group \((n=16)\) and high proficiency group \((n=16)\) were small and the data were not distributed normally.\(^5\) In addition, the majority of the data was skewed either positively or negatively. Thus, these results should be interpreted with caution. This chapter is divided into two main sections: First, the results that are relevant to the original research question investigated by Kim are presented in order to compare the findings of the present conceptual replication study and the original study. This is followed by the results in relation to the three additional research questions investigated in the present study, which seek to extend Kim’s inquiry.

Results Related to Original Research Question

The first research question investigated the effects of task complexity and task type on the occurrence of LREs during task-based interaction between ESL learners from different proficiency levels. Wilcoxon Signed Ranks tests were used to analyze the differences between the number of LREs the low and high proficiency learners produced during the simple and complex versions of the picture narration and picture difference tasks. The alpha level was set at .05. First, the results of the low and high proficiency learners’ performance on the picture narration tasks are presented, followed by their performance on the picture difference tasks.

**Picture narration tasks.** Following Kim (2009a), I hypothesized that the low proficiency learners would produce more LREs during the simple picture narration task than in the complex version, while an opposite result would be found for the high proficiency learners, who would produce more LREs during the complex picture narration task than in the simple version. Table 6 provides information on the average number of LREs produced per 100 words by the low and high proficiency learners during the picture narration tasks.

\(^5\) Non-parametric Wilcoxon Signed Ranks tests were also used to analyze the data in Kim (2009a). Although Kim (2009a) presented means and standard deviations in the descriptive statistics, a statistical consultant suggested the median as a more appropriate measure of central tendency for the non-parametric data in the present study. For the same reason, the interquartile range (IQR) was calculated as a measure of variability. The median is also presented in the present study as the majority of the data was not distributed normally.
Table 6 Average Number of LREs Produced per 100 Words by Proficiency Groups during Picture Narration Tasks

<table>
<thead>
<tr>
<th>Proficiency Complexity</th>
<th>Lexical LREs</th>
<th>Grammatical LREs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Mdn</td>
<td>IQR</td>
</tr>
<tr>
<td>Low Simple</td>
<td>9.92</td>
<td>1.40</td>
<td>.93</td>
</tr>
<tr>
<td>Complex</td>
<td>5.39</td>
<td>.62</td>
<td>.62</td>
</tr>
<tr>
<td>High Simple</td>
<td>8.28</td>
<td>1.05</td>
<td>.26</td>
</tr>
<tr>
<td>Complex</td>
<td>7.12</td>
<td>.82</td>
<td>.46</td>
</tr>
</tbody>
</table>

For the low proficiency learners, results indicated a significant difference ($z = 1.96, p < .05$), with learners producing more LREs during the simple picture narration task than the complex picture narration task. The mean of the ranks for LREs generated during the simple picture narration task was 4.57 ($Mdn = 2.32$), while the mean of the ranks for the complex version of the picture narration task was 4.00 ($Mdn = 1.38$). Results also indicated a large effect size ($r = .69$) (Cohen, 1992). Thus, for the low proficiency learners, my hypothesis was confirmed. However, for the high proficiency learners, no significant differences were found between LREs produced during the simple and complex picture narration tasks ($p > .05, r = .49$). Thus, my hypothesis for the high proficiency learners’ performance on the picture narration tasks was not confirmed.

**Picture difference tasks.** In line with Kim (2009a), I predicted that the low proficiency learners would produce significantly more LREs in the complex picture difference task than in the simple picture difference task. As for the high proficiency learners, based on Kim’s results, I predicted that no significant differences would be found between the number of LREs produced during the simple and complex picture difference tasks. Table 7 provides information on the average number of LREs produced per 100 words by the low and high proficiency learners during the picture difference tasks.

---

6 Effect size was calculated according to the formula for Wilcoxon Signed Ranks test: $r = z / \sqrt{N}$. 
Table 7 Average number of LREs Produced per 100 Words by Proficiency Groups during Picture Difference Tasks

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Complexity</th>
<th>Lexical LREs</th>
<th>Grammatical LREs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Mdn</td>
<td>IQR</td>
<td>Number</td>
</tr>
<tr>
<td>Low</td>
<td>Simple</td>
<td>9.01</td>
<td>1.13</td>
<td>.58</td>
</tr>
<tr>
<td>Complex</td>
<td>10.23</td>
<td>1.34</td>
<td>.46</td>
<td>6.02</td>
</tr>
<tr>
<td>High</td>
<td>Simple</td>
<td>10.07</td>
<td>1.29</td>
<td>.50</td>
</tr>
<tr>
<td>Complex</td>
<td>8.25</td>
<td>1.08</td>
<td>.20</td>
<td>7.88</td>
</tr>
</tbody>
</table>

The results of the Wilcoxon Signed Ranks test indicated no significant differences between the number of LREs produced in the simple and complex picture difference tasks for the low proficiency learners ($p > .05$, $r = .52$). Thus, my hypothesis for the low proficiency group was not confirmed. As for the high proficiency learners, no significant differences were found between the number of LREs produced in the simple and complex picture difference tasks ($p > .05$, $r = .10$). Thus, my hypothesis for the high proficiency group was confirmed. Tables 8 and 9 summarize the results of the analyses by proficiency, complexity, and task type.

Table 8 Summary of Wilcoxon Signed Ranks Test Results for All Tasks

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Picture Narration</th>
<th>Picture Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Z$</td>
<td>$p$-value</td>
</tr>
<tr>
<td>Low</td>
<td>-1.96</td>
<td>.05$^*$</td>
</tr>
<tr>
<td>High</td>
<td>-1.40</td>
<td>.16</td>
</tr>
</tbody>
</table>

$^*$Significant difference

Table 9 Summary of Results by Proficiency, Complexity, and Task Type

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Picture Narration</th>
<th>Picture Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Significantly more LREs during simple task</td>
<td>No differences between LREs during simple and complex tasks</td>
</tr>
<tr>
<td>High</td>
<td>No differences between LREs during simple and complex tasks</td>
<td>No differences between LREs during simple and complex tasks</td>
</tr>
</tbody>
</table>
Results for Additional Research Questions

As described in Chapter 1, the present study investigated additional research questions in an effort to extend Kim’s descriptive analysis of lexical and grammatical LREs as well as to investigate the effect of task complexity and task type on LREs when they are divided into interactive LREs (i.e., LREs involving both learners in a pair) and self-corrections (LREs involving one individual in a pair). The results in relation to the three research questions presented in Chapter 1 are presented as follows: (a) the occurrence of grammatical and lexical LREs, (b) the separate occurrence of interactive LREs and self-corrections, and (c) the proportion of interactive LREs and self-corrections according to lexical or grammatical features of English.

Lexical vs. grammatical LREs. Based on Kim’s descriptive findings, I predicted that both low and high proficiency learners would produce more lexical LREs during the simple and complex picture difference tasks. In addition, I hypothesized that the low proficiency learners would produce more lexical LREs during the simple picture narration task. Kim’s findings were based on a descriptive observation of the overall raw number of lexical and grammatical LREs produced by learners in each proficiency group during simple and complex tasks, whereas the present study compared the average number of LREs produced per 100 words during simple and complex tasks for each proficiency group using a Wilcoxon Signed Ranks test.

For the picture difference tasks, the low proficiency learners produced significantly more lexical LREs during both the simple \( (z = 2.52, p < .05, r = .89) \) and complex \( (z = 2.31, p < .05, r = .82) \) picture difference tasks. The high proficiency learners, on the other hand, produced significantly more lexical LREs during the simple picture difference task \( (z = 2.52, p < .05, r = .89) \), but not during the complex version \( (p > .05, r = 0.10) \). For the picture narration tasks, there were no significant differences between the number of lexical or grammatical LREs produced by both the low proficiency learners during the simple \( (p > .05, r = .10) \) and complex version \( (p > .05, r = .45) \). A similar result was obtained for the high proficiency learners on both the simple \( (p > .05, r = .35) \) and complex \( (p > .05, r = .10) \) picture narration tasks.

In summary, the low proficiency learners produced more lexical LREs during both simple and complex picture difference tasks, while the high proficiency learners only produced more lexical LREs during the simple picture difference task. In addition, no differences were found in either proficiency group during the picture narration tasks.
Separate occurrence of interactive LREs and self-corrections. The number of interactive LREs and self-corrections produced by learners during the simple and complex tasks were also analyzed separately. These analyses were conducted in order to determine whether learners produced significantly more or fewer LREs during the tasks depending on how the LREs were defined – i.e. when only LREs that were genuinely interactive or LREs as self-corrections were taken into account. The relative occurrence of interactive LREs and self-corrections are displayed in Table 10 for the picture narration tasks and Table 11 for the picture difference tasks.

**Table 10 Relative Occurrence of Interactive LREs and Self-corrections for Picture Narration Tasks**

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Complexity</th>
<th>Interactive LREs</th>
<th>Self-Corrections</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number</td>
<td>Mdn</td>
<td>IQR</td>
</tr>
<tr>
<td>Low</td>
<td>Simple</td>
<td>6.58</td>
<td>.76</td>
<td>.51</td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>2.66</td>
<td>.26</td>
<td>.62</td>
</tr>
<tr>
<td>High</td>
<td>Simple</td>
<td>5.04</td>
<td>.59</td>
<td>.57</td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>4.27</td>
<td>.52</td>
<td>.16</td>
</tr>
</tbody>
</table>

**Table 11 Relative Occurrence of Interactive LREs and Self-corrections for Picture Difference Tasks**

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Complexity</th>
<th>Interactive LREs</th>
<th>Self-Corrections</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number</td>
<td>Mdn</td>
<td>IQR</td>
</tr>
<tr>
<td>Low</td>
<td>Simple</td>
<td>6.67</td>
<td>.74</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>7.81</td>
<td>1.05</td>
<td>.47</td>
</tr>
<tr>
<td>High</td>
<td>Simple</td>
<td>6.86</td>
<td>.95</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>6.78</td>
<td>.86</td>
<td>.31</td>
</tr>
</tbody>
</table>

**Picture narration tasks.** The results of the low proficiency learners’ performance during the picture narration tasks indicated that when interactive LREs were analyzed separately from self-corrections significantly more interactive LREs occurred during the simple picture narration task than during the complex version ($z = 2.52, p < .05, r = $). The mean of the ranks for
interactive LREs generated during the simple picture narration task was 4.52 (Mdn = .76) while the mean of the ranks for the complex version of the picture narration task was 0.00 (Mdn = .26). Results also indicated a large effect size (r = 0.89). Thus, the low proficiency learners produced significantly more interactive LREs during the simple picture narration task than in the complex task. However, no significant differences were found between the number of self-corrections produced by the low proficiency learners in the simple and complex narration tasks (p > .05, r = .42). This was also the case for the high proficiency learners, with no significant differences found between the simple and complex picture narration tasks both in terms of interactive LREs (p > .05, r = .25) and self-corrections (p > .05, r = .25).

**Picture difference tasks.** As for the low proficiency learners’ performance on the picture difference tasks, no significant differences were found between the simple and complex tasks both in terms of interactive LREs (p > .05, r = .35) and self-corrections (p > .05, r = .40). A similar result was obtained for the performance of the high proficiency learners, with no significant differences found between the simple and complex picture difference tasks both in terms of the number of interactive LREs (p > .05, r = .50) and self-corrections produced (p > .05, r = .22).

Overall, the results of the separate analysis of interactive LREs and self-corrections yielded similar results obtained in the analysis of the first research question. In this initial analysis, LREs were analyzed as a whole with interactive LREs and self-corrections combined in response to the Kim’s (2009a) original research question. As discussed above, the low proficiency learners produced significantly more LREs during the simple picture narration task than the complex version, whereas no other significant differences were found between any of the other tasks for the two proficiency groups. When interactive LREs and self-corrections were analyzed separately in order to address the additional research question presented in this section, the same results were obtained as in the original analysis with the exception of the number of self-corrections produced by the low proficiency learners during the simple and complex picture narration tasks. During these tasks, the low proficiency learners produced significantly more interactive LREs during the simple picture narration task than the complex picture narration task but they did not produce significantly more self-corrections. No other significant differences were found between the number of LREs produced during the simple and complex picture difference tasks for the low proficiency learners when interactive LREs and self-corrections were
analyzed separately. This was also the case for the high proficiency learners’ performance during the simple and complex versions of both the picture difference and picture narration tasks. Table 12 summarizes the results of the analyses by proficiency, complexity, and task type.

**Table 12** Summary of Wilcoxon Signed Ranks Test Results for all Tasks Analyzed by the Separate Occurrence of Interactive LREs and Self-corrections

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Picture Narration</th>
<th>Picture Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Z</td>
<td>p-value</td>
</tr>
<tr>
<td>Low Interactive LREs</td>
<td>-2.52</td>
<td>.01*</td>
</tr>
<tr>
<td>Low Self-Corrections</td>
<td>-1.19</td>
<td>.23</td>
</tr>
<tr>
<td>High Interactive LREs</td>
<td>-.70</td>
<td>.48</td>
</tr>
<tr>
<td>High Self-Corrections</td>
<td>-.70</td>
<td>.48</td>
</tr>
</tbody>
</table>

*Significant difference

**Proportion of interactive LREs and self-corrections according to lexical or grammatical features of English.** I also analyzed separately the number of interactive LREs and self-corrections produced by learners during the simple and complex tasks according to whether they related to lexical or grammatical features of English. This analysis was conducted in order to determine whether particular LREs, for example, self-corrections, tended to be lexical or grammatical. The relative occurrence of interactive LREs and self-corrections according to lexical or grammatical features is provided in Table D1 for the picture narration tasks and Table D2 for the picture difference tasks. To provide more direct comparisons, the relative occurrence of interactive LREs and self-corrections is provided in the form of percentages in Table 13 organized as either lexical or grammatical.
Table 13 Proportions of the Occurrence of LREs Organized by Type of Interaction and Language Focus for All Tasks

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Language Focus</th>
<th>Type of Interaction</th>
<th>Picture Narration</th>
<th>Picture Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Simple</td>
<td>Complex</td>
</tr>
<tr>
<td>Low</td>
<td>Lexical</td>
<td>Interactive LREs</td>
<td>28</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>21</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammatical Interactive LREs</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>46</td>
<td>53</td>
</tr>
<tr>
<td>High</td>
<td>Lexical</td>
<td>Interactive LREs</td>
<td>22</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>21</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammatical Interactive LREs</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>52</td>
<td>48</td>
</tr>
</tbody>
</table>

**Picture narration tasks.** Results indicated that for the simple picture narration tasks, the low proficiency learners produced a similar number of lexical interactive LREs (28%) and lexical self-corrections (21%). This was also the case for the complex picture narration tasks for lexical interactive LREs (17%) and lexical self-corrections (26%). In contrast, the overwhelming majority of grammatical LREs produced by the low proficiency learners were the result of self-correction, not collaboration. For example, in the simple picture narration task, only 5% of the LREs produced by the low proficiency learners were the result of self-correction, while 46% were the result of self-correction. A similar trend was observed in the complex picture narration task for interactive LREs (4%) and self-corrections (53%).

The high proficiency learners produced very similar proportions of the four types of LREs as the low proficiency learners in both simple and complex picture narration tasks. During the simple picture narration task, 22% of the LREs were the result of collaboration and 21%
were the result of self-correction. Furthermore, the same pattern found for the low proficiency learners was observed for the high proficiency learners during the complex picture narration task for collaborative LREs (17%) and self-corrections (26%). As for grammatical LREs, self-corrections again occurred at a much higher rate, with only 5% of the LREs produced in the simple task being the result of collaboration and 52% the result of self-correction. A similar pattern was found during the complex picture narration task for interactive LREs (9%) and self-corrections (48%). Thus, it appears that learners at both proficiency levels tended to produce a similar number of lexical LREs that were either collaborative or individual in nature, while grammatical LREs tended to be produced individually as self-corrections.

**Picture difference tasks.** As for the simple picture difference task, the low proficiency learners produced more than twice as many lexical interactive LREs (45%) than lexical self-corrections (21%). For the complex picture difference task, a similar result was obtained for lexical interactive LREs (43%) and lexical self-corrections (20%). As in the picture narration tasks, grammatical self-corrections also occurred at a much higher rate than interactive LREs in the picture difference tasks. In the simple picture difference task, only 4% of the LREs were the result of collaboration while 30% were the result of self-correction. A similar result was found during the complex picture difference task for interactive LREs (5%) and self-corrections (32%).

The high proficiency learners also produced more lexical interactive LREs (39%) than lexical self-corrections (23%) during the simple picture difference task, and more than twice as many during the complex picture difference task, with 36% of the LREs the result of collaboration and 15% the result of self-correction. During the simple picture difference task, grammatical self-corrections (35%) once again occurred at much higher rates than grammatical interactive LREs (3%), with a similar result obtained for grammatical self-corrections (43%) and grammatical interactive LREs (6%) in the complex picture difference task. Thus, during the simple and complex picture difference tasks, learners of both proficiency levels tended to produce more lexical collaborative LREs than lexical self-corrections, whereas they produced many more grammatical self-corrections than grammatical interactive LREs.

In summary, learners at both proficiency levels produced similar numbers of lexical LREs that were interactive and individual in nature in the picture narration tasks. However, they produced more lexical LREs that were interactive during the simple and complex picture difference tasks. As for grammatical LREs, learners of both proficiency levels produced many
more grammatical self-corrections than grammatical interactive LREs across all tasks regardless of task complexity or task type. Figures 1 and 2 illustrate the relative occurrence of the four types of LREs for the low and high proficiency learners respectively.

**Figure 1** Occurrences of the Four types of LREs by Task and Complexity among the Low Proficiency Learners by Percentage

**Figure 2** Occurrences of the Four Types of LREs by Task and Complexity among the High Proficiency Learners by Percentage
Chapter 4: Discussion

This chapter is organized into two main sections. First, I discuss the results in relation to the original research question investigated in both the original study and the present replication study. Second, I discuss the results in relation to the three additional research questions investigated in the present study.

Original Research Question

The first research question that was investigated in both Kim (2009a) and the present replication study is: (1) Do task complexity and task type impact the occurrence of LREs during task-based interaction between ESL learners from different proficiency levels? The results are discussed below in relation to the two different task types used in the two studies.

Picture narration tasks. Based on Kim’s (2009a) results, I hypothesized that the low proficiency learners would produce significantly more LREs during the simple picture narration task. This hypothesis was confirmed lending support to Kim’s findings and contradicting the Cognition Hypothesis. As discussed in Chapter 1, Nuevo (2006) also found the simple version of a picture narration task to elicit more interaction moves (e.g. recasts, uptake of recasts) included within Swain and Lapkin’s (1998) definition of LREs. Nuevo speculated that learners engaged in low complexity tasks may have more attentional resources available to monitor their interlocutors’ output, making it more feasible to provide feedback to each other. Kim (2009a) argues that this might be particularly relevant for low proficiency learners whose attentional resources are likely to be focused on completing the complex task at the expense of producing interaction-driven learning opportunities. However, in Nuevo’s study, another interaction move (i.e., metalinguistic talk) did not occur more frequently in the simple versions of the tasks.

Overall, it is difficult to make direct comparisons between Nuevo’s study and Kim’s study (and by extension, Nuevo’s study with the present replication study) because LREs, the interaction measure used in the present study and Kim (2009a), are a composite of the multiple measures used in Nuevo (2006).

As for the high proficiency learners, Kim found that the complex picture narration task did result in significantly more learning opportunities, which supports Robinson’s Cognition Hypothesis. She concludes that high proficiency learners, with relatively freer attentional resources available, were pushed to interact more during the complex picture narration task manipulated by the + reasoning demands dimension. She argues that these learners produced
LREs particularly in the form of negative feedback such as recasts and explicit correction (2009a), which is supported by Robinson and Gilabert’s (2007) claim that more complex tasks lead to a greater use of focus on form techniques such as recasts. However, my second hypothesis, which was based on Kim’s results, was not confirmed. That is to say, no differences were found between the number of LREs produced by the high proficiency learners during completion of the simple and complex versions of the picture narration tasks.

There are a number of possible reasons for the different results obtained in the two studies. First, as discussed earlier, the present study altered Kim’s research design by having learners observe brief pre-task modeling videos in an effort to encourage them to produce LREs and participate in a collaborative manner during task completion. These videos may have alerted learners to pay increased attention to linguistic form as well as encouraged them to provide feedback to one another and to resolve linguistic issues collaboratively. Therefore, it is possible that learners in this study were more inclined to generate higher numbers of LREs in all tasks regardless of the degree of complexity or task type. However, this explanation does not account for the fact that the low proficiency learners produced more LREs during the simple picture narration task. It may be that the complex picture narration task was overwhelming for the low proficiency learners despite the possible advantages of pre-task modeling. Again, this is consistent with Kim’s finding that the low proficiency learners produced more LREs during the simple picture narration task. Thus, according to the results of Kim’s study and the present replication study, it seems that despite the potential mediating effect of pre-task modeling, low proficiency learners interact more during simple picture narration tasks that are not made more complex along the + reasoning demands dimension.

Furthermore, as described in Chapter 2, transcription lengths were corrected for all tasks following Gilabert et al. (2009) in an effort to make the simple tasks comparable to the more complex tasks. This was of particular concern with the simple picture narration tasks, as transcription lengths were on average much shorter than the complex versions due to the the learners being simply required to narrate a story in the former as opposed to ordering and narrating the story in the latter, which took significantly more time and discussion. However, it is not clear if this was done in Kim’s study. Thus, the calculation of ratios for the occurrence of LREs in the present study as opposed to the use of raw numbers in the original study may also have led to some of these differences in the results.
Finally, a Latin Square design was used in this study to avoid any possible ordering effects. Mackey and Gass (2005) note that such counterbalancing is particularly important when administering tasks to participants to ensure that tasks presented first, for example, are not completed by participants with more interest simply because they are completed before subsequent tasks. Kim (2009a) does not report whether this was a feature of her study design. This apparent difference could also have contributed to the different findings between the studies.

The results of the picture narration tasks in the present study also differ from Gilabert et al. (2009) who, like Kim, found that their complex picture narration task resulted in a higher occurrence of LREs among learners. However, as noted in Chapter 1, Gilabert et al. manipulated the complex picture narration task in two ways (+ reasoning demands and + there-and-then dimensions), resulting in a complexity condition in which learners not only had to order the pictures in the narrative, but were required to do so without the visual aid of the pictures while completing the task. Thus, there is a possibility that the more complex task condition in Gilabert et al. pushed learners to produce a higher number of LREs than the complex task condition in the present study. As described in Chapter 2, in the first pilot study, I manipulated the complex task by both the + there-and-then and + reasoning demands dimensions that were used in Gilabert et al. However, the low proficiency learners were largely unable to complete the complex task in this condition, and produced very few LREs. In response, I manipulated the complex task solely by the + reasoning demands dimension for both high and low proficiency learners in the second pilot study and the main study in order to maintain the same task complexity condition for both groups. Thus, there is a possibility that the task complexity condition was not sufficiently complex to lead high proficiency learners to produce significantly more LREs than they did in the simple version. Nonetheless, no firm conclusions can be drawn because these speculations are based on the results of the first pilot study, in which only four learners participated.

**Picture difference tasks.** Following Kim’s (2009a) results, I hypothesized that the low proficiency group would produce more LREs when participating in the more complex picture difference task, which supports the Cognition Hypothesis. However, this hypothesis was not confirmed. Unlike the low proficiency learners in Kim’s study who produced more LREs in the more complex picture difference task, the low proficiency learners in this study produced similar numbers of LREs in both the simple and complex versions of the picture difference task.
Kim (2009a) offers a number of explanations for her findings. She notes that the majority of LREs produced by low proficiency learners during the picture difference task were lexical in nature. Furthermore, this task was made more complex by increasing the number of items in the pictures (few elements). Since the task requires learners to interact with one another by specifying the particular items they are discussing, the amount of vocabulary-related interaction between low learners might be sensitive to increased complexity (increased number of items). Kim speculates that low proficiency learners with less vocabulary might be affected by the – few elements task complexity dimension more than high proficiency learners. However, as discussed above, the results of the present study contradict Kim’s findings, as the low proficiency learners did not produce more LREs in the complex picture difference task. The speculations offered earlier (the inclusion of pre-task modeling, corrected transcription lengths, and the counterbalancing of tasks) may have led to these different findings for the picture difference tasks.

While the low proficiency learners in the present study did not produce more LREs during the complex picture difference task, they did produce a higher proportion of lexical interactive LREs than lexical self-corrections during both the simple and complex picture difference tasks, which is described in detail in the discussion of the fourth research question below. This finding supports Kim’s conclusion that picture difference tasks require learners to collaborate more and be specific about the items they are discussing during task production, which may result in a higher occurrence of LREs involving both learners (defined in the present study as interactive LREs). The following example illustrates such collaboration between two low proficiency learners during the complex picture difference task in this study. Learner A attempts to describe an object in his or her picture (lighthouse), and Learner B speculates and offers suggestions as they negotiate a description of the object in order to identify it.

1. Learner A: And I have one… I don't know how, how to say, spelling, like... boat.
2. Learner B: Light?
3. Learner A: No light, it's like right to the boats... I don't know how.
4. Learner B: You mean like castle building?
5. Learner A: Yeah, and ups of the building have a light.
6. Learner B: Light. Maybe I don't have.
7. Learner A: I don't know how you say this, the name, this in English. In Portuguese I know, it's faro.
Thus, it appears that picture difference tasks lead to a higher occurrence of interactive LREs as opposed to self-corrections. These findings are discussed in greater detail below.

For the high proficiency group, I hypothesized that similar numbers of LREs would occur in the simple and complex versions of the picture difference task, in contradiction to the Cognition Hypothesis. This hypothesis was confirmed and is consistent with Kim’s (2009a) findings. As Kim suggests, task complexity manipulated by the number of items (-few elements) might not facilitate the production of increased interaction-driven learning opportunities for high proficiency learners. This point is also discussed in more detail below.

In summary, two of Kim’s findings were confirmed in the present study, both of which contradict the Cognition Hypothesis. In both the present study and the original study, less complex picture narration tasks that did not involve any + reasoning demands led to greater interaction among low proficiency learners. Furthermore, both studies found that more complex picture difference tasks manipulated by the – few elements dimension did not appear to lead to greater interaction among high proficiency learners. These findings are especially pertinent for the Cognition Hypothesis. While Robinson (2005) argues that the predictions of his hypothesis should hold for learners of any level of proficiency, he has also emphasized the need for empirical research that investigates his claim. The two findings found in both the original study and replication study demonstrate the mediating effect of proficiency on the relationship between task complexity on learner-learner interaction, suggesting that more complex tasks will not always lead to increased learner-learner interaction.

**Additional Research Questions**

The three additional research questions that were explored in order to deepen and extend Kim’s inquiry are discussed in the order in which they were presented in Chapter 1.

**Research question 2:** Do learners at different proficiency levels produce more lexical or grammatical LREs depending on task complexity and task type? As discussed above, Kim presented the occurrence of lexical and grammatical LREs through descriptive statistics, but did not submit her data to any statistical testing. Based on her descriptive results, I predicted that the low proficiency learners would produce more lexical LREs during the simple and complex picture difference tasks as well as the simple picture narration task. I also predicted that the low proficiency learners would produce similar numbers of these two types of LREs during the complex picture narration task. These hypotheses were only partially confirmed. As in Kim, the
low proficiency learners produced significantly more lexical LREs during both the simple and complex picture difference tasks, and they also produced similar numbers of both types of LREs during the complex picture narration task. However, contrary to Kim’s descriptive findings, the low proficiency learners did not produce more lexical LREs during the simple picture narration task.

As discussed earlier, it is worth noting that in the present study, the mean differences between lexical and grammatical LREs were analyzed using non-parametric as opposed to descriptive statistics, as reported in Kim (2009a). Thus, comparisons between the two studies must be made more tentatively. Furthermore, the majority of the data were skewed positively or negatively, meaning the results of the present study need to be interpreted with caution. In addition, the finding that the simple picture narration task did not lead low proficiency learners to produce more lexical LREs could have been a result of the inclusion of pre-task modeling in the present study. In addition to examples of lexical corrections, the videos demonstrated learners dealing with grammatical issues such as verb tense and correct preposition use. Thus, learners may have been alerted to pay increased attention to grammatical issues during task completion. This may have been especially influential during the picture narration tasks, in which learners spoke for longer turns than in the picture difference tasks as they described their pictures in detail and attempted to narrate the story. As a result of these longer turns, learners self-corrected at higher rates than during the picture difference tasks, particularly for grammatical reasons, which is described in more detail below in the discussion of the fourth and final research question. Thus, unlike in Kim’s study, the low proficiency learners’ increased use of grammatical LREs (particularly self-corrections) during the simple picture narration task in the present study led to more of a balance between lexical and grammatical LREs during this task. However, because self-corrections occurred less frequently during the picture difference tasks, the inclusion of pre-task modeling did not appear to balance lexical and grammatical LREs during these tasks. Similar to the original study, the low proficiency learners still produced more lexical LREs than grammatical LREs during the picture difference tasks.

In addition, in line with Kim’s descriptive findings, I predicted that the high proficiency learners would produce more lexical LREs only during the simple and complex picture difference tasks, not the picture narration tasks. My hypothesis was again partially confirmed, as these learners only produced more lexical LREs during the simple picture difference task, not the
complex version. However, similar to Kim, no significant differences between lexical and grammatical LREs were found during the high proficiency learners’ completion of the simple and complex picture narration tasks.

The finding that the complex picture difference task did not lead high proficiency learners to produce more lexical LREs, contradicting Kim’s descriptive findings, indicates that the high proficiency learners in this study produced more LREs focused on grammatical issues than in Kim’s study during this task. Research has shown that learner attention to form increases as proficiency advances (Kim & McDonough, 2008), which could explain why the high proficiency learners produced more grammatical LREs than the low proficiency learners during the complex picture difference task in the present study (which resulted in a balance between lexical and grammatical LREs). However, this does not explain why the high proficiency learners produced more lexical LREs during the simple picture narration task, which is in line with Kim’s findings. It is possible that the higher degree of complexity during the complex picture narration task could have led higher proficiency learners to produce more grammatical LREs, which led to more balanced occurrences of lexical and grammatical LREs. Thus, in the present study, while the simple picture difference task led high proficiency learners to produce more lexical LREs, the more complex task led them to produce similar numbers of lexical and grammatical LREs. Despite these speculations, it remains unclear why the high proficiency learners produced more lexical LREs during the simple picture difference task but not the complex version. Furthermore, as discussed below in the limitations section, the “high” proficiency learners in the present study may not have been directly equivalent to the “high” proficiency learners in Kim’s study, which may have contributed to these different findings.

**Research question 3:** Do task complexity and task type effect the occurrence of LREs produced interactively or individually among L2 learners of different proficiency levels? Given that no previous studies have measured LREs that occur either between two learners or individually as self-corrections, I hypothesized that there would be no differences between simple and complex tasks in either proficiency group when interactive LREs and self-corrections were analyzed independently. Results indicated that there were no differences across tasks aside from the number of LREs produced by the low proficiency learners during the simple picture narration task. During this task, learners produced significantly more interactive LREs than during the complex picture narration task, but not more self-corrections. These findings are
supported by Nuevo (2006), who found that learners produced similar numbers of self-corrections during both simple and complex versions of a picture narration task. On the other hand, these findings contradict those of Gilabert et al. (2009), who found self-corrections to occur at a higher rate in a complex picture narration task. As discussed above, the complex narration tasks used in Gilabert et al. were more complex than those used in the present study (i.e., they included the + there-and-then dimension in addition to the + reasoning demands dimension). Gilabert et al. argue that “the prompt to reconstruct the story in there-and-then forced learners to stretch their attentional and memory resources, to hold the [pictures] in memory for their scrutiny, and to be precise about the information they needed to communicate” (2009, p. 384). They argue that this complex condition led learners to produce more self-corrections because accuracy was essential in providing their interlocutors with a clear image of their pictures, and the difficulty involved in doing this from memory led them to make errors more frequently. Thus, the complexity conditions in Gilabert et al. and the present study were quite different given the higher degree of complexity involved in the former study. This may have contributed to the different findings in the present study.

As discussed in Chapter 1, Fortune and Thorp (2001) argue that self-corrections are not collaborative in nature and thus are not included in their definition of LREs. In their study, they refine the framework for LREs, and argue that while LREs are valuable in analyzing L2 interaction, they fail to capture the complexity of learner collaboration. Unlike the previous studies described in this thesis the present study attempted to explore this issue by comparing interactive LREs and self-corrections. The fact that interactive LREs (but not self-corrections) occurred at higher rates during the simple picture narration task among the low proficiency learners suggests that while manipulations in task complexity may lead to higher rates of some types of LREs (interactive LREs), they may not lead to higher rates of other types of LREs (self-corrections). However, given that neither of these two measures occurred at statistically significant higher rates in other tasks completed by both proficiency groups, no conclusions can be drawn. However, as described below in a discussion of the fourth research question, this subdivision appears to lead to clearer patterns when interactive LREs and self-corrections are further categorized as lexical or grammatical.

**Research question 4:** Do task complexity and task type effect the proportional occurrence of lexical or grammatical LREs that are produced interactively or individually among
L2 learners of different proficiency levels? The present study extended Kim’s inquiry to investigate any patterns that might exist in the relative proportions of interactive LREs and self-corrections that were lexical or grammatical in nature. The most notable finding in this study is that grammatical LREs occurred most often as self-corrections, not as interactive LREs across all tasks and proficiency groups. On the other hand, lexical episodes occurred most often as interactive LREs and not as self-corrections during the picture difference tasks, regardless of proficiency level. In the picture narration tasks however, interactive LREs and self-corrections occurred at more similar rates. This suggests that task type plays an important role in mediating the effect of task complexity on the occurrence of LREs that occur interactively and individually in relation to lexical or grammatical features of English. While picture difference tasks appear to lead learners to interact with one another more often and engage in collaborative interaction-driven learning opportunities in which they tend to negotiate lexical items together, picture narration tasks lead to more balanced levels of lexical LREs that occur individually (as self-corrections) and interactively (as interactive LREs). Along with the finding that grammatical LREs tend to occur as self-corrections, this finding demonstrates the importance of analyzing different types of interaction moves that comprise LREs. Future research should extend these findings by observing if other patterns emerge when different task types and different complexity variables are manipulated.

Limitations

There are a number of limitations in this conceptual replication study that should be acknowledged and addressed in future research. First and foremost, proficiency was determined through participants’ placement in their university language program as opposed to a standardized measure of proficiency. In Kim, however, proficiency was determined both through participants’ placement in their language program as well as their TOEFL PBT scores. Without a standardized proficiency measure, it is impossible, for example, to determine whether the “low” and “high” learners in the current study were of the same proficiency as the “low” and “high” learners in Kim’s study.

There are also several limitations that the present replication study shares with the original study. Both studies were conducted on a relatively small number of adult participants in a university ESL setting, and thus results cannot be generalized to other settings with great confidence such as EFL settings or classrooms with younger learners. In addition, both studies
operationalized task complexity dichotomously, and research suggests that complexity should be operationalized along a continuum as it appears to have a progressive impact as tasks advance from simple to + complex to ++ complex (Kim, 2009b, 2012; Robinson, 2007). Furthermore, the mediating effects of task type should be addressed in future research through the use of a wider variety of task types, as only two task types were employed in the replication study and original study. As Kim (2009a) suggests, future studies should employ a variety of task types manipulated by different complexity variables in order to provide more data for the Cognition Hypothesis.

Finally, while the present study investigated the occurrence of interaction-driven learning opportunities, it did not measure the effect of such occurrences on L2 learning. Research suggests that such opportunities that occur during learner-learner interaction may lead to learner development of specific L2 features such as question formation (Kim, 2009b, 2012). Future studies should continue to measure the occurrence of these opportunities as well as learner development of a wide variety of linguistic features in order to better understand how learners at different proficiency levels might benefit from increases or decreases in task complexity during task-based learner-learner interaction through their acquisition of specific linguistic features.

**Directions for Future Research**

Despite these limitations, a comparison of the original study and this conceptual replication leads to a number of implications for future replication research in SLA. First, replication studies are valuable not only in improving generalizability, as well as confirming or rejecting findings, but they may also provide a means of addressing potential methodological issues in original studies (Porte, 2012). The use of corrected transcription lengths and the counterbalancing of tasks were viewed as important methodological improvements that the present study made to the original study’s methodology. A further improvement was the use of non-parametric statistics as opposed to descriptive statistics in the analysis of the occurrence of lexical and grammatical LREs.

On the one hand the results of the present study that support Kim’s findings strengthen her conclusion that the relationship between task complexity and learner-learner interaction is indeed mediated by task type and learner proficiency. In addition, different complexity dimensions appear to affect this relationship differentially. Thus, some task types and task complexity dimensions, such as narration tasks that require learners to order and narrate a story
simultaneously, may result in less interaction among low proficiency learners despite the fact that pre-task modeling might encourage increased interaction on more complex picture narration tasks. Other task types and dimensions, such as complex picture difference tasks that include many elements, may result in a similar degree of interaction among high proficiency learners as in simpler tasks with fewer elements.

On the other hand, the results that contradict Kim’s findings raise an important issue regarding pre-task modelling. In this study the more complex picture difference task did not lead to greater interaction among low proficiency learners, and the more complex picture narration task did not lead to greater interaction among high proficiency learners, whereas in Kim (2009a), they did. As suggested above, the use of pre-task modeling may have encouraged learners to produce more LREs in the simple versions of these tasks. While Kim argues that the more complex tasks pushed learners to interact more, this pre-task modeling may have mitigated the effect task complexity had on their interactions, resulting in a similar degree of interaction in simple and complex tasks. Thus, the contrary results obtained in the present study suggest that pre-task modeling can be used as a powerful tool that encourages greater interaction among learners, which mediates the effect that task complexity may have on learner-learner interaction.

The present study also demonstrates the challenge involved in conducting replication work, particularly in conceptual replications. The inclusion of pre-task modeling in the current study represents a departure from the original study, thus rendering direct comparisons between the two studies difficult. However, this modification also serves as an improvement given that research published after Kim (2009a) demonstrates the advantages pre-task modeling may have for learners engaged in task-based interaction (Kim, 2013; Kim & McDonough, 2013). The use of pre-task modeling in the present study demonstrates that while proficiency mediates the effects task complexity and task type have on learner-learner interaction (Kim, 2009a), such modeling may also play a mediating role by encouraging greater interaction among learners at both low and high proficiency levels regardless of whether a task is simple or complex. Future research should acknowledge pre-task modeling as an intervening variable that may mediate the relationship between task complexity and the occurrence of interaction-driven learning opportunities. However, as discussed above, despite the mediating effect of pre-task modeling, it appears that for some task types, language proficiency may still play a more important role in the effects of task complexity on learner-learner interaction. Even with the support of pre-task
modeling, some complex tasks (i.e., picture narration tasks) may prove too difficult for lower proficiency learners, leading to less learner-learner interaction.

**Pedagogical Implications and Conclusion**

The present study has several pedagogical implications, particularly for the use of interactive tasks in L2 classrooms. Like Kim, the finding that the simple picture narration task led to higher rates of interaction between low proficiency pairs suggests that teachers, curriculum developers, and material designers should carefully consider students’ proficiency when designing and using interactive tasks for L2 learners. Higher complexity tasks may in fact lead to less interaction for lower proficiency learners, particularly for narration tasks that push learners to produce language for extended turns.

The results of the current study also suggest that though more complex picture difference tasks may increase learner production of interaction-driven learning opportunities at lower proficiency levels, pre-task modeling may balance this tendency across simple and complex picture difference tasks. Thus, low proficiency learners might be encouraged to produce more interaction-driven learning opportunities in both simple and complex tasks through the use of pre-task modeling, which may mitigate the need to manipulate task complexity in the first place. Thus, educators may have the option of relying more heavily on task complexity to encourage interaction or use pre-task modeling as an alternative. The use of pre-task modeling as opposed to more complex tasks may be appropriate if learners feel frustrated, overwhelmed, or demotivated by the degree of complexity involved in a task as the participants in the first pilot study were in the present study.

In regards to grammatical and lexical LREs, like Kim, this study suggests that in general, picture difference tasks lead to learner-learner interaction that is primarily focused on lexical issues as opposed to grammatical ones. Although the present study did not find this to be the case for the high proficiency learners’ performance during the complex picture difference task, the fact that high proficiency learners produced more lexical LREs in the simple picture difference task and the low proficiency learners produced more during both simple and complex versions of the picture difference tasks suggests that particular task types (i.e., picture difference tasks) might be effectively used in order to provide learners with increased opportunities to deal with lexical items. Other tasks, such as picture narration tasks, might be more appropriate for
increased exposure to and practice with grammatical issues (e.g., verb tense while narrating stories, more opportunity for grammatical self-correction due to longer turns, etc.).

Both the replication and the original study demonstrate that learners do attend to both lexical and grammatical features of English during task-based interaction, and that educators may use appropriately pre-designed tasks to provide learners of different proficiency levels with opportunities to interact with other learners, experiment with a L2, and provide one another with feedback. Such findings are particularly relevant for instructors who find it difficult to interact one-on-one with individual learners in large classrooms, as tasks can provide students with a means of both analyzing and practicing language collaboratively. As Long (1996) argues, such collaboration may lead to L2 acquisition. While neither study investigated L2 learning, other research suggests that learners can make measurable progress on specific L2 features through such task-based learner-learner interaction (Kim, 2009b, 2012). Furthermore, pre-task modeling may also be used by instructors, either through the use of video or more simply demonstrated live in the classroom to encourage collaborative patterns of interaction in pairs or small groups characterized by balanced participation among learners, the negotiation of lexical and grammatical items, the provision of feedback, and the use of self-correction as a strategy, particularly when learners are intent on improving L2 accuracy. Such modeling may be used in conjunction with appropriate degrees of task complexity in order to match pairs of learners to tasks appropriate to their level of proficiency in order to maximize the degree of interaction between them (Robinson, 2005).

Finally, it is important to remember that the objectives of such task-based research invariably lead to the larger question of how educators might organize a syllabus that sequences tasks in a principled manner. Kim’s (2009a) study along with the present replication study contribute to this effort by demonstrating how individual difference variables such as proficiency and task type must be taken into account when making predictions about the effects of task complexity on learner-learner interaction. Further contributions, including future replication work, will be necessary if researchers and educators alike hope to establish a set of principles that may help guide learners through the process of acquiring an L2 within the classroom as efficiently and productively as possible.
References


Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie and T. Bjatia (Eds.), *Handbook of second language acquisition* (pp. 413-


Replication research in Applied Linguistics (pp. 21-46). New York: Cambridge 
University Press.

In G. Porte (Ed.), Replication research in Applied Linguistics (pp. 21-46). New York: 
Cambridge University Press.

Language Acquisition, 19, 499-508.

(pp. 1-17). Cambridge: Cambridge University Press.


Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A 


framework for second language task design. International Review of Applied Linguistics 
in Language Teaching, 43(1), 1-32.

speech production, interaction, uptake and perceptions of task difficulty. International 

language learning and performance. IRAL, International Review of Applied Linguistics in 
Language Teaching, 45(3), 161-176.

Macmillan.

Press.


Appendix A
Sample Materials

Simple Narrative\(^7\) (Gilabert et al., 2009)

---

\(^7\) Permission to reproduce the illustrations on page 64 and 65 of this thesis was provided in the form of an email by De Gruyter Rights and Licenses Team. Illustrations are originally taken from the following publication: (Gilabert et al., “Manipulating cognitive complexity across task types and its impact on learners’ interaction during oral performance”; in: International Review of Applied Linguistics, 47, 2009, pp. 367-395, ill. 387-388.)
Complex Narrative (Gilabert et al., 2009)

**Student A**

**Student B**
Appendix B
Transcribing Conventions

If you cannot understand a word a speaker is trying to pronounce, just write what you hear and indicate this beside the word with a double question mark in parentheses (??) (e.g., “the child has takshee, takshee, (??) for the all”). If the recording is unclear to the point where you are unable to write out the sounds you hear, then indicate this with an asterisk *. Include one asterisk for each missed word. Then include (??) immediately after the asterisk. For example, “she go to the * (??)”

Please highlight mispronounced words by writing the word you believe the speaker intends to pronounce and include the word “pron” in parentheses followed by the way in which the speaker actually pronounced the word. Also, please bold the word “pron” so that these instances may be easily found in the transcription. For example, “there are four birds (pron beards”).

Include repetitions of words as well as false starts and write them as they are pronounced. For example, “in my picture, picture, there are a woman, a wo, a woman”. Please note that it is important that I know when a false start sounds the same as the word eventually pronounced, or when the word eventually pronounced includes a different sound than originally uttered. For example, (1) “a chi, a chi (as in child), a child” (in this case, the false start sounds included in “chi” are identical to the word eventually uttered, “child”). On the other hand, in the following example, the sounds originally uttered are different from the sounds included in the word eventually uttered: (2) “a choy, choy, chew, child”. Example two is considered a self-correction, whereas example one is considered a false start, which is deemed to be a result of a lack of fluency as speakers repeat a sound.
Appendix C
Coding Scheme for Language Related Episodes

General Definition

A language-related episode is defined by Swain and Lapkin (1995) as follows: “A LRE episode is defined as any part of a dialogue where the students talk about the language they are producing, question their language use, or correct themselves or others” (Swain & Lapkin, 1995).

Kim (2009) uses the above definition in her study and elaborates with the following details:

LREs comprise a variety of discourse functions during learner-learner interaction such as:

- Requests for assistance
  - Example
    1: But the picture, the picture how do you say…..
    2: Fall down?
    1: Yeah fall down because of the music
  - Negotiation sequences (eg: learners are unsure of language and experiment with possibilities either by themselves or with their partner, or a lack of understanding leads a learner to rephrase something with different words to make it more target-like)
    - Example
      1: And and what what they do? What they
      2: The man? The man do?
      1: Yeah (1’s use of ‘they’ is incorrect, so 2 asks for clarification, “the man?”)
  - Explicit and implicit feedback (including self or other correction)
  - Incidences where the learners notice a gap between their interlanguage and the target language system

In all cases, LREs are limited to areas of the dialogue related to form, not meaning. Thus, learner interactions that do not involve attention paid to grammar, vocabulary, pronunciation, spelling, and so on are NOT coded as LREs (Kim, 2009; Revesz, 2011; Swain & Lapkin, 1995; Williams, 1998).

Coding for LREs

1) Each LRE deals with one language item (one vocabulary word, one phrasal verb, one verb, one verb as its tense is questioned or self-corrected, one word that is pronounced in several different ways, etc.). This is the case even for LREs that are discontinuous (i.e., they occur in different areas of the transcript and are separated from each other by many lines). This specification is made in Fortune & Thorp (2001).

2) When a participant begins to pronounce a sound or word but then immediately self corrects by using a different sound or word BEFORE that word is completely uttered, it is considered a LRE (and is coded as lexical)
Example 1
1: The the child said goodbye to their mom and they felt happy, and they took the bucket to outside have picnic, finish.

10) Self-correction

If a participant corrects himself/herself several times and searches for the right word or form (e.g., tense) consecutively, this is counted as one single self-correction.

Example 1
1: He comes, he come, he's came, he's coming to

In the example above, even though the participant corrects himself/herself three times, he/she is dealing with one language item, which is the correct tense for ‘come’ in this sentence.

However, if the self-corrections are related to different language items, count them separately.

Example 1
1: But you, you can, you can see in the picture like he, they are, they were still, still listening to music.

In the example above, the first correction is coded as grammatical (pronoun selection, he/they), and the second is also coded as grammatical (tense, are/were). They are counted as TWO independent self-correction grammatical LREs (SELF GRAM).

Coding Lexical and Grammatical LREs

When distinguishing between lexical and grammatical LREs, the present study uses Kim’s definition of grammatical and lexical LREs (2009a) while also further clarifying this difference through the use of Swain and Lapkin (2002).

Grammatical LRE
- LREs in which learners discussed grammatical features of English, including verb tense, word order, preposition use and so on. (Kim, 2009a)

Lexical LRE
- LREs in which the learners discuss the meaning, pronunciation or spelling of lexical items (Kim, 2009a)

Following Swain and Lapkin (2002), the following clarifications have been made to Kim’s coding scheme in the effort to make categorization of lexical and grammatical LREs less ambiguous. It is important to note that Swain and Lapkin (2002) categorize LREs involving spelling as grammatical. However, in an effort to follow the same coding procedures in Kim’s study (2009a), spelling is categorized as lexical in the present study.
Lexical: includes adverbs, nouns, adjectives, verbs, and spelling.
Grammatical: includes articles, possessive pronoun/articles, prepositions, pronoun reference, sentence structure, verb tense, and verb form.

Additional Notes

The following areas of the dialogue are NOT coded as LREs.

1) Mere repetition (eg: I can see two ball, two ball in my picture.)
2) False starts that contain the exact same sounds as the word that is eventually spoken, except when a participant adds an ending to a verb to change its tense, to change it from the infinitive to a gerund or vice versa, to change a statement from affirmative to negative or vice versa, to change a word from singular to plural or vice versa, or to add or omit the possessive ‘s.

Example 1 (NO LREs)
1: He w, w, wearing a green shirt.
2: I see t, t, two balls.

The following examples illustrate exceptions that are coded as LREs:

Example 2 (Exception -> infinitive to gerund change)
1: He play, playing with a ball.

Example 3 (Exception-> verb tense change)
1: yes in the basket. And a dog look, a dog looked basket and the dog yes. OK number three. (look/looked)

Example 4 (Exception->singular to plural)
1: difference yeah yeah. Two ball, two beach balls has a lot of colours right?

Example 5 (Exception->addition of the possessive ‘s)
1: the boat the boat boat’s colour is red ♥F(possessive ‘s)
### Appendix D

**Supplemental Statistical Tables**

**Table D1** Average LREs produced per 100 words during the picture narration tasks organized by type of interaction and language focus

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Language Focus</th>
<th>Type of Interaction</th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sum</td>
<td>Mdn</td>
</tr>
<tr>
<td>Low Lexical</td>
<td></td>
<td>Interactive LREs</td>
<td>5.57</td>
<td>.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>4.35</td>
<td>.54</td>
</tr>
<tr>
<td>Grammatical</td>
<td></td>
<td>Interactive LREs</td>
<td>1.01</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>9.23</td>
<td>.83</td>
</tr>
<tr>
<td>High Lexical</td>
<td></td>
<td>Interactive LREs</td>
<td>4.30</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>3.98</td>
<td>.55</td>
</tr>
<tr>
<td>Grammatical</td>
<td></td>
<td>Interactive LREs</td>
<td>.74</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections</td>
<td>10.13</td>
<td>1.13</td>
</tr>
</tbody>
</table>
Table D2: Average LREs produced per 100 words during the picture difference tasks organized by type of interaction and language focus.

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>Language Focus</th>
<th>Type of Interaction</th>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td>Mdn</td>
<td>IQR</td>
</tr>
<tr>
<td>Low</td>
<td>Lexical</td>
<td>Interactive LREs 6.17</td>
<td>.64</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections 2.84</td>
<td>.35</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>Grammatical</td>
<td>Interactive LREs .50</td>
<td>.07</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections 4.04</td>
<td>.58</td>
<td>.62</td>
</tr>
<tr>
<td>High</td>
<td>Lexical</td>
<td>Interactive LREs 6.39</td>
<td>.85</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections 3.68</td>
<td>.43</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td>Grammatical</td>
<td>Interactive LREs .47</td>
<td>.03</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-Corrections 5.68</td>
<td>.68</td>
<td>.51</td>
</tr>
</tbody>
</table>