Studies in Second Language Acquisition, 2012, **34**, 627–658.
doi:10.1017/S0272263112000368

TASK COMPLEXITY, LEARNING OPPORTUNITIES, AND KOREAN EFL LEARNERS' QUESTION DEVELOPMENT

YouJin Kim
Georgia State University

Building on the cognitive and interactive perspectives of task research, the cognition hypothesis states that increasing task complexity promotes greater interaction and feedback and thus facilitates second language (L2) development (Robinson, 2001b, 2007a). To date, very little research has explored this claim during learner-learner interactions in English as a foreign language (EFL) classroom contexts in which a task-based syllabus is implemented. The purpose of this study was to examine the effects of task complexity on the occurrence of interaction-driven learning opportunities and question development in such a classroom context. Korean university students (N = 191) from four intact English classes were randomly assigned to one comparison group and three experimental groups with various task-complexity levels (i.e., simple, +complex, and ++complex) based on Robinson's framework. Their interactions were audiorecorded, and occurrences of learning opportunities, operationalized as language-related episodes (LREs), were identified. Question development was identified between pretest and posttests on the basis of the learners'

This research grew out of my PhD dissertation. I would like to thank my advisor, Kim McDonough, and dissertation committee members, Joan Jamieson, Mary McGroarty, and Joseph Collentine, for their constant guidance and support throughout the development of the project. I am also very grateful to Peter Robinson and to anonymous *SSLA* reviewers for their insightful comments and suggestions. Special thanks go to Nicole Tracy-Ventura, Heidi Vellenga, Bethany Grey, and Diana Trebing for their helpful comments at different stages of this article. The project was supported by the *Language Learning* dissertation grant and TESOL's Ruth Crymes Fellowship for Graduate Study. My sincere thanks also go to ETS for providing the TOEIC Bridge. Finally, I would like to thank all the learners who participated in the study. Any remaining errors, of course, are my own.

Address correspondence to YouJin Kim, Department of Applied Linguistics and ESL, Georgia State University, 34 Peachtree Street, Suite 1200, Atlanta, GA 30303; e-mail: ykim39@gsu.edu.

© Cambridge University Press 2012

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

stage advancement using Pienemann and Johnston's (1987) developmental sequence. Results indicated that more complex tasks promoted a greater number of LREs and particularly led to LREs targeting developmentally advanced question structures, which result in question development.

Over the past few decades tasks have been seen as a primary unit of second language (L2) instruction and have provided a fruitful common ground between L2 pedagogy and research (Bygate, Skehan, & Swain, 2001; Ellis, 2003; Robinson, 2011). Tasks can offer sufficient means for the authentic presentation of appropriate target language input to learners and can generate ample opportunities for language use in meaningful contexts while also promoting learners' attention to linguistic forms (Bygate et al., 2001). Although task-based instruction arose out of pedagogical proposals for greater emphasis on communicative activities in language teaching, SLA research has increasingly informed task-based instruction (Robinson, 2011). In particular, SLA researchers have made various predictions about different task design and implementation factors that affect L2 development in task-based instruction. For example, two competing accounts of how attention is deployed during task performance (i.e., the single-resource model vs. the multiple-resource model) have provided implications for task design, specifically in relation to task complexity.

With a multiple-resource view of attention, Robinson proposed the cognition hypothesis, stating that task complexity would promote the frequency of interactional features such as feedback and negotiation of meaning during task performance and thus facilitate L2 learning (Robinson, 2001b, 2003, 2005, 2007a). Although Robinson (2009) makes claims about how to design syllabi on the basis of the cognition hypothesis, very little research has been carried out testing his predictions in classroom contexts in which a task-based syllabus has been implemented. Thus, the purpose of the present study is to fill this gap by examining the role of task complexity in the occurrence of interaction-driven learning opportunities and L2 development in task-based English as a foreign language (EFL) classroom settings.

LITERATURE REVIEW

The Limited Capacity Hypothesis versus the Cognition Hypothesis

The two influential claims regarding the extent to which task characteristics can affect the allocation of the learners' attention during task

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

performance are Skehan's (1998) limited capacity hypothesis and Robinson's (2001b) cognition hypothesis. A theoretical construct of importance for these two hypotheses is the "notion of attention as capacity," and both hypotheses are interested in how increasing the attentional demands of tasks affects language production (Robinson, 2011, p. 17). On the basis of a single pool of attentional resources for L2 learners, the limited capacity hypothesis states that learners' attentional resources are limited and that increasing task complexity reduces a pool of attentional capacity during task performance. On the basis of the limited capacity hypothesis, Skehan (1998) and Skehan and Foster (1999, 2001) argued that a trade-off exists between attention to form and attention to meaning during task performance. As a result, more complex tasks will demand more attention to content and, correspondingly, will allow less attention to language. Furthermore, in Skehan's (1998) view, due to limitations in attentional resources, learners cannot attend to all aspects of language production at the same time (e.g., complexity, fluency, accuracy). Thus, tasks can promote either increased complexity or accuracy, but not both.

Whereas Skehan's (1998) limited capacity hypothesis argues for the single-resource model of attention, Robinson's (2001a, 2001b, 2003, 2005) cognition hypothesis predicts that learners are able to access multiple and noncompetitive pools of attention. According to Robinson, there is not a trade-off between attention to accuracy and attention to complexity of language production. Rather, he claims that increasing task complexity promotes more accurate and more complex language.

In his task-complexity framework, Robinson (2001a, 2001b, 2003) includes a series of predictions regarding the role of task complexity on L2 performance and development in which he classifies task complexity into two dimensions: cognitive-conceptual (i.e., resource-directing) and performative-procedural (i.e., resource-dispersing) demands. Robinson (2001b) argued that the two task-complexity categories identify an important difference in the way these dimensions affect resource allocation during L2 task performance. He thus claimed that the effects of task complexity in the two kinds of dimensions (i.e., resource-directing and resource-dispersing) are very different.

According to Robinson (2001b), resource-directing variables of task complexity make greater demands on attention and working memory in a way that redirects them to linguistic resources during task performance. Therefore, increasing task complexity along resource-directing dimensions—for example, by requiring learners to use reasoning skills [+reasoning demands] to consider many elements [–few elements] or to narrate events that are displaced in time and space [–here and now]—can direct learners' attention to specific, task-relevant linguistic features.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

On the contrary, resource-dispersing variables are those that make increased performative-procedural demands on participants' attentional and memory resources but do not direct them to any element of the linguistic system (Robinson, 2001b, 2005). Making tasks more complex along resource-dispersing dimensions—for instance, by requiring learners to perform more than one task simultaneously [–single task] or by providing no prior knowledge support [–prior knowledge] or planning time [–planning time]—leads learners to disperse attention over many nonlinguistic areas during task performance.

Whereas Skehan's limited capacity hypothesis (1998) predicts that increasing the cognitive demands of tasks would negatively affect both accuracy and linguistic complexity of learner production, Robinson's cognition hypothesis claims that making tasks more complex in the resource-directing dimensions will increase linguistic accuracy and complexity (e.g., Robinson, 2001b, 2005, 2007b). Robinson also predicts that increasing task complexity would encourage learners to look for more assistance in the input and attend to linguistic codes that are required for task completion (Robinson, 2001a; Robinson & Gilabert, 2007). With regard to L2 development, Robinson has claimed that increasing the task complexity of oral interactive tasks leads to L2 development, especially with developmentally more advanced forms (Robinson, 2007a, 2007b; Robinson & Ellis, 2008). In task-based learner-learner interaction contexts, increasing complexity along resource-directing dimensions has the potential to direct learners' attentional and memory resources to L2 structures, providing "learning opportunities" and thus ultimately leading to interlanguage development (Robinson, 2007b, p. 23).

To date, a variety of predictions about the effects of task complexity in Robinson's (2001b) framework have been tested, focusing mainly on L2 linguistic performance (i.e., complexity, accuracy, and fluency) during either oral or written tasks (Gilabert, 2007; Ishikawa, 2007; Kuiken, Mos, & Vedder, 2005; Kuiken & Vedder, 2007; Michel, Kuiken, & Vedder, 2007; Robinson, 2001a, 2007b). However, the findings of these studies have not been conclusive; they suggest that more complex tasks positively impact linguistic performance in general, yet more specific findings related to both accuracy and syntactic complexity only partially supported the cognition hypothesis (e.g., promoting either complexity or accuracy). Although language learning requires that learners increase the complexity, accuracy, and fluency of their language production, these measures do not capture all of the processes of L2 acquisition; particularly, they miss those related to development of specific linguistic forms in meaning-oriented language production. To address such concerns, a growing number of studies examined the relationship between task complexity, learners' attention to linguistic forms during interaction, and subsequent development of those structures.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

Task Complexity, Interaction, and L2 Learning

The construct of cognitive complexity has received growing attention in SLA studies that aim to explore how increases in the cognitive complexity of tasks differentially affect interaction and learning outcomes (Baralt, 2010; Gilabert, Barón, & Llanes, 2009; Kim, 2009a, 2009b; Nuevo, 2006; Révész, 2009, 2011; Robinson, 2001a, 2007b). Some studies were conducted in laboratory settings in which learners met with a researcher or a native speaker (e.g., Baralt, 2010; Révész, 2009) or in classrooms where learners were paired with other learners during task performance (e.g., Kim, 2009a, 2009b; Révész, 2011). Because the present study focuses on learner-learner interaction, studies examining interaction between learners will be discussed in this section.

Previous task-complexity studies operationalized interactional features in various ways, including the occurrence of specific feedback types such as recasts and clarification requests and the number of language-related episodes (LREs). These included both one-way (Robinson, 2001b, 2007a) and two-way (Kim, 2009a; Nuevo, 2006; Révész, 2011) tasks in either EFL or English as a second language (ESL) settings. The first study exploring task complexity and interaction was Robinson (2001b), in which he manipulated [+/–prior knowledge] and [+/–few elements] variables to create different levels of task complexity with map tasks. A simple version of the task used a small map of an area known to the Japanese learners, whereas a complex version used an authentic street map of a much larger area likely to be unknown to the learners. A total of 44 Japanese university students participated in two versions of these map tasks (simple and complex). In pairs, one participant (speaker) was asked to give directions from point A to point B to a partner (hearer) who had only point A on his or her map. The hearer was then required to draw the routes between points A and B described by the speaker. The findings showed significantly greater amounts of interaction with more confirmation checks on the complex version of the task.

Robinson (2007b) examined the effects of task complexity on L2 speech production, interaction, uptake, and perceptions of task difficulty with 42 Japanese EFL learners. He used a [+/–reasoning demands] variable to manipulate three degrees of task complexity. During learner-learner interaction, the speaker decided on the correct sequence of a series of pictures and then narrated this story to the listener, who had to sequence the pictures on the basis of the order of narration. To assess the relationship between the task complexity of the narration tasks and the extent of interaction and learning opportunities, Robinson analyzed the number of turns taken, clarification requests, and confirmation checks for the listener-sequencer on each task version. Findings indicated that task complexity, along the resource-directing dimension,

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

significantly promoted the amount of interaction (i.e., more turns, which consisted of greater numbers of clarification requests and confirmation checks). Overall, Robinson's previous studies in which learners carried out one-way tasks in pairs suggested that task complexity along resource-directing dimensions positively promoted interaction-driven learning opportunities, and that it particularly resulted in significantly more confirmation checks with complex tasks in both experiments (Robinson, 2001b, 2007b).

Because previous interaction studies suggested that two-way tasks are more effective in promoting interaction (Ellis, 2003), two-way tasks during learner-learner interaction have been increasingly used. Additionally, a growing number of studies operationalized interaction in various ways and explored how other task and learner factors mediate the relationship between task complexity and interaction. For instance, Kim (2009a) explored to what extent learner proficiency mediates the relationship between task complexity and the occurrence of interaction-driven learning opportunities during two different task types (i.e., picture-difference tasks and narration tasks). She examined the impact of task complexity along resource-directing dimensions on the occurrence of LREs with ESL learners (n = 34) from two different proficiency levels. In pairs, learners carried out a picture-difference task and a narration task in which task complexity was manipulated using two variables: [+/–few elements] for the picture-difference task and [+/–reasoning demands] for the narration task. The results suggested that the effects of task complexity on the occurrence of learning opportunities—operationalized as the occurrence of LREs—differed depending on task type and learner proficiency. More specifically, the complex narration task elicited more LREs by the higher-proficiency learners than the simple narration task, whereas the opposite findings were obtained with the lower-proficiency learners. For the picture-difference task, the lower-proficiency learners produced significantly more LREs, particularly lexical LREs, during the complex task than the simple task. From a methodological point of view, such findings suggest that in addition to learner variables, task variables other than complexity should be taken into consideration when designing task-complexity studies.

Gilabert and colleagues (2009) also explored how the impact of task complexity on interaction can be mediated by other factors by examining increases in cognitive complexity across three different types of tasks. In pairs, a total of 54 EFL learners carried out two complex versions of three tasks (i.e., a narrative task, an instruction-giving task, and a decision-making task), which were manipulated by [+/–here-and-now], [+/–few elements], and [+/–causal reasoning]. During task-based interaction, interactional features such as confirmation checks, clarification requests, comprehension checks, recasts, LREs, and repairs were

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

identified. The findings showed that in general the more complex versions of tasks resulted in more interactional moves and opportunities for negotiation of meaning. Similar to Kim (2009a), however, they found that predictions of the cognition hypothesis did not hold for all task types. For example, the complex version of the narrative task resulted in significantly more clarification requests but significantly fewer comprehension checks. There were significantly more LREs and repairs (but not recasts) during the complex version of the narrative task. These same findings held for the map task as well. The complex version of a map task yielded significantly more confirmation and comprehension checks, whereas the decision-making task only showed an effect for task complexity in terms of number of repairs. Such findings suggested that task complexity tended to promote more interactional features, particularly during the narrative task and the map task, but not during the decision-making task.

In a recent classroom-based study, Révész (2011) tested Robinson's cognition hypothesis with 43 ESL learners by manipulating a [+/–reasoning demands] variable. Each participant carried out two versions (simple vs. complex) of the same argumentative task in groups. To identify interaction-driven learning opportunities, Révész analyzed LREs, confirmation checks, clarification requests, recasts, and metalinguistic talk. Results showed that the complex task led to a significantly higher number of LREs and metalinguistic talk. However, other measures (i.e., clarification requests, confirmation checks, and recasts) also increased in number descriptively but not statistically. In general, her findings supported the cognition hypothesis, in that students who carried out more complex tasks produced more interactional features, which indicated a larger amount of attention to linguistic forms during task performance.

These studies predominantly examined the amount of interaction; however, the direct relationship among task complexity, the amount of interactional features during learner-learner interaction and L2 development has not been widely investigated. Nuevo (2006) took a first step to examine such a relationship by researching the role of task complexity manipulated by [+/–reasoning demands] on L2 learning opportunities and development, and focused on past tense and locative prepositions. The study used a pretest–posttest–delayed posttest research design, and three groups (n = 113) of learners participated in two different tasks with different levels of complexity—that is, narration tasks and decision-making tasks. She analyzed nine different interactional processes, and findings showed that the simple task version led to greater uptake of recasts, more comprehension checks, and more other repetitions than the complex version. Additionally, the simple version of the decision-making task resulted in a greater amount of metalinguistic talk, whereas the complex version of the tasks led to

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

more confirmation checks. These findings indicated that different task-complexity conditions promoted different sources of learning opportunities. In terms of L2 development, results showed no association between task complexity and L2 learning, counter to the prediction of the cognition hypothesis.

Overall, the existing studies suggest that task complexity can impact the amount of interaction-driven learning opportunities and L2 learning during task-based interaction. However, previous studies have provided inconclusive findings regarding the nature of the relationship between the three constructs. For instance, although Nuevo (2006) found contradictory evidence for the cognition hypothesis—especially with regard to the role of task complexity and L2 development—other studies supported the hypothesis in terms of complex tasks leading to more interaction, with some variations among different interactional features (Gilabert et al., 2009; Révész, 2011; Robinson, 2001b, 2007b).

These mixed findings may be due to some methodological issues in the previous literature. With regard to learning opportunities, previous studies (Kim, 2009a; Nuevo, 2006; Révész, 2011; Robinson, 2001a, 2007b) either did not specify the different types of linguistic forms to which learners paid attention or presented very broad categories for linguistic structures that were discussed during task-based interaction (e.g., lexis vs. grammar). Thus, the previous literature reviewed here has not been able to address the issue of the areas in which the L2 learners tended to notice gaps in their interlanguage during task-based interaction. To address this issue, the present study defined learning opportunities more narrowly with a specific linguistic feature (i.e., question formation).

When interpreting the results of studies focusing on learning opportunities, it should be noted that learning opportunities are not necessarily a direct indicator of subsequent L2 linguistic development. Thus, separate measures for L2 learning should be provided to determine L2 development (Philp, Oliver, & Mackey, 2006). To date, only one study (Nuevo, 2006) investigated the role of task complexity in L2 linguistic development as an outcome of task performance during learner-learner interaction. Because Nuevo did not support Robinson's claim that more complex tasks promote L2 development, there is no positive evidence for the hypothesis in learner-learner interaction contexts. Furthermore, Nuevo's study analyzed two grammatical structures (i.e., past tense and locative prepositions) in terms of accuracy and did not consider the developmental sequences of the structures. Consequently, Robinson's claim that task complexity encourages learners to use developmentally advanced forms during task performance and facilitates their subsequent L2 learning has not been directly tested so far. Additionally, according to Robinson (2007a), functional demands created by a task with different levels of task complexity should be related to linguistic

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

demands imposed on the learners; thus, the target structures should be carefully chosen considering the demands of different task complexities. Studies that focus on such methodological issues are much needed to examine the relationships between task complexity, learning opportunities for developmentally advanced forms, and L2 development.

In terms of the operationalization of task complexity, the majority of previous studies, except Robinson (2007b) and Lee (2002), operationalized task complexity dichotomously (i.e., simple vs. complex) by manipulating the presence and absence of each variable (e.g., +/–reasoning, +/–few elements). Because the concept of task complexity can be determined along a continuum, it is important to include multiple degrees of task complexity in the research design to provide solid evidence for or against the cognition hypothesis. Previous studies, which did not find full support for the cognition hypothesis, have also suggested that the lack of empirical support for the hypothesis could have been due to the lack of task-complexity distinctiveness among tasks (e.g., Kuiken et al., 2005; Nuevo, 2006). As a result, by including multiple degrees of task complexity (simple, +complex, and ++complex), the current study addresses both methodological concerns regarding distinctive task-complexity levels and theoretical issues related to the construct of task complexity. With regard to contexts of research, previous task complexity studies often draw pedagogical implications from either lab-based or classroom-based studies in which tasks were provided outside course curricula. However, as Robinson (2011) argued, there is a clear need to examine the generalizability of pedagogical implications drawn from such research to actual classroom settings in which tasks are implemented following task-based syllabi.

In sum, the purpose of the current study is to contribute to this line of research by examining whether the manipulation of multiple levels of task complexity during task-based instruction has the potential to impact the occurrence of interaction-driven learning opportunities for question formation as well as subsequent question development. Additionally, by implementing tasks in task-based L2 classrooms, the current study attempts to serve as a bridge addressing both the theoretical and pedagogical implications of the cognition hypothesis. The study is guided by the following research questions and hypotheses:

1. How does task complexity impact the occurrence of learning opportunities for questions during task-based learner-learner interaction?

   Following Robinson's cognition hypothesis (2001b, 2007a) and some of the previous classroom-based studies, which provided at least partial support for the hypothesis (Révész, 2011), it was hypothesized that increasing task complexity would encourage learners to produce more learning opportunities for question formation during learner-learner interaction.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

2. What effect does task complexity have on the extent to which task-based instruction promotes question development in comparison to traditional instruction?

   Following previous literature that focuses on the benefits of pedagogical tasks (e.g., Ellis, 2003), it was hypothesized that task-based instruction would be more beneficial for question development compared to traditional instruction. Furthermore, based on the cognition hypothesis, it was predicted that more complex tasks would promote learners' question development. This prediction can also be supported by previous interaction studies that have demonstrated the benefits of interactional features (e.g., feedback, LREs) during learner-learner interaction on subsequent L2 development (e.g., Kim, 2008; McDonough & Sunitham, 2009).

METHOD

Participants

The participants were 191 (females, n = 99; males, n = 92) adult EFL students at a large private university in South Korea who were enrolled in a two-credit freshman English course required for their bachelor's degree. The learners' ages ranged from 18 to 26 years (M = 18.83; SD = 1.08), and the length of their previous English study ranged from 6 to 15 years (M = 9.5; SD = 1.88). The learners had taken at least 6 years of mandatory English classes at junior high and high schools before enrolling at this university. Approximately 30% of the learners (n = 56) reported having basic knowledge of another language besides Korean or English, including Japanese, Chinese, French, and German. Only 2 learners reported studying English in another country (the Philippines and the United States), with their lengths of stay in those countries ranging from 1 to 3 months. Based on the Test of English for International Communication (TOEIC) Bridge scores, the participants' proficiency ranged from beginners to intermediates (see the Results section for more information). They were majoring in various undergraduate degree programs, including education, mass communication, Korean literature, Chinese, law, social work, agricultural science, and computer science.

Four sections of the same course were selected for the current study, three of which were randomly assigned to experimental groups (i.e., task groups) with three levels of task complexity (n = 45 for the simple group, n = 47 for the +complex group, and n = 51 for the ++complex group), and one of which was assigned to a comparison group (i.e., no task group; n = 48). Task-based instruction was provided to the three experimental groups, whereas the comparison group was provided with instructional treatments similar to those provided by the four instructors who taught the other seven sections of

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

the same course.[1] The class met once a week for 2 hr over a 16-week semester.

Target Linguistic Form

Because the cognition hypothesis claims that increasing the cognitive demands of tasks results in the learning of developmentally more advanced forms, a structure with an empirically proven developmental sequence was considered appropriate for the study (Robinson, 2007a, 2007b; Robinson & Ellis, 2008). Additionally, given that increasing task complexity is related to increasing the functional demands of using a particular linguistic feature, it was important to investigate a linguistic feature that is related to the functional demands of the different task complexity levels used in the current study (see the Treatment Tasks section). Although the participants in the current study were expected to be familiar with the target structure because it was introduced in their junior high school textbooks, their oral production opportunities for questions were very limited. Thus, learners' knowledge and use of English questions were investigated in the current study.

The development of English questions has been widely studied in previous L2 interaction research (Adams, 2007; Loewen & Nabei, 2007; Mackey, 1999, 2006; Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006; Philp, 2003). The majority of these studies examined learners' question development by building on the developmental model by Pienemann and Johnston (1987). Following the previous literature, the present study used this question-development model, which is based on emergence criteria rather than accuracy.[2]

Pienemann and Johnston's (1987) developmental stages are governed by processing mechanisms that constrain the movement from one stage to the next. As learners develop questions through the sequence (i.e., Stage 1 through Stage 6), they progressively start to include question words and auxiliary verbs, to invert subjects and auxiliary verbs, to avoid unnecessary auxiliary verbs, and to omit subject auxiliary inversion in embedded questions (see Appendix for examples). Learners often produce questions from multiple stages in addition to their highest stage.

Instructional Materials

Treatment Tasks. Throughout one semester, the three experimental groups were provided with task-based instruction. They participated in a total of 12 tasks, which were organized according to three thematic

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

topics of tasks throughout one semester: *The World of Work*, *Travel*, and *University Life Events*. This study includes learners' task performance during four tasks in *The World of Work* unit: (a) finding a part-time job, (b) working as a matchmaker, (c) discussing promotion opportunities, and (d) hiring employees. Each topic was chosen on the basis of learners' needs and interests in both Korean and English.

Based on Robinson's (2001b, 2003, 2005) task-complexity framework, the complexity variable—[+/–reasoning demands] along resource-directing dimensions—was operationalized using three different levels: simple task [–reasoning], +complex task [+reasoning], and ++complex task [++reasoning]. To manipulate different degrees of task complexity between the +complex task and the ++complex tasks, the [+/–few elements] variable was used for controlling the number of factors learners needed to make a decision (see Kuiken & Vedder, 2002, 2007). Other than the task-complexity factor, task features in all task-based groups were held constant following Ellis's (2003) positive task characteristics for promoting interaction to minimize other task variable effects: (a) tasks that are relevant to learners' current or future needs either in Korean or English and (b) tasks that would naturally require information exchange as two-way tasks, which have been found to elicit significantly more conversational adjustments than one-way tasks (e.g., Gass & Varonis, 1985; Long, 1981; Pica, Young, & Doughty, 1987). Furthermore, these four tasks were designed as convergent tasks, which have been found to elicit more interactive features (e.g., more turns per task, more questions, and more confirmation checks) than divergent tasks (Duff, 1986; Pica & Doughty, 1985).

Table 1 provides more details of a sample treatment task manipulated along the dimension of task complexity. For instance, learners were provided with the same task input across the three groups. For simple tasks, learners were required to exchange necessary information. For +complex tasks and ++complex tasks, they were asked to make a decision. Between +complex tasks and ++complex tasks, the number of considerations for making a decision was the main factor in manipulating the degree of task complexity.

Tasks were piloted several times with different populations to ensure that the target learners would carry out the tasks as they were designed in the given time. All four tasks followed a similar pattern of task design, as illustrated in Table 1, to minimize the effects of any other task variables.

Traditional Instruction Materials. The syllabus of the comparison group (i.e., traditional instruction) was designed based on the information provided by the four instructors of the course. The same themes as in the experimental groups were used in the comparison group, and all activities were adapted from the syllabi used by the four instructors. For instance, listening and reading materials dealing with the characteristics of *The World of Work* were adapted from commercialized

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

Table 1. Descriptions of a treatment task ("University Students' Part-Time Job") by task complexity

	Simple [–reasoning]	+Complex [+reasoning]	++Complex [++reasoning]
Task input	Profile cards and students' part-time job descriptions	Profile cards and available part-time job descriptions	Profile cards and available part-time job descriptions
Task outcome	Prepare for a report on university students' part-time jobs	Decide appropriate part-time jobs for students	Decide appropriate part-time jobs for students
Procedure	1. Both learners had profiles of four different university students, including their part-time jobs. 2. Both learners were missing some information about each student's background. 3. **In order to complete a report of university students' part-time jobs, both learners collected information from their partners.**	1. Both learners had profiles of four different university students, including their part-time jobs. 2. Both learners were missing some information about each student's background. 3. **Based on each learner's background, learners were asked to suggest appropriate part-time jobs for each student.**	1. Both learners had profiles of four different university students, including their part-time jobs. 2. Both learners were missing some information about each student's background. 3. **Based on each learner's background, learners were asked to suggest appropriate part-time jobs for each student.**
		[+few elements] **Two considerations should be met for their decision.**	[–few elements] **Four considerations should be met for their decision.**

textbooks (e.g., Schoenberg, 2006). Different types of listening and reading comprehension questions such as fill-in-the-blanks, true-false questions, and vocabulary practice questions were used in the comparison group. Furthermore, one main difference between the task-based and comparison groups was related to learner-learner interaction. The members of the comparison group were not provided with any type of learner-learner interaction activities during their regular instruction. In addition to their regular instruction, the comparison group members participated in all of the pretests, posttests, and delayed posttests.

Testing Materials

Measure of General English Proficiency. To assess learners' general English proficiency, the TOEIC Bridge was administered on the second day of the course. The TOEIC Bridge is a short version of the TOEIC, which is an English proficiency test for nonnative speakers of English. The TOEIC Bridge was developed by the Educational Testing Service (ETS) for beginning- to high-intermediate-level learners of English. It measures the emerging receptive skills (i.e., listening and reading comprehension) of early learners in an international environment. The TOEIC Bridge is a paper-and-pencil, multiple-choice test that includes 50 questions in each of two sections: listening and reading. The test took about 1 hr to complete (25 min for the listening section and 35 min for the reading section). Test results were used to determine learners' general proficiency level and to ensure that the participants in the experimental and comparison groups were comparable at the start of data collection.

Measures of Question Development. To measure the use of question forms in a variety of communicative settings, three tests—two individual oral production tests and one paired oral production test—were designed. Three different versions of each test were designed for the pretest, posttest 1, and posttest 2. The first individual oral test provided six different cartoons with short descriptions and asked learners to create questions or make conversation on the basis of prompts (e.g., picture: Tom talking to a secretary at a gym; scenario given: Tom wants to join a fitness center. Tom is asking several questions about the gym.) Different scenarios were provided for three versions of the tests. The second individual oral test asked the learners to make at least five interview questions for four interview situations (e.g., an interview for a potential scholarship recipient to come to the United States). Different interview descriptions were provided for three versions of the tests (pretest, posttest 1, and posttest 2). For the paired oral production tasks, the learners were asked to select various topics (e.g., e-mail, school) and

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

interview one another in a natural communication setting. The learners were given 20 min for the two individual oral tests and 20 min for the paired tests. The three versions of the oral production tests were different from the treatment tasks to avoid problems associated with the use of identical treatment and testing materials.

Procedure

The present study was carried out over 6 weeks, as illustrated in Figure 1. In Week 1, the learners took the TOEIC Bridge and filled out a background questionnaire. The pretest was given during the second week, and interaction tasks were provided to task groups (i.e., simple, +complex, and ++complex) during Weeks 3 and 4, whereas the traditional group followed the regular syllabus adapted from other instructors. During task performance, learners were asked to choose their own partners and keep the same partners while carrying out the four tasks to avoid any interlocutor factors. Each task was carried out in 40–45 min.

The first posttest was conducted during Week 5, and the second posttest was carried out during Week 6, which was 2 weeks after instruction. Each posttest was composed of three test tasks (two individual test tasks and one paired test task) and took 40 min to complete.

Data Coding

Treatment Task Performance. Learners' recordings during treatment sessions were transcribed by the researcher and four research assistants.

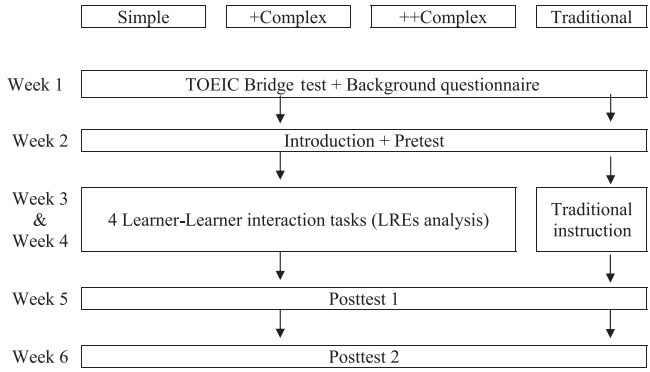


Figure 1. Procedure of the study.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

To identify the occurrence of learning opportunities during task-based learner-learner interaction, the number of LREs involving the target structures during the four tasks were identified and summed. In the current study, LREs are defined following Swain and Lapkin (1998) 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" (p. 326). Language-related episodes involving questions were defined as any instance of dialogue in which learners provided explicit and implicit feedback or asked explicit questions about question formation using either Korean or English. Depending on the target question form, more than one linguistic issue regarding question formation (e.g., *wh*-word, auxiliary choice) addressed during several turns could be included within one LRE, as illustrated in example (1). In example (1), Learner 1 asked a question regarding the *wh*-word choice (line 1). In line 4, Learner 2 provided feedback (i.e., a recast on Learner 1's wrong question formation). Then, Learner 2 also suggested the object of the verb *read* in a question. Overall, these two learners in example (1) took several turns to create one question, and these turns were counted as one LRE focusing on one target question.

(1) LRE involving questions
1 Learner 1: *Does he read . . . how long?* 얼마나가 *how long* 이야? (When asking for time length, do we use "how long"?)
2 Learner 2: 어 (*Yes*).
3 Learner 1: *How long read?*
4 Learner 2: *How long does he read?*
5 Learner 1: *How long does he read?*
6 Learner 2: *Read book?*
7 Learner 1: *Yes. How long does he read book?* ⟶ Stage 5
8 Learner 2: *He reads books for two hours.*

Once LREs involving questions were determined, the question that the learners had discussed in each LRE was coded for its developmental stage (see Appendix for question development stages). Although the total number of LREs involving questions provides useful information regarding the extent to which learners paid attention to form, not all LREs would necessarily trigger L2 development. For instance, when a learner's question stage was identified as Stage 3 on the pretest, producing LREs involving Stage 3 questions might not be as beneficial for L2 learning as those targeting developmentally advanced forms (e.g., Stage 4 or Stage 5). Therefore, producing Stage 4 and Stage 5 question types during LREs generates learning opportunities for developmentally advanced forms.[3] To provide further insight into learning opportunities for such forms, a subset of LREs involving Stage 4 and Stage 5 was considered separately for further analysis.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

When assigning stages, the last utterance in the LRE was considered. For instance, in example (1), after several turns, the last question form produced by Learner 1 was *How long does he read book*, and Stage 5 was assigned to this LRE. However, not all LREs involving questions were relevant to the question-stage assignment in the study (e.g., when learners asked for explicit questions regarding question formation without producing actual questions, or when there was no resolution of question-related LREs). The researcher coded all data, and a second rater coded data from 25% of the participants. The Cohen's kappa for the interrater reliability was .84 for identifying LREs involving questions.

Oral Production Tests. To assess performance, individual questions produced during each test task (i.e., pretest, posttest 1, and posttest 2) were coded according to the question development stages discussed previously. When identifying learners' developmental stages on each test, a conservative emergence criterion was used: Each learner was assigned to the highest level on the scale for which the learner produced two questions with unique lexical items for at least two out of three test tasks. This was similar to the criterion used in other SLA studies involving question development (e.g., Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006; Philp, 2003; Spada & Lightbown, 1993). The test data from 30% of the participants were coded by the second coder, and the Cohen's kappa for interrater reliability was .91 for question-stage assignment, and .96 for identifying stage increase or no stage increase from pretest to posttests.

Analysis

Task Performance. For research question 1, once the question-related LREs were coded, the number of LREs from the four treatment tasks was added, and the three task groups' total numbers were compared. To identify to what extent task complexity mediated the occurrence of developmentally advanced target features, LREs involving developmentally advanced questions were further analyzed in addition to LREs involving questions in general. For the occurrence of LREs, thus, two separate one-way ANOVAs were computed for research question 1 with LREs involving all stages of questions and with LREs involving developmentally advanced question forms (i.e., Stage 4 and Stage 5).

Oral Production Tests. After determining each learner's question stage in each of the tests (pretest, posttest 1, and posttest 2), learners were classified as either *developed* or *not developed*. Following previous interaction studies (e.g., McDonough, 2005), development was operationalized as

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

a learner's movement to a higher stage in the developmental sequence on both posttests.[4] Learners who did not advance to a higher stage on either posttest or advanced to a higher stage on only one posttest were classified as *no development*. To examine whether there was a main effect for task complexity, a chi-square test was conducted. In terms of identifying statistical significance, alpha was set at .05 for all statistical tests. Furthermore, for all statistically significant results, the effect size using partial η^2 was used.

RESULTS

Results of the Proficiency Test: TOEIC Bridge

To ensure the four groups were comparable in their general English proficiency, the TOEIC Bridge was administered. The maximum score on the TOEIC Bridge is 180: 90 for each of the two sections (listening and reading). Descriptive statistics for the TOEIC Bridge for each of the four classes participating are provided in Table 2.

Overall, the total TOEIC Bridge scores ranged from 64 to 164 (M = 115.66, SD = 19.26), which is comparable to the range of a paper-based TOEFL between 331 to 489, indicating that the participants' proficiency ranged from that of beginners to that of high intermediates. A one-way ANOVA was carried out to compare learners' English proficiency in the four groups. Results showed that there was no significant difference between the four classes, $F(3, 187) = 2.69$, $p > .05$, which suggested that the English proficiency level of each class was comparable.

Task Complexity and Learning Opportunities for Questions

Research question 1 investigated the role of task complexity in the occurrence of learning opportunities for question formations. Because

Table 2. Descriptive statistics for TOEIC Bridge

	Simple (n = 45)		+Complex (n = 47)		++Complex (n = 51)		Traditional (n = 48)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Listening	61.91	11.65	57.15	9.65	59.96	10.71	58.38	9.02
Reading	58.31	11.47	53.87	9.52	58.78	9.15	54.29	10.74
Total	120.22	21.86	111.02	17.48	118.82	19.09	112.58	18.52

Note. M = mean, SD = standard deviation.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

the comparison group did not participate in any interactive tasks, only results from the three experimental groups (i.e., task groups) are considered for this question. Two dependent variables were included. First, the total number of LREs involving questions during each of the four tasks was examined. Additionally, to gain further insight into the learning opportunities for developmentally advanced questions, LREs involving Stage 4 and Stage 5 were analyzed separately. Table 3 shows the descriptive statistics for the number of LREs.

Descriptive statistics for the occurrence of LREs indicated that the ++complex group (M = 57.55) generated the greatest number of question-related LREs. The +complex group (M = 50.09) produced more LREs involving questions compared to the simple group (M = 42.58). In terms of LREs involving developmentally advanced questions, as shown in Table 3, the ++complex group (M = 32.41) produced a larger number of LREs involving Stage 4 and Stage 5 than the simple group (M = 24.80) and the +complex group (M = 27. 21).

To answer the first research question, which focuses on the relationship between task complexity and the occurrence of learning opportunities, two separate one-way ANOVAs were conducted after checking the assumptions for using the ANOVAs. The results demonstrated that task complexity significantly affected the occurrence of LREs involving questions, $F(2, 140) = 14.35$, $p < .05$, partial $\eta^2 = .17$. The Bonferroni post hoc comparisons indicated that the ++complex group (M = 57.55) produced significantly more LREs involving questions than both the +complex (M = 50.09) and the simple (M = 42.85) groups, and the +complex group produced significantly more LREs than the simple group. Therefore, the findings of the post hoc analysis showed that the significant main effect for task complexity was due to the differences among all three levels of task complexity.

With regard to LREs involving developmentally advanced questions, the results of the one-way ANOVA revealed that there was a significant main effect for task complexity $F(2, 140) = 4.56$, $p < .05$, partial $\eta^2 = .06$.

Table 3. Descriptive statistics for LREs involving questions

	Simple (n = 45)		+Complex (n = 47)		++Complex (n = 51)	
	M	*SD*	*M*	*SD*	*M*	*SD*
All LREs involving questions	42.58	17.25	50.09	9.51	57.55	13.36
LREs involving advanced questions	24.80	14.73	27.21	9.21	32.41	13.48

Note. M = mean, *SD* = standard deviation.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

The Bonferroni post hoc comparisons pinpointed a significant difference between the ++complex (M = 32.41) group and the simple (M = 24.80) group, indicating the ++complex group had significantly more learning opportunities for developmentally advanced questions than the simple group. However, no difference was found between the +complex (M = 27.21) and the ++complex groups on the number of LREs with developmentally advanced question stages. Also, no significant difference was found between the simple group and the +complex group in the number of LREs involving Stage 4 and Stage 5 questions.

To summarize, results of research question 1 indicated that there was a significant main effect for task complexity on the learning opportunities for questions as measured by the number of LREs, with the ++complex group producing the greatest amount of LREs involving questions. Additionally, the +complex group had a greater number of LREs than the simple group. A similar pattern was found when focusing on the LREs that involved developmentally advanced questions. However, a significant difference was found between the ++complex and the simple groups only.

Benefits of Task-Based Instruction and the Role of Task Complexity in Question Development

The second research question investigated the relationship between instruction types and Korean EFL learners' question development. Additionally, the relationship between task complexity and question development among task-based groups was explored. To ensure there were no preexisting differences among groups in terms of learners' question developmental stage on the oral production test, learners' initial developmental stage as observed on the pretest was identified. As shown in Table 4, 131 of 191 students were determined as initially being in Stage 3, whereas 60 students were in Stage 4. Results of a chi-square test suggested that there was no significant relationship between the number of learners in each stage and task-complexity group, suggesting that the groups were comparable (χ^2 = 1.39, p > .05).

Table 4. Learners' initial question stages on the pretest

	Traditional (n = 48)	Simple (n = 45)	+Complex (n = 47)	++Complex (n = 51)
Stage 3	31 (65%)	29 (64%)	34 (72%)	37 (73%)
Stage 4	17 (35%)	16 (36%)	13 (28%)	14 (27%)

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

Table 5. Summary of question development by groups

	Traditional (n = 48)	Simple (n = 45)	+Complex (n = 47)	++Complex (n = 51)
No development	26 (54%)	15 (33%)	13 (28%)	9 (18%)
Development	22 (46%)	30 (67%)	34 (72%)	42 (82%)

Note. χ^2 = 15.78, df = (2, 191), p < .05.

Table 5 displays the number of learners in the traditional group and the three task groups who advanced to a higher stage or who failed to advance to a higher stage. As described previously, after determining each learner's question stages on each of the tests (pretest, posttest 1, and posttest 2), the learners were classified as either *developed* or *not developed*. Among the task-based groups, 30 (67%) out of 45 learners from the simple group, 34 (72%) out of 47 learners from the +complex group, and 42 (82%) out of 51 learners from the ++complex group advanced to a higher stage of question development. In contrast, only 22 (46%) out of 48 learners in the comparison group (i.e., traditional instruction group) advanced to a higher stage of question development.

The assumptions for using the chi-square test were met, and the results indicated that the relationship between the classes and English question development was significant (χ^2 = 15.78, p < .05). Analysis of the adjusted standardized residual (Haberman, 1973) located the significance with the performance of the ++complex group (±2.7) and the comparison group (±3.5). Overall, the results of research question 2 revealed that task-based language teaching, which provides ++complex tasks, promoted question development in comparison to traditional instruction.

DISCUSSION

The cognition hypothesis claims that increasing the complexity of tasks, which requires more cognitive-conceptual demands, induces more interaction, uptake of linguistic forms, and subsequent L2 development (Robinson, 2007b). The current study tested this claim by operationalizing learning opportunities as the number of LREs involving questions and operationalizing question development as question-stage advances based on Pienemann and Johnston (1987).[5] In terms of learning opportunities, the findings of the current study were in line with Robinson's claims that complex tasks induced significantly more LREs than less cognitively complex tasks. These findings echo those of previous classroom-based learner-learner interaction studies (e.g., Révész, 2011;

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

Robinson, 2007b), which showed that more complex tasks promoted the occurrence of beneficial interactional features such as LREs, clarification requests, and confirmation checks. One important methodological issue when examining the role of task complexity was to ensure that the target structure is related to the task demands of different levels of task complexity. Overall, question forms were appropriate target structures in that they were related to linguistic demands imposed on the learners who participated in more complex tasks in the current study (i.e., requiring learners to compare information while making a decision).

The simple group had the fewest learning opportunities involving questions, which was perhaps due to the fact that the nature of the interaction is different in simple versus complex tasks. Learners in the simple group took shorter turns and exchanged phrases by asking questions on the basis of information they were provided in their task input, rarely discussing issues on the task input further. To illustrate this observation, example (2) is taken from two learners in the simple group. In this example, two learners were exchanging information regarding university student profiles to complete a simple version of the task "University Students' Part-Time Job." Because the simple task only required students to exchange information without reasoning to complete the task, their question formations seemed to be limited to the amount of missing information.

(2) Two learners in the simple group during "University students' part-time job"
1 Learner 1: *What is her dream job in ten years?*
2 Learner 2: *She want to open her own hair salon.*
3 Learner 1: *Where she want to place?* ◄——— Stage 3
4 Learner 2: *He want in Pusan.*
5 Learner 1: *When she move?*

As shown in example (2), these two learners basically asked questions to complete the task without any LREs. For instance, when Learner 1 asked about a student's dream place to open a hair salon, she produced a Stage 3 question (line 3) without inversion. Neither Learner 1 nor Learner 2 paid attention to the question form at that time, and Learner 2 answered the question with the name of the location. It appeared that, because of the nature of the task (e.g., information gap), if there were no communication breakdowns, the learners in the simple group tended to pay less attention to question forms. Additionally, the question types and structures were rather limited to the information that they needed to collect to complete the task.

In contrast, as Robinson (2007b) claimed, more complex tasks generated more turns and consisted of a greater amount of negotiation and LREs. Furthermore, increasing the cognitive and conceptual demands

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

of the tasks led the learner to produce pushed output to meet those demands, causing reanalysis and restructuring of current linguistic resources. Example (3) illustrates two learners' interaction in the ++complex group during the same task. These two learners were discussing the same part of the task as the students in (2). For the decision-making requirement of the ++complex tasks, learners were asked to suggest an appropriate part-time job for the university students on the basis of their profiles provided in the task input.

(3) Two learners in the ++complex group during "University Students' Part-Time Job"
1 Learner 1: *What is her dream job?*
2 Learner 2: *She wants to have her hair salon.*
3 Learner 1: *What she wants?* ⟶ Stage 3
4 Learner 2: *What? Where . . Where she . . . no . . . Where does she want to open?*
5 Learner 1: *Yes. Where . . . Where . . . Where does she want?* ⟶ Stage 5
6 Learner 2: *In Pusan. Have you been Pusan?*
7 Learner 1: *Of course, I love Kwanghali. Let's go.* (laugh) ——— (several turns later)
8 Learner 2: *What part-time job for her?* ⟶ Stage 3 (without a verb phrase)
9 Learner 1: 무슨 말이야? (*What do you mean?*)
10 Learner 2: *What job is good for her major?* ⟶ Stage 4
11 Learner 1: *Work at a hair salon. Her major is Beauty Design. What do you think?*
12 Learner 2: *I agree*.

When Learner 1 asked the question, *What she wants* (line 3), Learner 2 tried to clarify the question by questioning the *wh*-word choice (line 4). After receiving Learner 2's feedback, Learner 1 restated her question with an inverted question (line 5). Also, the target question that triggered the LRE in line 4 was a Stage 5 question. It was interesting to see how both the learners in (2) and those in (3) commonly had difficulty in producing a developmentally advanced question form at the same point. The nature of the ++complex tasks entailed greater communicative demands because while suggesting an appropriate part-time job for students, learners were also required to ask more questions to compare information and get their partner's opinion, thus leading to increased comprehension and production difficulty and a correspondingly larger amount of noticing target structures. By resolving this LRE with a Stage 5 question, two learners in the ++complex group had the opportunity to construct a developmentally advanced question. Additionally, as they were trying to make a decision, they created original questions to review the task input, and those questions were often

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

developmentally advanced due to the nature of decision making (lines 10–11). Furthermore, learners who were in the +complex and ++complex groups used the task input to expand their conversation (e.g., asking personal experiences), which naturally generated more questions during their interactions (lines 6–7).

Overall, as Robinson (2007a) argued, the findings of the study suggested that complex task features manipulated by reasoning demands promoted learners' analysis of their current interlanguage system of the target features (line 4 in example [3]). Learners attempted to draw attention only to task-relevant constructions and those whose use would be affected by different degrees of task complexity. For instance, results of the current study indicate that reasoning demands provide more learning opportunities for asking questions during decision-making processes through language output and interaction (Long, 1996; Swain, 1995).[6] In particular, as shown in example (4), complex versions of tasks required learners to compare the information (line 3) and to make a decision (line 8). Such features of the complex versions of the tasks instigated more interaction-driven language learning opportunities because they created more communication breakdowns between learners due to the need to ask for more specific information (lines 5 and 9) (Gass, Mackey, & Ross-Feldman, 2005; Révész, 2011).

(4) Two learners in the ++complex group during "Working as a matchmaker"
1 Learner 1: *Minho Seo, what is his hobby?*
2 Learner 2: *Go swimming and watch movie.*
3 Learner 1: *Who who enjoy enjoys similar hobbies?*
4 Learner 2: *Similar hobby? Suha. What she like?*
5 Learner 1: *What?*
6 Learner 2: *What is Suha like?*
7 Learner 1: *Read book.*
——————————————————————— (several turns later)
8 Learner 1: 결정하자. (*Let's decide.*) *You think Suha want man who?* ⟶ Stage 2
9 Learner 2: 뭐? 이상해 who 로 시작해야지. (*What? [That sounds] strange. You should start with "who."*)
10 Learner 1: 어떤 남자를 원하냐고 어떻게 물어봐? (*How can I ask "what kind of man does she want?"*)
11 Learner 2: *What did she want?* 맞냐? *Is this right?* ⟶ Stage 5
12 Learner 1: *Oh . . . Baesuk Jang.*

Furthermore, while justifying and explaining their reasons for making different decisions, learners were pushed to produce more complex language to meet the greater functional and conceptual communicative demands. Although learners' pushed output may not necessarily lead to more accurate language use (see lines 6 and 11 in example [4]), it allowed learners to encounter linguistic problems, which, in turn, led

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

them to notice the gaps in their own interlanguage (Doughty & Williams, 1998; Swain, 1995).

The cognition hypothesis was generally confirmed for question development in that there was a significant relationship between task complexity and question development identified dichotomously following previous oral interaction studies: stage increase (i.e., developed) versus no stage increase (i.e., not developed) (e.g., Mackey, 2006; McDonough, 2005; McDonough & Mackey, 2006). The more complex tasks pushed learners to use more advanced forms because cognitively complex tasks may require the learners to ask various types of questions (Robinson, personal communication, March 22, 2009). For instance, as shown previously in (3) (lines 8–10) and (4) (lines 8–12), it was found that learners often attempted to ask more advanced questions when asking their partner to make a decision.

Results suggest that linguistic forms that were relevant to the cognitive-conceptual demands of the tasks received more attention during complex tasks than during simple tasks, which in turn led to linguistic development. Such pushed output with a target structure during task performance should contribute to L2 development, especially as measured by oral production (Swain, 1995, 1998).

Research question 2 investigated to what extent tasks with different degrees of task complexity promote Korean EFL learners' question development compared to the traditional instruction. Descriptive results demonstrated a clear trend that task groups outperformed the traditional instruction group in terms of question development. However, the statistical analysis indicated that a significant relationship was seen in the traditional instruction group and the ++complex group only. In other words, the results suggested that +complex and simple tasks did not significantly drive learners' question-stage increase.

It is important to notice that the current study was different from previous ones in that this investigation chose target structures (i.e., question formations) that are acquired by L2 learners following their developmental sequence proposed by Pienemann and Johnston (1987). This allowed more specific evidence of the relationship between task complexity, learning opportunities for developmentally advanced L2 structures, and subsequent L2 learning. For instance, as noted previously, although the +complex group produced significantly more LREs involving all question forms than the simple group, there was no difference in the number of LREs targeting developmentally advanced question forms (i.e., Stage 4 and Stage 5 questions) between the simple and the +complex groups. Based on these findings, it can be inferred that producing LREs involving developmentally advanced forms would directly drive L2 development, whereas LREs involving a lower stage than or the same stage as the learners' current stage may not have the same positive impact on L2 development.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

To date, very little research has investigated the direct relationship between task complexity and L2 development of specific linguistic constructions during learner-learner interaction. One example is Nuevo (2006), a study focusing on two linguistic features: past tense and locative pronouns. Her findings were not in line with Robinson's claims in that no differences were found between simple and complex task versions on learners' oral production test scores. As interaction researchers have addressed (Jeon, 2007; Williams & Evans, 1998), learner-learner interaction may be differentially effective in terms of the development of different areas of language. It may be, as Mackey (2006) suggests, that learners notice their interlanguage gaps with certain structures more than others. Therefore, the effects of tasks with various degrees of complexity may not be the same for different target structures. Perceptually salient linguistic structures for completing tasks—such as the questions in the current study—are likely to result in communication breakdowns and thus are more sensitively influenced by different levels of task complexity.

From a methodological perspective, the current study has provided information regarding the importance of treating task complexity along a continuum rather than as a dichotomous construct. This finding is supported especially with the learning opportunity data. For instance, although significant differences were found among the three groups when comparing LREs involving questions in general, LREs involving developmentally advanced questions were produced substantially more often by the ++complex group than the other two groups. Such findings suggest that only the ++complex tasks led to significantly more learning opportunities for developmentally advanced constructions during which learners actually produced the advanced question forms. As a result, among task groups, only the ++complex tasks were positively associated with L2 development.

The current study addresses several pedagogical implications for adapting task-based syllabi in large L2 classroom contexts. With higher ecological validity than studies in laboratory contexts (Gass et al., 2005), the findings suggest that task groups showed more positive learning outcomes compared to the traditional instruction group. A central concern for L2 language teachers is how to design tasks to help the learning and teaching of particular aspects of language. Even though some researchers argue that this should not be the principal contribution of tasks (e.g., Long, 1996; Skehan, 1998), most L2 practitioners would agree that a primary challenge is to design tasks that provide learners with opportunities both to engage in meaningful interaction and to direct their attention to linguistic forms (McDonough & Mackey, 2000). The current study took the position of directing learners' attention to linguistic forms. It also explored how task complexity would impact the occurrence of meaningful interaction, which entails learning

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

opportunities for linguistic structures as well as the learners' linguistic development.

The findings indicated that manipulating task complexity was found to be an important factor in promoting a greater amount of interactional features. As a result, findings suggest that task complexity is an important variable that can impact the amount of meaningful interaction as well as learners' attention to linguistic forms in task-based instruction contexts. However, to design such pedagogic tasks, one must be certain about functional and linguistic demands of tasks with different levels (e.g., what kinds of tasks would meet the needs of target learners, what kinds of task-complexity variables are appropriate, or what kinds of linguistic demands are related to task demands?). Thus, the vexing question is not just to determine what level of task complexity to create but to identify what task demands and linguistic demands should be targeted to promote L2 learning when designing tasks with various levels of complexity.[7] More specifically, carefully designed pedagogic tasks can influence the quantity and quality of learning opportunities for developmentally advanced linguistic forms that occur in L2 classrooms, in terms of both interaction and production practice.

There are several limitations to the current study that need to be addressed in future research. The first limitation is related to external validity. The study included relatively homogeneous participants in terms of their age, previous English instruction, and overseas living experiences to control for possible intervening variables other than the target variables of interest. However, due to the characteristics of participants included in the study, findings cannot be generalized to other populations or teaching contexts. Therefore, replication studies in a variety of contexts with different language learners are highly desirable. There is a need for more research to learn about the role of task complexity with different age groups and various target languages as well as pedagogical topics targeting different aspects of L2 learning.

The second limitation is related to the operationalization of learning opportunities and data coding. The current study took a global approach when determining learning opportunities by analyzing LREs as defined by Swain and Lapkin (1998). Most previous studies have operationalized learning opportunities with more specific analysis units by including different types of feedback, including clarification requests (e.g., Révész, 2011; Robinson, 2007b). Such studies found that task complexity impacted the occurrence of certain feedback types more than others. Because the current study focused on the resolution of LREs (i.e., question stages), a global approach was more appropriate. Due to the nature of the data analysis, these results cannot provide insights into what types of feedback learners offered to one another during LREs. Therefore, additional data analyses with more specific

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

feedback types would provide insightful information on the nature of learning opportunities.

The current study operationalized task complexity on the basis of one variable (i.e., +/–reasoning demands), and the validity of different degrees in task complexity was not tested. Future studies would benefit from validating different levels of task complexity, as proposed by Révész (2011). Additionally, this study manipulated only one task-complexity variable, with two types of tasks. To gain better insight and more pertinent data for the cognition hypothesis, there is a need for more experiments that test the hypothesis with a variety of task types that are manipulated by other variables in Robinson's framework (Kim, 2009a).

Furthermore, as Skehan (2003) addressed, an obviously relevant set of variables, which has received increasing attention in previous literature, concerns learner variables such as task motivation (Dörnyei & Tseng, 2009), anxiety, and working memory. What has been lacking in previous studies is systematic research that probes just how learner variables may impact the role of task complexity. Furthermore, as instructional settings have expanded due to the benefits of technology, not only face-to-face interaction but also other types of interaction settings such as computer-mediated communication contexts need to be explored further (Baralt, 2010). More research on this issue may be possible to increase the sensitivity of research designs and pedagogical implications in task-based language teaching.

CONCLUSION

Building off of previous research testing the cognition hypothesis, the inclusion of a specific target structure in the current study provided further insight into how different levels of learning opportunities mediated by task complexity may result in different degrees of L2 learning during meaning-oriented tasks. Results of this study suggest that, to maximize learning opportunities and L2 development in task-based language teaching contexts, variables that manipulate task complexity in Robinson's triadic framework need to be considered by L2 practitioners when designing tasks. In particular, as highlighted previously, when considering different levels of task complexity, the functional demands—and not just the cognitive demands—created by tasks with different task-complexity levels should be taken into consideration, particularly those related to the linguistic demands imposed on the learners. As Robinson (2011) suggests, more classroom-based research is needed to demonstrate the generalizability of the findings of laboratory studies for the role of task complexity on interaction-driven learning in classroom contexts with learners from diverse backgrounds. The current study tried to help fill this research gap, keeping in mind that

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

different task-complexity variables under Robinson's triadic components should continue to be examined to implement optimal L2 learning environments in task-based language teaching contexts.

(*Received 14 November 2011*)

NOTES

1. The four instructors are university faculty members who hold PhDs in English. Individual interviews were carried out to collect information about their courses. The results showed that no unified textbook was chosen among the four instructors, and their syllabi were designed based on topics and relevant vocabulary. The instructors mainly used both reading and listening materials that focused on improving vocabulary, pronunciation, and grammar. They provided different reading and listening materials adapted from newspapers, magazines, and movies and mainly focused on comprehension skills. Because of the large class size (approximately 55–60 students in one section), learner-learner interaction was not highlighted in their course plan.

2. See Gass and Selinker (2008) and Norris and Ortega (2003) for an extended discussion on how to measure L2 acquisition.

3. An anonymous *SSLA* reviewer pointed out that identifying Stage 4 and Stage 5 questions as developmentally advanced structures for learners whose question stage was assigned as Stage 4 on the pretest might not be appropriate. However, because the study included learners at both Stages 3 and 4 on the pretest, and both types of learners generated LREs in pairs, it was not possible to separate learners at Stage 3 and Stage 4 in the analysis. When interpreting the results, these issues need to be considered.

4. In the current study, question development was identified within a two-week timeline. Thus, to be considered as *developed*, learners must demonstrate stage increases on the immediate posttest and must maintain their stage increases on the delayed posttest.

5. An anonymous *SSLA* reviewer questioned whether task accomplishment was comparable in all three groups. Additionally, the reviewer suggested that how well each group accomplished tasks should be taken into consideration. The current study controlled for time on task in each group only. The actual quality or outcome of the task accomplishment, which is an important issue in task-based language teaching, needs to be addressed in future studies.

6. As an anonymous *SSLA* reviewer noted, other linguistic forms can be addressed due to reasoning demands (e.g., subordinate clauses, comparatives). The current study only focused on question forms due to the existing developmental sequence for questions to test Robinson's claim regarding promoting acquisition of developmentally advanced forms by increasing task complexity.

7. I appreciate the comments made by an anonymous *SSLA* reviewer, who pinpointed this issue.

REFERENCES

Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 29–51). Oxford: Oxford University Press.

Baralt, M. (2010). *Task complexity, the cognition hypothesis and interaction in CMC and FTF environments*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Bygate, M., Skehan, P., & Swain, M. (Eds.) (2001). *Researching pedagogic tasks, second language learning, teaching and testing*. London: Longman.

Dörnyei, Z., & Tseng, W. (2009). Motivational processing in interactional tasks. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 117–134). London: Routledge.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. New York: Cambridge University Press.
Duff, P. (1986). Another look at interlanguage talk: Taking task to task. In R. Day (Ed.), *Talking to learn* (pp. 147–181). Rowley, MA: Newbury House.
Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
Gass, S. M., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning*, *55*, 575–611.
Gass, S., & Selinker, G. (2008). *Second language acquisition: An introductory course* (3rd ed.). London: Routledge.
Gass, S. M., & Varonis, E. (1985). Task variation and non-native/non-native negotiation of meaning. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp.149–161). Rowley, MA: Newbury House.
Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching*, *45*, 215–240.
Gilabert, R., Barón, J., & Llanes, A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics in Language Teaching*, *47*, 367–395.
Haberman, S. (1973). The analysis of residuals in cross-classified tables. *Biometrics*, *29*, 205–220.
Ishikawa, T. (2007). The effect of manipulating task complexity along the +/– dimension on L2 written narrative discourse. In M. P. Garcia Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–176). Bristol, UK: Multilingual Matters.
Jeon, K. S. (2007). Interaction-driven L2 learning: Characterizing linguistic development. In A. Mackey (Ed.), *Interactional feedback and second language acquisition* (pp. 379–403). Oxford: Oxford University Press.
Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *Modern Language Journal*, *92*, 114–130.
Kim, Y. (2009a). The effects of task complexity on learner-learner interaction. *System*, *37*, 254–268.
Kim, Y. (2009b). *The role of task complexity and pair grouping on the occurrence of learning opportunities and L2 development*. Unpublished doctoral dissertation, Northern Arizona University, Flagstaff.
Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. Foster-Cohen, M. P. García-Mayo, & J. Cenoz (Eds.), *Eurosla Yearbook* (Vol. 5, pp. 195–222). Amsterdam: Benjamins.
Kuiken, F., & Vedder, I. (2002). The effect of interaction in acquiring the grammar of a second language. *International Journal of Educational Research*, *37*, 343–358.
Kuiken, F., & Vedder, I. (2007). Cognitive task complexity and linguistic performance in French L2 writing. In M. P. García-Mayo (Ed.), *Investigating tasks in formal language settings* (pp. 117–135). Bristol, UK: Multilingual Matters.
Lee, Y. (2002). *Effects of task complexity on the complexity and accuracy of oral production in L2 Korean*. Unpublished doctoral dissertation, University of Hawai'i, Manoa.
Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 361–378). Oxford: Oxford University Press.
Long, M. (1981). Input, interaction and second language acquisition. In H. Winitz (Ed.), *Annals of the New York Academy of Sciences conference on native and foreign language acquisition* (pp. 159–278). New York: New York Academy of Sciences.
Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of research on language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, *21*, 557–587.
Mackey, A. (2006). Feedback, noticing and second language development: An empirical study of L2 classroom interaction. *Applied Linguistics*, *27*, 379–394.
Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, *82*, 338–356.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

McDonough, K. (2005). Identifying the contributions of negative feedback and learners' responses to L2 development. *Studies in Second Language Acquisition*, *27*, 79–103.
McDonough, K., & Mackey, A. (2000). Form and meaning: Designing communicative tasks to target grammar in Thai classrooms. *Foreign Language Annals*, *33*, 82–92.
McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, *56*, 693–720.
McDonough, K., & Sunitham, W. (2009). Collaborative dialogue between Thai EFL learners during self-access computer activities. *TESOL Quarterly*, *43*, 231–254.
Michel, M., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, *45*, 241–259.
Norris, J., & Ortega, L. (2003). Defining and measuring SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 717–761). Oxford: Blackwell.
Nuevo, A. (2006). *Task complexity and interaction: L2 learning opportunities and development*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.
Philp, J. (2003). Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, *25*, 99–126.
Philp, J., Oliver, R., & Mackey, A. (2006). The impact of planning time on children's task-based interactions. *System*, *34*, 547–565.
Pica, T., & Doughty, C. (1985). The role of group work in classroom second language acquisition. *Studies in Second Language Acquisition*, *7*, 233–248.
Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, *21*, 737–758.
Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45–141). Adelaide, Australia: National Curriculum Resource Centre.
Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, *31*, 437–470.
Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *Modern Language Journal*, *95*, 162–181.
Robinson, P. (2001a). Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). New York: Cambridge University Press.
Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*, 27–57.
Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies*, *21*, 45–105.
Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, *43*, 1–32.
Robinson, P. (2007a). Criteria for classifying and sequencing pedagogic tasks. In M. P. Garcia-Mayo (Ed.), *Investigating tasks in formal language settings* (pp. 7–26). Bristol, UK: Multilingual Matters.
Robinson, P. (2007b). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching*, *45*, 193–213.
Robinson, P. (2009). Syllabus design. In M. H. Long & C. J. Doughty (Eds.), *Handbook of second and foreign language teaching* (pp. 294–310). Oxford: Blackwell.
Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning*, *61*, 1–36.
Robinson, P., & Ellis, N. (2008). Conclusion: Cognitive linguistic, second language acquisition and L2 instruction-issues for research. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 489–545). London: Routledge.
Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching*, *45*, 161–176.
Schoenberg, I. E. (2006). *Focus on grammar: An integrated skills approach* (3rd ed.). London: Longman.

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
Skehan, P. (2003). Task-based language teaching. *Language Teaching*, *36*, 1–14.
Skehan, P., & Foster, P. (1999). The influence of structure and processing conditions on narrative retellings. *Language Learning*, *49*, 93–120.
Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183–205). New York: Cambridge University Press.
Spada, N., & Lightbown, P. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, *15*, 205–224.
Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). New York: Cambridge University Press.
Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, *82*, 320–337.
Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139–155). New York: Cambridge University Press.

APPENDIX

DEVELOPMENTAL STAGES OF QUESTIONS

Developmental Stage	Examples
Stage 1: Single words, formulas, or sentence fragments	Your major?
Stage 2: Declarative word order, no inversion, no fronting	You have blue jeans? Your cell phone what color?
Stage 3: Fronting (do-fronting, *wh*-fronting, no inversion)	Do you have other experience? How much you drink?
Stage 4: *Wh*+copular, yes-no questions with other auxiliaries, yes-no copula	What's your favorite movie star? Can you give a hint?
Stage 5: Inversion in *wh*-questions with both an auxiliary and a main verb	What's she eating? What degree do you have? How long did you go here?
Stage 6: Complex questions: tag questions, negative questions, embedded questions	You like pizza, don't you? Don't you have pet? Can you tell me where she study?

https://doi.org/10.1017/S0272263112000368 Published online by Cambridge University Press