
Task Complexity, Focus on L2 Constructions, and Individual Differences: A Classroom-Based Study

ANDREA RÉVÉSZ

Lancaster University

Linguistics and English Language

County South C70

Lancaster LA1 4YT

United Kingdom

Email: a.revesz@lancaster.ac.uk

Motivated by cognitive-interactionist frameworks for task-based learning, this study explores whether task complexity affects the extent to which learners focus on form–meaning connections during task-based work in a classroom setting, and whether this relationship is modulated by 3 individual difference factors—linguistic self-confidence, anxiety, and self-perceived communicative competence. Forty-three English as a second language learners from 6 intact classes worked in self-selected groups during their normal English classes. Each group performed 2 versions of the same argumentative task—a simple version and a complex version. The topics of the discussions were comparable, and the sequence of the tasks was counterbalanced. Twenty-three hours of audio- and videotaped data were collected and coded in terms of global and specific measures of speech production and various interactional features hypothesized to facilitate attention to second language constructions. Self-report questionnaires were employed to determine the mediating effects of the individual difference variables. Quantitative analyses revealed a few significant trends. One such trend is that participants demonstrated lower syntactic complexity but greater accuracy and lexical diversity when task complexity was increased. In addition, the more complex task proved more effective in inducing the use of specific, developmentally advanced constructions and in promoting interaction-driven language learning opportunities. However, no significant effects were observed for individual differences.

THE FIELD OF INSTRUCTED SECOND language acquisition (SLA) research has seen an enormous growth of interest in the construct of task in recent years (e.g., Bygate, Skehan, & Swain, 2001; Ellis, 2003; García Mayo, 2007; Long & Crookes, 1992; Robinson & Gilabert, 2007; Samuda & Bygate, 2008). This increased interest has largely been motivated by the growing consensus among SLA researchers that, if the goal of second language teaching is to achieve advanced second language (L2)

capacities, instruction is most effective when it is primarily meaning-based but, at times, allows for the processing of L2 constructions or form–meaning mappings (Doughty & Williams, 1998; Lightbown & Spada, 1990; Long & Robinson, 1998; Robinson & Ellis, 2008). Tasks, which are defined as activities that require “learners to use language, with emphasis on meaning, to attain an objective” (Bygate et al., p. 38), appear to be ideal pedagogical tools in implementing instructional approaches based on this understanding. They have the potential to generate plentiful opportunities for meaningful language use while providing a framework for promoting learner attention to L2 constructions.

Cognitive-interactionist L2 researchers have made several proposals on ways to integrate a focus on form–meaning connections into task

performance. One approach that has attracted a considerable amount of interest from task-based researchers involves proactively manipulating the design characteristics of tasks. Among other things, it has been argued that task design manipulations may, for one, induce learners to place enhanced attention on the quality of their linguistic output (e.g., Skehan, 1998; Robinson, 2001a; Swain, 1995) and, for another, generate interactional features, which are hypothesized to promote the acquisition of form–meaning relationships (e.g., Gass, Mackey, & Ross-Feldman, 2005; Long, 2007; Robinson, 2001a, 2007). Although the impact of a relatively large number of task design features has been examined to test the outcomes of one or the other of these processes, few empirical studies (e.g., Robinson, 2001b, 2007) exist that have investigated task characteristics in relation to both simultaneously, and even fewer have done so in a classroom setting. One goal of the present study, therefore, is to help fill this gap by exploring the effects of one element of task design—task complexity—on the quality of learner linguistic output, as well as the number and type of interaction-driven learning opportunities arising during task-based work in the classroom.

Another aim of this study is to address whether and, if so, how individual difference (ID) variables might modulate the link between task manipulations and the extent of learner attention to form-function mappings during task performance. Thus far, only a few ID variables have been examined with respect to this relationship, including aptitude, working memory, intelligence, and anxiety (Niwa, 2000; Robinson, 2007). In order to extend this line of research, the present study focuses on 3 learner factors—self-perceived communicative competence (McCroskey & McCroskey, 1988), linguistic self-confidence (Dörnyei & Kormos, 2000), and language-use anxiety (Dörnyei & Kormos, 2000; Robinson, 2007).

TASK COMPLEXITY, FOCUS ON FORM–MEANING MAPPINGS, AND L2 PRODUCTION

Task complexity refers to the “attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner” (Robinson, 2001a, p. 29).¹ There are two competing accounts of how task complexity may affect attentional allocation and, correspondingly, the extent of focus on L2 constructions during task performance. Skehan (1998, 2007; Skehan & Foster, 2001) presumes, in

what he terms the “Trade-Off Hypothesis,” that humans have limited attentional capacity, and thus, for L2 learners, there is a constant competition for attentional resources between content and language. When learners’ attentional limits are reached, and when resources can be allocated freely, content is likely to be prioritized over language. Therefore, more complex tasks, which demand more attention for message conveyance, will allow learners to devote less attention to the linguistic aspects of their performance. Conversely, decreasing task demands will result in less attentional resources needed for content, and hence more attention available for linguistic features.

Alternatively, Robinson (2001a, 2003), following Wickens (1992, 2007), puts less emphasis on capacity constraints and distinguishes between separate resource pools along three dichotomous dimensions: processing stages (perception vs. response), modality (auditory perception/vocal response vs. visual perception/manual response), and codes of processing (verbal vs. spatial), each being responsible for a different aspect of task performance. In this view, the relative ease or difficulty of a task depends on the interference or competition within these resource pools during task work. As Wickens (2007) explains, the more levels that two tasks share along these three dimensions, the higher interference will develop—which, in turn, will negatively affect task performance.

In addition, Robinson (2001a, 2005) divides task variables into two groups: resource-dispersing and resource-directing dimensions. Increasing task complexity along resource-dispersing dimensions, for instance, by requiring learners to perform more than one task simultaneously, simulates real-world performance conditions and, in so doing, disperses attention over many non-specific areas of the L2. On the other hand, making tasks more complex along resource-directing dimensions, for example, by requiring learners to reason about and/or refer to many elements instead of a few, increases the conceptual/cognitive demands of the task and, in turn, directs learners’ attention to specific L2 constructions in the L2 relevant to the appropriate conceptual/cognitive domain. Specifically, tasks requiring reasoning will probably induce the use of logical connectors and subordinate clauses, whereas tasks with many elements might lead to greater reliance on complex noun phrases, relative clauses, and demonstrative pronouns.

Against these backgrounds, Robinson and Skehan make a series of predictions as to how changes in task complexity will affect the linguistic

aspects of L2 output or, to be precise, the accuracy (i.e., correct use of the L2) and/or complexity (i.e., use of advanced and elaborate interlanguage constructions) of production (Skehan & Foster, 2001). The Trade-Off Hypothesis predicts that increasing the cognitive demands of tasks will negatively affect both accuracy and linguistic complexity. A further trade-off between accuracy and complexity is expected, given that learners will only be able to pay attention to one of these areas at a time. Just as Skehan does, Robinson hypothesizes, within the scope of what he calls the "Cognition Hypothesis," that more complex tasks along resource-dispersing dimensions will degrade both accuracy and linguistic complexity. These putative effects, however, are attributed to increased demands on task-switching and scheduling mechanisms caused by interference. The two frameworks also differ in that the Cognition Hypothesis assumes no competition between accuracy and complexity, since these two aspects of production are closely connected. For resource-directing dimensions, in line with insights drawn from cognitive linguistics (Talmy, 2000) and first language (L1) acquisition (Cromer, 1974), Robinson predicts that more complex monologic tasks will lead to both greater accuracy and complexity. More complex interactive tasks, however, while also expected to lead to improved accuracy and lexical complexity, are presumed to result in less syntactically complex language, owing to the many phrasal and one-word responses that tend to occur during negotiation work (Robinson, 2005).

Stimulated by these theoretical frameworks, a substantial number of empirical studies have been directed at exploring the link between task design variables and the accuracy and linguistic complexity of L2 output (see Ellis, 2003; Samuda & Bygate, 2008 for recent reviews). Of particular relevance here is previous work investigating the task dimensions +/- many elements and +/- reasoning demands, given that the tasks in the current study were manipulated along these design characteristics. Kuiken and Vedder, as in the study reported here, set out to gauge the combined effects of these task dimensions on L2 performance, in a series of studies examining written production (Kuiken, Mos, & Vedder, 2005; Kuiken & Vedder, 2007). The general pattern emerging from their work is that increased task complexity can lead to improved accuracy (fewer errors per t-unit) and lexical diversity (higher uncorrected/corrected type-token ratios), but it is unlikely to affect syntactic complexity (clauses per t-unit and dependent clauses per clause).

These trends were partly confirmed by Michel, Kuiken, and Vedder (2007) and Robinson

(2001b) for oral production in relation to task manipulations along the +/- many elements dimension. In Michel et al.'s study, accuracy was measured by errors per AS-unit, lexical errors and omissions per AS-unit, and ratios of self-repair. Guiraud's Index, a type-token ratio corrected for text length, and rate of lexical words per number of words, were computed to measure lexical diversity. As in Kuiken and Vedder's previous work, more complex tasks, overall, yielded more accurate and lexically diverse language. Similarly, in Robinson's study (2001b), the task with more elements led to greater type-token ratio than its counterpart with fewer elements. Task complexity, however, had no impact on the rate of error-free c-units, which was used to gauge accuracy. Importantly, neither study revealed an effect for syntactic complexity, which was measured by subordination indices.

In contrast, Niwa (2000) and Robinson (2007) did find an effect on syntactic complexity for the task feature +/- reasoning demands. In Niwa's study, accuracy was expressed as a ratio of error-free t-units, and syntactic complexity as a ratio of clauses per t-unit. The study yielded a positive effect for task complexity in terms of syntactic complexity but found no link between task complexity and accuracy. In Robinson's study, accuracy was measured by error-free c-units, and measures of syntactic complexity included clauses per c-unit, words per turn, and use of terms describing psychological states and the extent to which they were associated with the incidence of infinitival phrases, conjoined clauses, and *wh*-clauses. Robinson detected a positive impact of task complexity for two of the syntactic complexity indices, words per turn, and terms describing psychological states, but not for accuracy.

In sum, existing empirical research suggests that increasing the reasoning demands and/or number of elements in a task can induce improved linguistic performance during task-based work, as predicted by the Cognition Hypothesis. The specific findings regarding accuracy and syntactic complexity, however, appear more congruent with the Trade-Off Hypothesis, given that most studies detected an impact for only one of these performance dimensions.

One possible explanation for the mixed findings may be that, except for Robinson (2007), the studies reviewed exclusively relied on global measures of accuracy and complexity. It has been argued, however, that it would be desirable to supplement general indices with specific, task-related measures because these, being more closely connected to the particular task design variable examined, may be more likely to capture relevant

performance effects (Bygate & Samuda, 2005; Mochizuki & Ortega, 2008; Norris & Ortega, 2008; Robinson, 2007). The use of specific, developmentally motivated measures is especially critical, Robinson explains, when investigating the predictions of the Cognition Hypothesis for resource-directing dimensions because these task features are, by definition, linked to certain conceptual/linguistic domains.² Indeed, in Robinson's (2007) study, the global measures proved less successful in detecting task effects than did the specific measures (terms describing psychological states). In view of these theoretical and empirically supported insights, the present study employs both global and task-specific measures to explore the link between task complexity and the quality of L2 production.

TASK COMPLEXITY, FOCUS ON FORM–MEANING MAPPINGS, AND INTERACTION

Besides exploring the effects of task features on L2 output, SLA researchers (e.g., Pica, Kanagy, & Falodun, 1993) have devoted a considerable amount of attention to identifying task design characteristics that may induce interactional elements hypothesized to promote the now-documented link between interaction and second language development (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006; Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007). Initially, inspired by the early version of the Interaction Hypothesis (Long, 1983), researchers were mainly concerned with isolating task variables that can promote interactional features such as clarification requests, comprehension checks, confirmation checks, and recasts. These strategies, back then, were mainly defined as constituting negotiation for meaning, the process in which interlocutors detect and, in response, attempt to resolve a communication problem. More recently, in line with current understandings of the acquisitional benefits of participating in conversational interaction (Gass, 1997; Long, 1996, 2007; Pica, 1994; Swain, 1995, 2005), the research focus has been extended to discerning task variables that can promote interaction-driven opportunities to attend to L2 constructions, regardless of whether communication breakdown occurs. These occasions when learners attempt to draw attention to L2 form–meaning connections in the context of meaning-based communication have either been termed as language-related episodes (LRE) (Swain & Lapkin, 1998, 2001) or focus-on-form episodes (Ellis, Basturkmen, &

Loewen, 2001). By now, a large array of task variables has been discovered that can impact the quantity and quality of negotiation and/or LREs arising during task-based interaction (see Ellis, 2003; Samuda & Bygate, 2008, for reviews). However, relatively little empirical research has considered the effects of task complexity on these interactional features.

The communicative consequences and attentional requirements of increasing task complexity, however, seem to have important implications for the number and type of interaction-driven language learning opportunities occurring during task work. First, as Robinson (2001a) explains, more cognitively complex interactive tasks (along both resource-directing and resource-dispersing dimensions) are likely to entail greater communicative demands, thus leading to increased comprehension difficulty and a correspondingly larger amount of negotiation for meaning. Second, it would appear probable that, in addition to negotiation sequences, manipulating task complexity will have an impact on the quantity of language-related episodes initiated by purely linguistic needs, in the absence of meaning negotiation. In particular, task complexity manipulations that can prompt increased attention to L2 form–meaning mappings (Robinson, 2003; Skehan, 1998) might more likely be associated with the incidence of such LREs. When learners' attention is more oriented towards L2 constructions, they will probably be more inclined to make an attempt at drawing their interlocutors' focus of attention to linguistic features even when understanding between them is not an issue.

Following this line of thought, Skehan's and Robinson's frameworks seem to carry distinct implications as to how task complexity will affect the incidence of LREs, triggered solely by code-related issues, when no problems in communication are perceived among interlocutors. The Trade-Off Hypothesis seems to suggest that more complex tasks will induce fewer opportunities for such LREs, as the increased task demands will require more attentional resources for content and hence, less available for language. In terms of Robinson's model, the prediction for the observable effects of increasing task complexity along resource-dispersing dimensions, albeit via a different underlying logic (see previous section), seems to be the same. In contrast, when task complexity is increased along resource-directing dimensions, there might be a higher likelihood that learners attempt to draw attention to task-relevant constructions, given that these task manipulations are

presumed to direct attention to task-relevant linguistic elements.

It is important to emphasize that these insights are derived from, rather than explicitly stated by, the two frameworks. Skehan does not seem to make overt predictions for the relationships among task complexity, attentional allocation, and interactional features. Although Robinson is concerned with this link, he seems to suggest that more complex tasks, due to their increased communicative demands, will generally lead to more LREs. That is, he does not make distinct predictions for the communicative and attentional correlates of enhancing task complexity, hence the potentially different effects of greater task demands on LREs prompted by message incomprehensibility versus those triggered by code-related problems in the absence of breaks in communication.

Investigating whether task complexity indeed differentially affects these two types of LREs and if so, how, is not only of theoretical relevance, but also of pedagogical relevance. As several researchers have noted, errors in communicatively redundant, perceptually non-salient L2 constructions, which are notoriously difficult for learners to acquire on their own (Sato, 1986), rarely cause communication failure (Foster & Ohta, 2005; Han, 2007). Hence, it would appear that task features with a capacity to promote the incidence of LREs unaccompanied by comprehension problems may have an especially great value in providing opportunities for improving grammatical competence because such LREs, being solely triggered by problems in the code, might be more likely to focus attention on L2 constructions, which are non-meaning bearing and non-salient and thus, less noticeable.

To date, few empirical studies exist that can provide information regarding the putative relationship between task complexity and the incidence of LREs. Among them are Robinson's previously cited studies (Robinson, 2001b, 2007), which also examine whether increasing the cognitive demands of tasks affects the extent of participants' tendency to interact and negotiate. In counting clarification requests and confirmation checks, Robinson observed significantly more confirmation checks on the more complex task versions in both experiments. Although the results showed trends for clarification requests in the same direction, they reached significance only in Robinson's study (2007).

Similar to Robinson (2007), Nuevo (2006) focused on the task dimension +/- reasoning demands, and investigated its effects on interaction-

driven learning opportunities in relation to two task types—a narrative task and a decision-making task. The interactional elements that Nuevo considered were recasts, clarification requests, confirmation checks, comprehension checks, hypothesis formulation, metalinguistic talk, self-repair, noticing of linguistic deficiency, and other repetition. The study revealed greater uptake of recasts, more comprehension checks, and more other repetitions on low complexity task versions than high ones. The less complex version of the decision-making task also led to a greater amount of metalinguistic talk. During the same task, however, more confirmation checks were detected on the more complex version. Likewise, learners appeared to test hypotheses to a greater degree when engaged in the more complex version of the narrative task. On the rest of the interactional measures, Nuevo detected no significant effects. In sum, the study failed to reveal any clear-cut patterns.

Gilabert, Barón, and Llanes (2009) generated somewhat clearer trends in examining the effects of task complexity on interactional patterns. In this study, participants carried out a narrative reconstruction task, an instruction-giving map task, and a decision-making task, which were manipulated along the degree of past time reference required, the number of elements contained, and the reasoning demands posed, respectively. The data were coded for negotiation of meaning (confirmation checks, clarification requests, and comprehension checks); recasts; LREs (excluding negotiation moves); and self-repairs. Gilabert et al. found that increased task complexity resulted in greater incidence of all interactional features except for recasts on both the narrative reconstruction task and the instruction-giving task, but not on the decision-making task. Interestingly, however, higher complexity led to more self-repairs on each task type.

Finally, Kim (2009) examined whether task complexity affects the occurrence of LREs by learners at different proficiency levels. Participants carried out two task types, a picture narration and a picture difference task, which were manipulated in terms of the +/- number of elements and +/- reasoning demands dimensions. On the picture narration task, the low-proficiency group produced more LREs when working on the simple version, whereas the high-proficiency learners engaged in a larger number of LREs during the complex version. For the picture difference task, the complex version generated more LREs by the low-proficiency group, but no significant link was observed between task complexity and

number of LREs for the high-proficiency group. In sum, Kim's results suggest that the effects of task complexity may not only differ across task types but also across proficiency levels.

Overall, the small number of existing empirical studies suggests that task complexity can affect interactional patterns occurring during task-based interaction. The results as yet, however, appear to supply limited information about the exact nature of its influence. Among other things, no direct evidence is available regarding the effects of task complexity on LREs with meaning negotiation versus without. None of the previously discussed studies (Gilabert et al., 2009; Kim, 2009; Nuevo, 2006; Robinson, 2001b, 2007) distinguished LREs in terms of this distinction. Another goal of this study, therefore, was to further elucidate the relationship between task complexity and interaction-driven language learning opportunities, while distinguishing LREs spawned by communication breakdowns and those that occur in their absence.

TASK COMPLEXITY, FOCUS ON FORM–MEANING MAPPINGS, AND INDIVIDUAL DIFFERENCES

Last but not least, this study set out to examine whether the extent of learners' focus on L2 constructions during simple tasks versus complex ones would vary depending on ID variables. The effects of individual factors on second learning processes and outcomes have been extensively researched for the past three decades (for an overview, see Dörnyei, 2005). Although there is a growing amount of evidence that IDs can affect second language learning in a broad sense, relatively little is known about the interaction between learner factors and specific L2 instructional events (Dörnyei, 2002). More recently, however, research attention began to focus on these relationships by adopting a micro (instead of a macro) perspective, and researchers started to examine learner factors in relation to L2 task-based production (e.g., Dörnyei, 2002; Dörnyei & Kormos, 2000; Kormos & Dörnyei, 2004; Niwa, 2000; Robinson, 2007) and interactional features (Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Mackey, Adams, Stafford, & Winke, 2010; Mackey & Sachs, 2011; Révész, *in press*; Sagarra, 2007; Sheen, 2007, 2008; Trofimovich, Ammar, & Gatbonton, 2007).

In order to continue this line of inquiry, the present study investigates the extent to which three affective variables—L2 use anxiety (Dörnyei & Kormos, 2000), linguis-

tic self-confidence (Dörnyei & Kormos, 2000), and self-perceived communicative competence (McCroskey & McCroskey, 1988)—modulate the link between task complexity and focus on L2 form–meaning mappings. Three previous studies are directly relevant to this research. Kormos and Dörnyei (2004) explored the impact of linguistic self-confidence and anxiety, along with other ID factors, on the quantity and quality of linguistic output during task performance. They employed a variety of linguistic measures: speech size, number of turns, accuracy, syntactic complexity, lexical richness, and number of arguments and counter-arguments. For anxiety and self-confidence, they found only two significant correlations: learners with lower anxiety used a wider array of lexical items, and those with greater self-confidence produced a larger number of words. Sheen (2008) also reported benefits for learners with low classroom anxiety in a study investigating individual differences in relation to the efficacy of recasts. These two studies, respectively, suggest that the quantity and quality of linguistic output and the extent of interaction-driven language learning can be influenced by ID variables. They provide no insights, however, as to whether the magnitude of these associations may vary depending on the cognitive complexity of the task in which learners are engaged.

Robinson (2007), besides investigating the effects of task complexity on speech production and interaction-driven language learning opportunities, addressed the link between task complexity and IDs. The Cognition Hypothesis predicts that “individual differences in ability and affective factors relevant to the cognitive demands of tasks will increasingly differentiate learners' . . . interaction and uptake, as tasks increase in complexity” (p. 196). Robinson's results partly confirm this prediction: learners with low output anxiety produced more syntactically complex speech on more cognitively complex tasks, whereas task complexity had no impact on the syntactic complexity of high anxiety learners' linguistic output. Robinson, however, found no effects for anxiety with respect to accuracy. Clearly, further research is needed to confirm and determine the extent and nature of the interaction between learner factors and task complexity manipulations.

RESEARCH QUESTIONS AND HYPOTHESES

1. Does task complexity affect the accuracy and complexity of L2 speech production? Following the Trade-Off and the Cognition Hypotheses, it was predicted that task complexity would

influence the accuracy and linguistic complexity of learners' output. The directions of the relevant hypotheses were not specified, given that the two theoretical models hold contradictory predictions for task dimensions that Robinson terms resource-directing.

2. Does task complexity affect the quantity and quality of interaction-driven learning opportunities occurring during task work? In line with the Cognition Hypothesis, it was expected that learners would engage in a greater number of LREs triggered by communication breakdowns when task complexity was increased. Additionally, based on insights derived from the Trade-Off and the Cognition Hypotheses, task complexity was hypothesized to modulate the number of LREs arising in the absence of communication problems. Again, the direction of the prediction was not specified, given the conflicting prognoses that the two frameworks appear to generate for the task manipulation in focus.

3. To what extent do individual differences in self-perceived communicative competence, linguistic self-confidence, and language-use anxiety moderate these relationships? Motivated by the Cognition Hypothesis, it was predicted that these ID factors would moderate the quality of linguistic output and quantity of interaction-driven language learning opportunities to a greater extent on more complex tasks.

METHOD

Participants

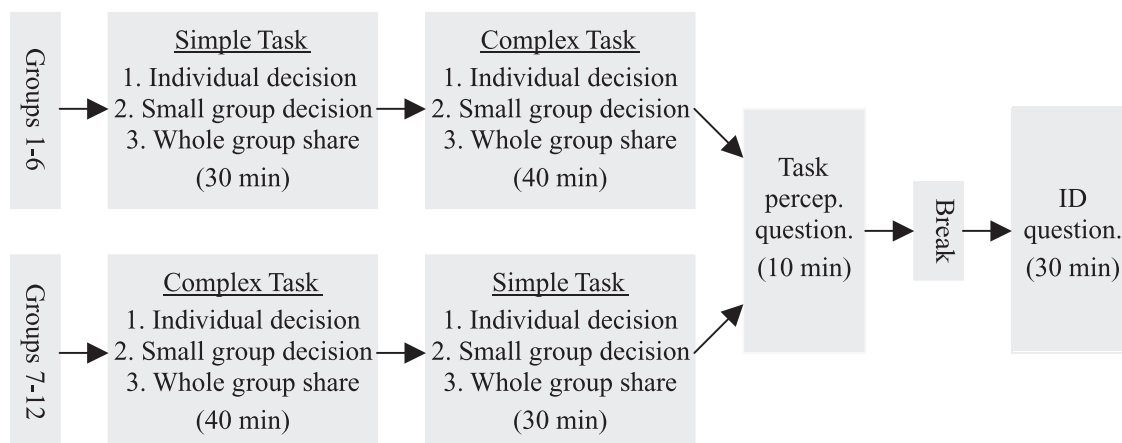
The participants in the study were 43 English as a second language (ESL) learners who were enrolled in intensive English courses in the Commu-

nity English Program of a university in New York City. Based on their scores on the program's placement test, the learners were placed in classes of high-intermediate to advanced levels. Altogether, 6 intact classes took part in the study. They each met 3–4 times a week for a period of 2.5 hours. At the time of data gathering, the classes had been running for 3–5 weeks. As for the students' background information, most of the learners were females; only 6 male students participated. The learners ranged in age from 21 to 45, with a mean of 32. Although the participants came from a variety of L1 backgrounds, the majority were native speakers of Spanish, Korean, or Japanese. Their average length of residence in the United States was 11 months, with a range from 3 months to 6 years. Overall, the makeup of the classes was typical of community English programs in the New York City metropolitan area. The participating instructors were the regular ESL teachers of the 6 classes.

Design

As part of the experiment, the students worked in self-selected groups of 3 or 4 during their normally scheduled English classes. Altogether, data from 12 groups (2 groups per class³) were included in the study. Each group worked on two versions of the same interactive task—a simple version and a complex version. As shown in Figure 1, these were counterbalanced for the groups across classes in order to avoid practice effects. The independent variable in focus was task complexity, a within-subject factor in the present design, and various speech production and interactional measures constituted the dependent variables (see Figure 1). The ID variables,

FIGURE 1
Design and Procedures



measured by the means of self-reported questionnaire items, were treated as moderating factors.

Materials

Task. The experimental task was adapted from Watson, DeSanctis, and Poole (1988). It was contextualized in the imaginary scenario that the participants were members of a board for a personal trust foundation. The students were told that the aim of the foundation was to finance projects in the New York City borough of the Bronx, and their task was to evaluate competing requests for funding and make judgments about their relative merits during the course of a board meeting. Each group had to decide about two foundations (Task 1 and Task 2) in the meeting. One of them was worth \$500,000, which the board could distribute among three projects. The other amounted to \$10,000,000, and could be allocated to six programs. The potential projects were all community programs worthy of support, such as purchasing additional volumes for the library system, providing subsidized medical care for low-income families, constructing bike paths, and renovating government facilities. Given the limited resources, participants were asked to allocate funds only to those programs that they considered most deserving of a donation.

The two versions of the task were designed to differ along the +/- reasoning and the +/- few elements dimensions. In the fields of SLA (Ellis, 2003; Prabhu, 1987; Robinson, 2001a, 2005) and cognitive psychology (e.g., Halford, Cowan, & Andrews, 2007), it is generally accepted that tasks that induce more reasoning and/or reference to a greater number of elements are more cognitively complex than tasks with decreased demands in terms of these design features. In light of this, the task version with the larger fund and greater number of projects was considered higher in complexity, given that it required learners to present reasons and justifications for the ordering of a greater number of proposals and the allocation of more substantial resources. In order to ensure that the task design manipulation was realized as intended, 5 independent raters (all experienced ESL teachers familiar with the target student population) were asked to estimate the difficulty involved in accomplishing the two task versions and explain their respective ratings. Both the raters' judgments and explanations were consistent and in harmony with the researcher's task specifications.

The rationale for choosing an argumentative task for the study was threefold. First, Foster and

Skehan (1996; Skehan & Foster, 1997), in a series of studies investigating various task types, found that a similar task—what they refer to as a “decision-making” task—yielded the most consistent patterns for the general linguistic measures: accuracy and linguistic complexity. Second, previous task-based research in the interactionist framework suggests that convergent tasks, which require students to reach an agreement about a problem, are more likely to induce repair negotiation resulting from local non-understanding (Pica, Holliday, Lewis, & Morgenthaler, 1989) than divergent tasks, in which students are allowed to maintain different positions on an issue (Duff, 1986; Long, 1989).⁴ Lastly, argumentative tasks are typical in communicative classrooms and were, in fact, often used in the participating students' language classes.

Questionnaires. In order to assess participants' linguistic self-confidence, language use anxiety, and self-perceived communicative competence, a self-report questionnaire was employed. The items tapping participants' linguistic self-confidence (three items, Cronbach $\alpha = .71$) and language use anxiety (three items, Cronbach $\alpha = .84$) were adapted from Kormos and Dörnyei (2004). Learners' self-perceived communicative competence was measured by using a questionnaire developed by McCroskey and McCroskey (1988) for this purpose.

A task perception questionnaire was also included in the design to obtain information concerning the participants' perspectives, both teachers' and students', on their own subjective experiences throughout the tasks. In particular, participants were asked to rate on a 7-point scale which version of the task they found more useful for L2 learning, more difficult, more interesting, more stressful, more effective in drawing their attention to the quality of their output, and more successful in directing their attention to the quality of their peers' production.

Data Collection Procedures. All phases of the data collection were completed during one of the students' regular English classes. The classes lasted 2.5 hours. The tasks were presented to the students by their regular teacher as part of the normal class instruction. The researcher had shared all the materials and instructions with the teachers approximately 2 weeks prior to the lesson. Subsequently, she met with each of the instructors individually and clarified any questions that they had regarding the tasks and procedures. The tasks had originally been drafted by the researcher for

the purposes of the study. However, to ensure that they were feasible and relevant to the targeted student population, they were subjected to several cycles of revision based on comments from four experienced ESL instructors. These instructors were employed in the same program where the research was conducted but were not involved in the actual study. The tasks were then piloted with a class of ESL students and, based on the teacher and student comments, were slightly modified once again. The tasks were also administered to 2 groups of 3 native speakers each, all graduate students at the university, in order to obtain baseline data for identifying task-relevant linguistic measures.

As shown in Figure 1, the classes proceeded as follows: In the first part of the lesson, the two versions of the tasks were performed. Each task was carried out in three phases. First, each student individually considered their own priorities (5 minutes). Then, students negotiated in their groups to arrive at a joint decision (simple version: 15 minutes, complex version: 20 minutes). Finally, each small group shared their decision with the whole group (10–15 minutes). Then, this three-stage cycle was repeated for the other version of the task. At the beginning of each phase of the tasks, the teachers gave brief directions. While the students worked on the tasks, the teachers typically circulated among them and responded to any questions that they posed (only a small number of questions arose). Once the oral tasks had been completed, both the students and the teachers were administered the task perception questionnaire (10 minutes). Finally, after a short break, student participants completed the questionnaire assessing their self-perceived communicative competence, linguistic self-confidence, and language use anxiety (30 minutes).

All classroom interactions were recorded with mini digital audio-recorders, which were placed among the participants. The researcher was present in the classrooms throughout the whole course of the data collection in order to take observational notes and to monitor the equipment. The students were unlikely to be disrupted by the presence of the researcher, since class observations by visiting teachers and teacher trainees were frequent in the language program. When the data collection was finished, all small group interactions were transcribed by research assistants and then 10% of the data, randomly selected, were checked by the researcher. Next, all the transcripts were coded by the research assistants and the researcher.

Data Analysis Procedures

Speech Production Measures. To examine the effects of task complexity on the linguistic complexity and accuracy of speech production, general measures were used. Syntactic complexity was computed by dividing the number of syntactic clauses by the number of AS-units (Foster, Tonkyn, & Wigglesworth, 2000). As a measure of lexical diversity, values of D were obtained (Malvern & Richards, 2002) by using the computer program *vocd* in CLAN (MacWhinney, 2000). Accuracy was expressed as the ratio of errors to AS-units and as the ratio of error-free AS-units to AS units. A measure of error repair was also computed for the proportion of self-repairs in relation to total number of errors. This was included because self-repair behavior is claimed to indicate “accuracy in process,” as it reflects self-monitoring and learners’ resulting attempts at correcting their own speech production (Gilbert, 2007, p. 221; Kormos, 2000, 2006).

As a specific, developmentally motivated measure, participants’ use of conjoined clauses was employed. This linguistic feature was considered relevant to the +/- reasoning task dimension. Compared to tasks that do not involve reasoning, tasks with reasoning demands are likely to require speakers to justify their opinions and interpretations by presenting arguments about contrastive, causal, and conditional relationships. These communicative and conceptual demands, in turn, are expected to be reflected in language use, for example, in the incidence of coordinate conjunctions such as “but” and adverbial clauses with “because,” “so,” and “if” (Robinson, 2005). Indeed, these constructions appeared essential to task completion based on the baseline data collected from the native speakers.

According to Diessel (2004), the emergence of conjoined clauses in first language acquisition involves certain stages. First, conjoined clauses functioning as independent sentences emerge (e.g., “But the Snoopy is asleep.”), followed by biclausal coordinate sentences (e.g., “You push it and it goes up.”) and final adverbial clauses (e.g., “We both sleep on the floor when we take naps.”). Finally, children begin to use adverbial clauses that precede the matrix clause (e.g., “When it’s got flat tire, it’s need to go to the station.”). In terms of this developmental sequence, it was examined whether the learners in the present study would be more likely to show use of biclausal as opposed to unbound, and initial versus final adverbial clauses

on the simple task version versus the complex one. As far as the emergence of more specific clause types is concerned, Diessel observed the following developmental order: children first use “and” clauses, then “but,” “because,” and “so” clauses appear in their output. The last stage involves the emergence of “when,” “if,” and other temporal and concessive clauses such as “after,” “before,” and “until” constructions. Again, it was examined whether learners would show use of more or less developmentally advanced clause types as a function of task complexity.

Cohen’s (1992) kappa values indicated strong and above chance inter-coder agreement (.91–.96) for all measures of speech production.

Interactional Measures. In order to assess the effects of task complexity on interaction-driven language learning opportunities, the data were analyzed as follows: First, the LREs were identified in the dataset. Following Swain and Lapkin (2001) and Williams (1999), an LRE was defined as any part of the discourse “where students talk about language they are producing, question their language use, or other- or self-correct their language production” (Swain & Lapkin, 2001, p. 104). In other words, learner exchanges that had no potential in directing attention to code-related problems were not considered. LREs typically started either with a learner producing an utterance that was perceived and addressed as an error by another learner or with a learner query about a linguistic item. Once the LREs were identified, it was determined whether they were generated by a break in communication or whether they were triggered by a linguistic issue in the absence of an obvious comprehensibility problem (apparently, only those cases where a miscommunication was overtly discernable could be identified by the coders). Finally, the exchanges contained in the LREs were considered and coded in terms of the following interactional features: confirmation checks, clarification requests, recasts, and metalinguistic talk. The rationale for choosing these particular interactional elements as coding categories was that they all are hypothesized to be facilitative of SLA and have been the focus of previous task-related research. Definitions of these elements and examples of them, all taken from data collected for the current study, are provided here.

Confirmation checks are “expressions... immediately following an utterance by the interlocutor which are designed to elicit confirmation that the utterance has been correctly heard or un-

derstood by the speaker” (Long, 1983, p. 137). Example 1 shows an instance of a confirmation check, in which Learner 2, while providing implicit negative feedback to her peer, confirms that she has understood Learner 1’s view about which of two projects merits more financial support.

EXAMPLE 1
Confirmation Check

- Learner 1: I think to take care of our health is most important than emergency assistance.
Learner 2: You said health is more important than emergence? (confirmation check)
Learner 1: Yeah.

Clarification requests are expressions “designed to elicit clarification of the interlocutor’s preceding utterance(s)” (Long, 1983, p. 137). Example 2 illustrates a clarification request in which Learner 2 needs assistance in understanding Learner 1’s interpretation of the aim of a competing project.

EXAMPLE 2
Clarification Request

- Learner 1: OK number five to create emergency assistance to families. This is to help family to get temporary place to live.
Learner 2: To get what? (clarification request)
Learner 1: To get a temporary place in case they lost their house.

Recasts, in the current study, are defined as reformulations of a learner’s utterance by altering one or more incorrect forms therein while retaining its semantic content (Long, Inagaki, & Ortega, 1998). Example 3 illustrates a recast in which Learner 1 provides the correct adjectival form “safe” in response to Learner 2’s utterance.

EXAMPLE 3
Recast

- Learner 1: It is safety maybe the people think it is safety and would visit Bronx if there not many homeless maybe so.
Learner 2: Safety streets and improve quality.
Learner 1: Safe streets. (recast)
Learner 2: Safe, ah, safe streets and improve quality.

Finally, metalinguistic talk (Williams, 1999) refers to utterances in which learners discuss or query about language form and/or meaning. In Example 4, Learner 1 explicitly asks for a linguistic item that would better express his intended meaning.

TABLE 1
General Measures of Speech Production on Complex Versus Simple Tasks ($N = 43$)

	Clauses/AS <i>M/SD</i>	<i>D-values</i> <i>M/SD</i>	Errors/AS <i>M/SD</i>	Error Free AS/AS <i>M/SD</i>	Self-Repairs/ Errors <i>M/SD</i>
Simple	1.71/.60	64.84/22.31	.51/.29	.68/.17	.11/.22
Complex	1.45/.63	81.09/21.25	.41/.26	.76/.12	.10/.10
Sig/Effect	$p < .01, d = .48$	$p < .01, d = .74$	$p < .01, r = .18$	$p < .05, d = .37$	$p = .21, r = .03$

AS = AS-unit; Sig = significance; Effect = effect size; Simple = simple task; Complex = complex task

EXAMPLE 4

Metalinguistic Talk

Learner 1: And art makes people happy and then and then after grow up children they can create after graduate they can create better than now in Bronx. How to say they can create better? (metalinguistic talk)

Learner 2: 'Improve.'

Learner 1: They can improve themselves they can improve themselves for living and themselves ok. And what else? Number one is ok?

Except for Example 2, none of the interactional sequences listed in this section were coded as involving a communication breakdown.

Cohen's kappa values showed strong inter-coder agreement (.91-.94) for all interactional features (LREs, +/-communication breakdowns, types of LREs).

Statistical Analyses

First, descriptive statistics were calculated for the speech production, interactional, and ID measures. To address the first and second research questions, paired samples *t*-tests, Wilcoxon Signed Ranks, or Sign tests were conducted depending on the normality of the distributions and the nature of the outcome variables (interval or binary). An alpha level of $p < .05$ was set for all tests. To measure effect sizes, Cohen's *d* was calculated for the *t*-tests, *r* was calculated for the Wilcoxon tests, *g* was calculated for the Sign tests. By convention, *d* values of .20, .50, and .80, *r* values of .10, .30, and .50, and *g* values of .05, .15, and .25 were considered small, medium, and large, respectively (Cohen, 1992). The third research question was addressed by computing correlations between the ID variables and the measures of speech production and interaction-driven language learning opportunities.

RESULTS

Research Question 1: Effects of Task Complexity on the Accuracy and Complexity of L2 Speech Production

The descriptive and inferential statistics for the general measures of speech production are provided in Table 1. As the table demonstrates, participants produced more syntactically but less lexically diverse language on the simple task when compared to the complex task. Paired samples *t*-tests found the effects for both syntactic complexity ($t(2, 42) = -3.11, p < .01$) and lexical diversity ($t(2, 42) = 4.88, p < .01$) statistically significant, and the effect sizes were small and medium ($d = .48$ and $d = .74$, respectively).

Turning to accuracy, the more complex task resulted in a lower rate of errors and self-repairs and a higher rate of error-free AS-units. A Wilcoxon Signed Ranks test and a paired-samples *t*-test yielded a significant effect for error rate ($z = -2.67, p < .01$) and for rate of error-free AS-units ($t(2, 42) = -2.43, p < .05$). However, the observed difference in the rate of self-repairs was not statistically significant ($z = -1.25, p = .21$). The effect sizes for errors per AS-unit ($r = .18$) and for error-free AS units per total number of AS-units ($d = .37$) were in the small range.

Tables 2 and 3 give the results obtained for the specific, developmentally motivated measure of speech production—that is, the use of conjoined clauses (Diessel, 2004). As shown in Table 2, approximately the same number of participants showed use of each type of conjoined clause in focus—unbound, bound, final, and initial—on the simple versus complex task versions. Differences, if at all observed, were not found significant by Wilcoxon Sign tests ($p > .05$).

Table 3 provides the data obtained for the appearance of specific types of conjunctions in conjoined clauses as a function of whether learners performed simple or complex tasks. As the table demonstrates, a greater number of

TABLE 2
Types of Conjoined Clauses Used on Complex Versus Simple Tasks ($N = 43$)

	Unbound N (Percentage)	Bound N (Percentage)	Final Adverbial N (Percentage)	Initial Adverbial N (Percentage)
Simple	43 (100%)	42 (97.7%)	41 (95.3%)	36 (83.7%)
Complex	43 (100%)	43 (100%)	42 (97.7%)	37 (86.9%)
Sig/Effect	- / -	- / -	$p = 1.00/g = .02$	$p = 1.00/g = .03$

N = number of participants showing use of the construction; Simple = simple task; Complex = complex task; Sig = significance; Effect = effect size

TABLE 3
Types of Conjunctions Used on Complex Versus Simple Tasks ($N = 43$)

	<i>and</i> N (Per)	<i>but</i> N (Per)	<i>so</i> N (Per)	<i>because</i> N (Per)	<i>when/if</i> N (Per)
Simple	37 (86.9%)	34 (79.1%)	35 (81.4%)	32 (74.4%)	33 (76.7%)
Complex	39 (90.7%)	41 (95.3%)	42 (97.7%)	41 (95.3%)	39 (90.7%)
Sig/Effect	$p = .69, g = .04$	$p < .05, g = .16$	$p < .05, g = .16$	$p < .01, g = .21$	$p = .07, g = .14$

N = number of participants showing use of the construction; Per = percentage; Simple = simple task; Complex = complex task; Sig = significance; Effect = effect size

participants showed use of all types of conjunctions when learners engaged in complex (as opposed to simple) task performance. Except for “and” clauses ($p = .69$) and “when” and “if” clauses ($p = .07$), Wilcoxon Sign tests confirmed these task-related variations to be statistically significant. The magnitude of these effects were all medium size ($.15 < g < .25$).

Research Question 2: Effects of Task Complexity on the Quantity and Quality of Interaction-Driven Learning Opportunities

Table 4 demonstrates the descriptive and inferential statistics obtained for the total rate of interaction-driven language learning opportunities across the complex and simple task versions. As shown, for all interactional measures, greater task complexity generated a higher rate of language-learning opportunities. Results of Wilcoxon Signed Ranks tests, however, yielded a significant effect for only two measures: LREs per AS-unit ($z = -3.64, p < .01$) and metalinguistic talk per AS-unit ($z = -2.80, p < .01$). Both of these effects were of medium size ($r = .39$ and $r = .30$, respectively). The observed differences on the remaining measures—clarifications requests, confirmation checks, and recasts—did not reach significance ($-1.72 < z < -1.00, p > .05$).

Table 5 shows the descriptive and inferential statistics for the rate of interaction-driven language learning opportunities as a function of task complexity and the presence or absence of communication problems. Similar trends were observed as for the total rate of interactional elements. Higher rates of language learning opportunities were found on the complex task when compared to the simple task for LREs, both with and without communication problems (CP). However, none of the differences were found significant by Wilcoxon Signed Ranks tests ($-1.447 < z < -1.718$) except for LREs per AS-unit (with CP: $z = -2.39, p = .02$, without CP: $z = -2.82, p < .01$) and metalinguistic talk per AS-unit (with CP: $z = -2.80, p < .01$, without CP: $z = -2.80, p < .01$). The effect size for the rate of LREs with CPs was small ($r = .26$), whereas the rest of the significant effects were in the medium range ($r = .30$).

Research Question 3: Modulating Effects of Individual Differences on the Relation of Task Complexity to Speech Production and Interaction-Driven Learning Opportunities

Table 6 provides the correlations between the individual difference variables and general measures of speech production. As shown, none of the correlations between the ID and speech production measures were found to be significant.

TABLE 4

Total Rate of Interaction-Driven Language Learning Opportunities on Complex Versus Simple Tasks
($N = 43$)

	LRE/AS <i>M/SD</i>	ClarReqs/AS <i>M/SD</i>	ConfChs/AS <i>M/SD</i>	Recasts/AS <i>M/SD</i>	MetaTalk/AS <i>M/SD</i>
Simple	.022/.035	.009/.024	.001/.007	.008/.015	.000/.000
Complex	.049/.060	.010/.018	.007/.017	.010/.016	.014/.031
Sig/Effect	$p < .01, r = .39$	$p = .29, r = .11$	$p = .09, r = .19$	$p = .32, r = .11$	$p < .01, r = .30$

AS = AS-unit; ClarReqs = clarification requests; ConfChs = confirmation checks; Simple = simple task; Complex = complex task; Sig = significance; Effect = effect size

TABLE 5

Rate of Interaction-Driven Language Learning Opportunities With/Without Communication Problems on Complex Versus Simple Tasks ($N = 43$)

	LRE/AS <i>M/SD</i>	ClarReqs/AS <i>M/SD</i>	ConfChs/AS <i>M/SD</i>	Recasts/AS <i>M/SD</i>	MetaTalk/AS <i>M/SD</i>
+ComProb					
Simple	.012/.025	.007/.021	.001/.003	.000/.001	.000/.000
Complex	.021/.032	.008/.014	.003/.008	.001/.002	.007/.015
Sig/Effect	$p < .05, r = .26$	$p = .26, r = .12$	$p = .09, r = .19$	$p = .66, r = .05$	$p < .01, r = .30$
-ComProb					
Simple	.010/.017	.002/.010	.001/.003	.007/.015	.000/.000
Complex	.028/.042	.003/.009	.003/.008	.009/.016	.007/.015
Sig/Effect	$p < .01, r = .30$	$p = .58, r = .06$	$p = .09, r = .19$	$p = .32, r = .11$	$p < .01, r = .30$

AS = AS-unit; ClarReqs = clarification requests; ConfChs = confirmation checks; ComProb = communication problem; Simple = simple task; Complex = complex task; Sig = significance; Effect = effect size

TABLE 6

Spearman Correlations of the Individual Difference Variables and General Measures of Speech Production on Complex Versus Simple Tasks ($N = 43$)

	Clauses/AS		D-values		Errors/AS		Error-free AS/AS		Self-repairs/AS	
	Simple	Comp	Simple	Comp	Simple	Comp	Simple	Comp	Simple	Comp
Confidence	-.11	-.04	.07	.13	-.26	-.16	.17	.04	.18	.23
Anxiety	.18	-.11	.04	.20	-.09	-.09	.16	-.06	.09	-.09
SPCC	-.10	-.05	-.13	-.01	-.12	-.11	.22	.13	.08	.12

AS = AS-unit; Simple = simple task; Comp = complex task

As Table 7 demonstrates, the correlations computed between the individual difference variables and measures of interaction-driven language learning opportunities did not yield any statistically significant results in the present study.

Task Perception Questionnaire

Table 8 shows the results on the task perception questionnaire. As the table indicates, the participants, on average, perceived the task demands in

the intended way: both the students and teachers rated the more complex task as more difficult. Learners and teachers' perceptions largely overlapped on the rest of the items as well. On average, they felt that the more complex task was more useful for L2 learning, as well as more interesting and more effective in directing learner attention to their own and others' linguistic output. Students and teachers' perceptions, however, differed as to their perception of learner anxiety during task work. Although teachers, in general, reported that students were equally relaxed

TABLE 7
Spearman Correlations of the Individual Difference Variables and Measures of Interaction-Driven Language Learning Opportunities on Complex Versus Simple Tasks (*N* = 43)

	LRE/AS		ClarReqs/AS		ConfChs/AS		Recasts/AS		MetaTalk/AS	
	Simple	Comp	Simple	Comp	Simple	Comp	Simple	Comp	Simple	Comp
Confidence	-.14	-.21	-.17	-.07	-.11	-.14	.13	.01	-	-.21
Anxiety	.05	.11	.06	.26	.14	-.06	.01	.14	-	-.11
SPCC	.16	.01	.24	.07	.19	-.11	-.02	-.17	-	-.07

AS = AS-unit; ClarReqs = clarification requests; ConfChs = confirmation checks; Simple = simple task; Comp = complex task.

TABLE 8
Perceptions of Task Difficulty (Students, *N* = 43; teachers, *N* = 6)

	Usefulness <i>M/SD</i>	Task Difficulty <i>M/SD</i>	Task Motivation <i>M/SD</i>	Anxiety <i>M/SD</i>	Attention to Own Output <i>M/SD</i>	Attention to Peer Output <i>M/SD</i>
Students	3.79/1.88	3.76/2.27	3.91/1.97	4.20/1.91	3.95/1.84	3.57/1.63
Teachers	3.67/1.03	4.17/1.83	4.83/1.83	3.50/2.07	4.17/0.41	4.00/0.63

Note. The values in the table represent participants' ratings on a 7-point scale. Values of 0 and 7, respectively, would indicate the highest possible rating in favour of the simple versus complex task versions.

during the two tasks, learners (on average) felt more stress on the complex version.

DISCUSSION AND CONCLUSION

The first research question asked whether the cognitive complexity of tasks in which learners are engaged will impact the accuracy and complexity of their speech production. Motivated by Skehan's Trade-Off Hypothesis and Robinson's Cognition Hypothesis, it was expected that task complexity would influence the accuracy and linguistic complexity of learners' output. The results provide strong confirmation of this hypothesis. As Table 1 demonstrates, all the general measures of speech production (except for self-repair) yielded significant differences between learners' performance on simple and complex tasks in the current study. In particular, participants produced more lexically diverse and accurate but syntactically less complex language when task complexity was enhanced. The more specific measure employed—learners' use of conjoined clauses—also demonstrated some variation as a function of task complexity (see Table 3). Participants were significantly more likely to show use of developmentally advanced constructions on complex tasks than on simple tasks in terms of Diessel's (2004) acquisition order for conjoined clauses. It is also worth noting that, in line with these overall patterns, par-

ticipants reported placing more attention to the quality of their linguistic output while performing the complex task version than while performing the simple one, according to their ratings in the task perception questionnaire (see Table 8).

The findings on the general speech production measures can be interpreted as follows: The improved performance, in terms of global accuracy and lexical diversity on the complex task, implies that the task with greater cognitive demands was more successful in directing attention to form-meaning connections. That is, enhanced cognitive complexity, overall, did not appear to afford learners less attentional resources for their linguistic performance. This finding seems more compatible with the Cognition Hypothesis than the Trade-Off Hypothesis. The former model presumes that greater complexity along resource-directing dimensions, and the increased conceptual demands associated with them, has the capacity to draw learner attention to task-relevant constructions. The latter framework, in contrast, predicts less attention available for language, manifest in both lower accuracy and linguistic complexity when task complexity is enhanced.

However, unlike the results for accuracy and lexical diversity, the decreased syntactic complexity observed on the more complex task version could be interpreted as providing support, at least in part, for both Skehan's and Robinson's

predictions, albeit following distinct lines of reasoning. Robinson (2005) predicts lower structural complexity on more complex interactive tasks owing to the amplified amount of negotiation that tends to result when the cognitive demands of interactive tasks are increased, a trend that was attested in the present study. Skehan, on the other hand, in harmony with his limited attentional capacity model, projects a competition between accuracy and complexity, leading to trade-off effects between these two areas, as was observed here. Such trade-off effects, however, do not seem to be corroborated by the results on the specific, developmental measure. As mentioned, learners were more likely to show use of complex, developmentally more advanced conjoined clauses (Diessel, 2004) on complex tasks when compared to simple tasks. The more complex task, probably given the enhanced reasoning demands it entailed, required and elicited more complex use of conjoined clauses with the words “but,” “because,” “so,” and “if”—that is, constructions associated with the process of reasoning (Robinson, 2005).

With regards to speech production, it is also worth highlighting that the global and specific measures generated distinct findings. Although participants showed greater syntactic complexity on complex tasks in terms of the general index employed (clauses per AS-unit), they were observed to use more complex language on the same task in terms of the task-relevant, developmental measure (use of conjoined clauses). This finding confirms the importance of recent calls (Bygate & Samuda, 2005; Mochizuki & Ortega, 2008; Norris & Ortega, 2009; Robinson, 2007; Robinson & Ellis, 2008) for employing both general and specific measures when investigating the link between task complexity and the quality of speech production. The general practice of relying exclusively on global measures, this study suggests, might result in an incomplete description of task effects.

The second research question was concerned with whether task complexity would affect the quantity and quality of interaction-driven language learning opportunities that arise during task-based interaction. The results of this study indicate that, as predicted, there were differences in the overall rate of LREs that occurred during simple tasks versus complex tasks (see Table 4). In line with the Cognition Hypothesis, tasks with greater cognitive demands induced a significantly higher rate of LREs than less cognitively complex tasks. Although similar trends were observed for the more specific interactional measures (i.e., clarification requests, confirmation checks, recasts, and

metalinguistic talk), these only reached significance for utterances involving metalinguistic talk. The task perception data largely conform to these patterns: participants, on average, reported slightly greater attention to their peers' output when completing the more complex task.

These tendencies for LREs are similar to Robinson's (2001b, 2007) findings for clarification requests and confirmation checks and to those of Nuevo (2006) for confirmation checks. The results are also in line with Gilabert et al.'s (2009) overall findings for interactional features, and partly overlap with those of Kim (2009) for LREs. As mentioned earlier, however, the effects detected in previous research reached statistical significance. One reason for the discrepancy between Robinson's, Nuevo's, and Gilabert et al.'s findings and those reported here may be that, unlike in their analyses, only interactional moves that arose as part of LREs were included. That is, negotiation sequences that appeared to have no potential in directing attention to form-function mappings were not considered here. Similar to the present study, however, Kim exclusively investigated LREs, and yet found different patterns. Evidently, further research is required to disentangle the potential effects of variables such as task type and proficiency level.

With regard to the quantity of LREs observed, it is also worth noting that the overall incidence of LREs was considerably small (3.5% of AS-units). This is not an entirely unexpected finding. First, previous research comparing native-speaker and learner-learner interaction suggests that learner-learner interaction may, in general, provide fewer opportunities for learners to receive feedback (e.g., García Mayo & Pica, 2000; Mackey et al., 2003). Another possible factor that may have affected learner use of LREs is the nature of the task which learners performed. The argumentative task that they completed was more of an opinion than an information gap task. Although it required some information exchange, the bulk of the information exchange was optional. It is well-documented that required information exchange tasks are more effective in instigating interaction-driven language learning opportunities than optional information exchange tasks (see Ellis, 2003 for a review; Gass et al., 2005).

Next, the two hypotheses addressing the relationship between task complexity and types of interaction-driven language learning opportunities are discussed. The first hypothesis predicted that LREs occasioned by communication breakdowns would be more frequent on tasks with greater cognitive complexity. Drawing on the

Cognition Hypothesis, it was assumed that more complex tasks, and the increased communicative demands they induce, will lead to comprehension problems more frequently and, hence, a larger amount of negotiation for meaning. The results of the present study uniformly confirm this prediction. As Table 5 shows, a significantly higher rate of LREs was observed in the absence of communication breakdowns when participants performed high (as opposed to low) complexity interactive tasks.

The second hypothesis regarding the effects of task complexity on specific types of interaction-driven language opportunities was concerned with LREs initiated in the absence of communication breakdowns. Inspired by insights from the Trade-Off and Cognition Hypotheses, it was hypothesized that task complexity, owing to its respective impact on attentional processes, would influence the extent to which learners engage in such LREs. The results confirm the second hypothesis. The rate of LREs prompted by purely code-focused problems was found significantly higher on the complex task version when compared to the simple task version (see Table 5). Notably, the direction of this trend seems to be more compatible with insights derived from the Cognition than the Trade-Off Hypothesis. As discussed earlier, the former model suggests that there will be more opportunities for learners to draw attention to form–meaning mappings when task complexity is enhanced along resource-directing task dimensions. More complex tasks along such task features, by definition, are more effective in attracting attention to relevant L2 constructions, and hence, learners will arguably be more inclined to make an attempt at directing their interlocutors' attention to form–meaning connections during task performance.

It is an important finding that more complex tasks along resource-directing dimensions are more prone to inducing LREs when understanding between interlocutors is sufficient for communication. As mentioned earlier, there might be greater likelihood that linguistic features, which are difficult for learners to acquire without assistance, become the focus of LREs when no communication breakdown occurs. Constructions that tend to prove hard to acquire are often communicatively redundant and/or perceptually non-salient linguistic elements (e.g., morphosyntactic features), thus unlikely to result in communication failure. Indeed, a follow-up Wilcoxon test revealed that morphosyntactic features were significantly more frequently targeted ($p < .01$, $r = .21$) in LREs unaccompanied by comprehension

problems (LREs/AS-unit: $M = .025$, $SD = .03$) than in LREs caused by breaks in communication (LREs/AS-unit: $M = .012$, $SD = .02$).

The third research question investigated the extent to which IDs in self-perceived communicative competence, linguistic self-confidence, and language use anxiety modulate the effects of task complexity on the quality of linguistic output and on the amount and type of interaction-driven language learning opportunities arising during task-based work. Inspired by the Cognition Hypothesis, it was hypothesized that learner factors would differentiate these relationships to a greater extent on the complex task version than on the simple task version. This prediction was rejected by the results of the present study. As shown in Tables 6 and 7, simple correlational analyses revealed no effects for any of the ID indices.

A question that arises from these results is why the ID variables in the present study had no association with the learners' speech production and the occurrence of interactional features. As far as speech production is concerned, one possible explanation for the absence of effects might have to do with the proficiency level of the students. The participants here were all advanced speakers of English. In other words, they were all successful language learners. Therefore, it is possible that, even if they had less favorable qualities with respect to the three learner factors in focus, they might have developed strategies to overcome or minimize the impact of these variables. The fact that the ID variables did not influence the quantity and quality of interaction-driven language learning opportunities is more puzzling. It would appear reasonable to assume that learners with lower anxiety, greater self-confidence, and higher self-perceived communicative competence would be more inclined to correct their peers' errors and initiate LREs more generally. Clearly, more research is needed to test the validity of this speculation and the generalizability of these patterns to other learners and contexts.

Finally, it is worth discussing some pedagogical implications of this research. Apparently, one needs to be careful in drawing conclusions for pedagogy based on this study. Although it was performed in a classroom setting, it had features that are typical of laboratory, rather than pure classroom, research (e.g., the tasks and procedures were originally designed by the researcher for experimental purposes) and, as it has been argued extensively, the results of laboratory studies may not be directly applicable to the classroom (e.g., Lightbown, 2000; Samuda & Bygate, 2008). With this caveat in mind, this study, along with

those of previous research investigating learner-learner interaction (e.g., Adams, 2007; Gass et al., 2005; Nuevo, 2006; Williams, 1999; see, however, Foster, 1998), suggests that learners working together, albeit not very frequently, do negotiate matters of L2 constructions in the context of meaningful communication. In addition, the results of the current study imply that teachers may enhance the extent to which learners pay attention to form–meaning mappings during task-based work by cautiously designing and choosing pedagogic tasks. In particular, careful task manipulations may induce learners to engage in increased monitoring behavior, more extensive experimentation with the L2, and higher incidence of interaction-driven language learning opportunities.

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

There are a number of limitations to this study that need to be acknowledged and considered in future research. First, a major weakness resides in the fact that the present study exclusively focuses on learner performance rather than development. Although there are theoretical reasons to believe that the effects identified for speech production (Robinson, 2005; Skehan, 1998) and interactional patterns can affect L2 development (e.g., Mackey & Goo, 2007; Robinson, 2005), only studies that establish a direct link between these constructs can provide reliable evidence of this. To date, however, only a few empirical studies (Nuevo, 2006; Révész, 2009) have set out to investigate the longer-term impact of task complexity. Therefore, further studies in this area are warranted. A second shortcoming has to do with the fact that a difference in task complexity between the two versions of the experimental task was assumed rather than established. In a follow-up study, an attempt could be made to confirm independently the construct validity of task complexity as operationalised here. A third limitation of the study is that the specific measure adopted—the use of conjoined clauses—was motivated by research findings from first (rather than second) language acquisition research. A developmental pattern established for SLA would have, no doubt, constituted a more valid basis for a specific analytical tool. Lastly, the present study was situated in a cognitive-interactionist framework. Thus, the notions of task complexity and language learning opportunities were not considered from other viewpoints such as sociocultural and conversation analysis perspectives (cf., Nakahama, Tyler, & van

Lier, 2001). In a follow-up study, it would be interesting to reanalyze the dataset in terms of these approaches.

Another important avenue for future research would involve extending the research questions posed here to other tasks, conditions, and contexts. The current study included a single task type that involved learners in small group interaction in the context of a community ESL program in the United States. It is not necessarily the case that the effects found here would transfer to different task types, distinct interactive conditions, and/or other instructional and sociocultural contexts.

ACKNOWLEDGMENTS

I would like to thank Judit Kormos, Peter Robinson, Rebecca Sachs, and Eun Sung Park for their insightful comments on this article. I am also grateful to Monika Ekiert, Jae Yong Choi, Marina Gotovina, Tia Lazarus, Jennifer Lu, and Steve Mercier for their assistance with data collection, and to Moon Jeung Chang, Mika Hama, YunKyoung Kang, and Sun Hee Wang for helping me code the data.

NOTES

¹Task complexity is different from task difficulty, which refers to learners' perception of the cognitive demands of the task (Robinson, 2001a).

²Robinson proposes that increasing task complexity along resource-directing dimensions "involves some recapitulation of a sequence of conceptual development in childhood" (Robinson, 2005, p. 6). Thus, increasing task demands along such dimensions will provide learners with ideal, "ontogenetically" natural contexts for acquiring L2 constructions, just as cognitive and conceptual development has been argued to drive, to a great extent, L1 acquisition (Cromer, 1974). For example, as Cromer points out, reference to here-and-now events emerges earlier in childhood than the ability to describe there-and-then events, and a similar acquisition order was observed in L2 acquisition (Robinson, 2005). This is reflected in Robinson's Cognition Hypothesis, which argues that increases in task complexity along \pm here-and-now dimension will promote L2 development in relevant linguistic constructions such as past time reference.

³All the students in each of the 6 classes participated in the study. Two groups were formed in 4 classes. In the remaining 2 classes, originally 3 and 4 groups performed the tasks. One group from both classes, however, was excluded from the analysis because a student was late for the session and thus missed part of the first task. Data gathered from another group had to be removed due to problems with the digital recorder.

⁴Note that this observation does not seem to hold when negotiation is defined in a broader sense,

including discourse level phenomena that help achieve global, rather than local, cohesion and coherence during interaction (see Nakahama et al., 2001).

REFERENCES

- Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 29–51). Oxford: Oxford University Press.
- Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37–74). Amsterdam/New York: Benjamins.
- Bygate, M., Skehan, P., & Swain, M. (2001). (Eds.). *Researching pedagogic tasks, second language learning teaching and testing*. Harlow: Longman.
- Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, *112*, 155–159.
- Cromer, R. (1974). The development of language and cognition: The cognition hypothesis. In B. Foss (Ed.), *New perspectives in child development* (pp. 184–252). Harmondsworth, England: Penguin.
- Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.
- Dörnyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 137–158). Amsterdam: Benjamins.
- Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Erlbaum.
- Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, *4*, 275–300.
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197–261). New York: Cambridge University Press.
- Duff, P. (1986). Another look at interlanguage talk: Taking talk to task. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Rowley, MA: Newbury House.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, *51*, 281–318.
- Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, *19*, 1–23.
- Foster, P., & Ohta, A. (2005). Negotiation for meaning and peer assistance in second language classrooms. *Applied Linguistics*, *26*, 402–430.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *19*, 199–323.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *21*, 354–375.
- García Mayo, M. P. (Ed.). (2007). *Investigating tasks in formal language learning*. Clevedon, Buffalo, Toronto: Multilingual Matters LTD.
- García Mayo, M. P., & Pica, T. (2000). L2 learner interaction in a foreign language setting: Are learning needs addressed?, *International Review of Applied Linguistics*, *38*, 35–58.
- Gass, S. M. (1997). *Input, interaction and the development of second languages*. Mahwah, NJ: Erlbaum.
- Gass, S. M., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning*, *55*, 575–611.
- Gilbert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics*, *45*, 215–240.
- Gilbert, R., Barón, J., & Llanes, À. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics*, *47*, 367–395.
- Halford, G., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, *11*, 236–242.
- Han, Z.-H. (2007). On the role of “meaning” in focus on form. In Z.-H. Han (Ed.), *Second language process* (pp. 45–79). Clevedon, England: Multilingual Matters.
- Keck, C. M., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. M. Norris and L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Amsterdam: Benjamins.
- Kim, Y. (2009). The effects of task complexity on learner–learner interaction. *System*, *37*, 254–268.
- Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*, *50*, 343–84.
- Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Erlbaum.
- Kormos, J., & Dörnyei, Z. (2004). The interaction of linguistic and psychological variables in second language task performance. *Zeitschrift für Interkulturellen Fremdsprachunterricht*. [Online], 9(2), <http://zif.spz.tu-darmstadt.de/jg-09-2/beitrag/kormos2.htm>
- Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. H. Foster-Cohen, M. del Pilar García Mayo, & J. Cenoz (Eds.), *Eurosla Yearbook 5* (pp. 195–222). Amsterdam: Benjamins.

- Kuiken, F., & Vedder, I. (2007). Cognitive task complexity and linguistic performance in French L2 writing. In M. del Pilar García Mayo (Ed.), *Investigating Tasks in Formal Language Learning* (pp. 117–135). Clevedon, England: Multilingual Matters.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309–365.
- Lightbown, P. M. (2000). Anniversary article: Classroom SLA research and second language teaching. *Applied Linguistics*, 21, 431–462.
- Lightbown, P., & Spada, N. (1990). Focus on form and corrective feedback in communicative language teaching: Effect on second language learning. *Studies in Second Language Acquisition*, 12, 429–448.
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, 126–141.
- Long, M. H. (1989). Task, group, and task-group interaction. *University of Hawaii Working Papers in ESL*, 8, 1–26.
- Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & B. K. Bahtia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.
- Long, M. H. (2007). Recasts: The story so far. In M. H. Long (Ed.), *Problems in SLA* (pp. 75–116). Mahwah, NJ: Erlbaum.
- Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 24, 27–56.
- Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative evidence in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, 82, 357–371.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 15–41). Cambridge: Cambridge University Press.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265–302.
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60, 501–533.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 407–452). Oxford: Oxford University Press.
- Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS–NNS and NNS–NNS adult and child dyads. *Language Learning*, 53, 35–66.
- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209). Philadelphia: Benjamins.
- Mackey, A., & Sachs, R. (2011). Older learners in SLA research: A first look at working memory, feedback and L2 development. *Language Learning*. DOI: 10.1111/j.1467/9922.2011.00649.x
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19, 85–104.
- McCroskey, J. C., & McCroskey, L. L. (1988). Self-report as an approach to measuring communication competence. *Communication Research Reports*, 5, 108–113.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Erlbaum.
- Michel, M., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics*, 45, 241–259.
- Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12, 11–37.
- Nakahama, Y., Tyler, A., & van Lier, L. (2001). Negotiation of meaning in conversational and information gap activities: A comparative discourse analysis. *TESOL Quarterly*, 35, 377–405.
- Niwa, Y. (2000). *Reasoning demands of L2 tasks and L2 narrative production: Effects of individual differences in working memory, intelligence, and aptitude* (Unpublished master's dissertation). Aoyama Gakuin University, Tokyo.
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578.
- Nuevo, A. (2006). *Task complexity and interaction: L2 learning opportunities and interaction* (Unpublished doctoral dissertation). Georgetown University, Washington DC.
- Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes and outcomes? *Language Learning*, 44, 493–527.
- Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, 11, 63–90.
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. In G. Crookes & S. M. Gass (Eds.), *Tasks and Language Learning* (pp. 9–34). Clevedon, England: Multilingual Matters.
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
- Révész, A. (2009). Task complexity, focus on form, and second language development.

- Studies in Second Language Acquisition*, 30, 437–470.
- Révész, A. (in press). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning*.
- Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating 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). Attention and memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631–678). Malden, MA: Blackwell.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43, 1–32.
- Robinson, P. (2007). 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*, 45, 193–213.
- Robinson, P., & Ellis, N. (2008). Conclusions: Cognitive linguistics, second language acquisition and L2 instruction—Issues for research. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 489–545). New York/London: Routledge.
- Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *International Review of Applied Linguistics*, 45, 161–176.
- Sagarra, N. (2007). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 229–248). Oxford: Oxford University Press.
- Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Palgrave MacMillan.
- Sato, C. (1986). Conversation and interlanguage development: Rethinking the connection. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 23–45). Rowley, MA: Newbury House.
- Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 301–322). Oxford: Oxford University Press.
- Sheen, Y. (2008). Recasts, language anxiety, modified output and L2 learning. *Language Learning*, 58, 835–874.
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skehan, P. (2007, September). Tradeoff and Cognition: Two hypotheses regarding attention during task-based performance. Plenary talk at the 2nd International Conference on Task-based Language Teaching, University of Hawaii.
- Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1, 185–211.
- Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press.
- 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. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 471–484). Mahwah, NJ: Erlbaum.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82, 320–337.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 99–118). London: Longman.
- Talmy, L. (2000). *Toward a cognitive semantics: Vol. 1. Concept structuring systems*. Cambridge, MA: MIT Press.
- Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171–195). Oxford: Oxford University Press.
- Watson, R., DeSanctis, G., & Scott Poole, M. (1988). Using a GDSS to facilitate group consensus: Some intended and unintended consequences. *MIS Quarterly*, 12, 463–478.
- Wickens, C. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.
- Wickens, C. (2007). Attention to the second language. *International Review of Applied Linguistics*, 45, 177–191.
- Williams, J. (1999). Learner-generated attention to form. *Language Learning*, 49, 583–625.