

# Effects of Task Reasoning Demand and Task Condition on Learner Written Output in ESL Classrooms

Lilliati Ismail<sup>1,\*</sup>, Arshad Abd. Samad<sup>2</sup>, Wong Bee Eng<sup>3</sup> & Nooreen

Noordin<sup>4</sup>

<sup>1</sup>S.A.M Unwanus Saadah, 42700 Banting, Selangor, Malaysia

<sup>2</sup>Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

<sup>3</sup>Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

<sup>4</sup>Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

\*Corresponding author: S.A.M Unwanus Saadah, 42700, Banting, Selangor, Malaysia

E-mail: lilliati75@yahoo.com

Received: August 15, 2012   Accepted: October 7, 2012   Published: November 25, 2012

doi:10.5296/ije.v4i4.2249

URL: <http://dx.doi.org/10.5296/ije.v4i4.2249>

---

**Abstract**

Considering the growing interest in task-based language teaching, classroom-based research that investigates the effects of task complexity on L2 development is needed. Despite the inclusion of task reasoning demand (TRD) as a dimension of task complexity in Robinson's Cognition Hypothesis (2007), there is insufficient classroom-based research that investigates the language learning outcomes that may occur as a result of engaging in tasks of differing reasoning demands in a variety of task conditions. This study aims to fill in some of the gap by identifying the main and interaction effects of task reasoning demand and individual versus dyadic task conditions (TC) on the grammatical accuracy and syntactic complexity of learner written output. Modified versions of the dictogloss task and the opinion-gap task were used to provide a relatively high reasoning demand task (+TRD) and a relatively low reasoning demand task (-TRD) to the learners respectively. A repeated-measures design was used with 76 participants consisting of 18 year-old learners in a public secondary school randomly assigned into four groups. Data were analysed using descriptive statistics and repeated-measures ANOVAs. Results indicated that both TRD and TC had significant main effects on grammatical accuracy. Also, TRD and TC had significant main and interaction effects on syntactic complexity. The results point to differential effects of using tasks of high and low reasoning demand in dyadic and individual task conditions. The results have pedagogical implications on task design and task selection to elicit higher rates of grammatical accuracy and syntactic complexity in learner written output.

**Keywords:** Task complexity; Task reasoning demand; Task condition; Dictogloss; Opinion-gap

## **1. Introduction**

Engaging learners in pair or group discussions is a common feature of task-based language teaching (TBLT). This approach in second language (L2) teaching advocates the use of activities and tasks that could elicit learner talk. A growing interest among practitioners and researchers on TBLT has led to several questions such as ‘How does talking to each other while engaging in tasks help learners develop their interlanguage?’ and ‘Does manipulating task conditions and task demands have differential effects on L2 production?’ These are some of the questions that have motivated this study.

TBLT constitutes the strong version of communicative language teaching (Ellis, 2003). However, through the use of focus on form in TBLT, lessons can be interrupted by providing focus to linguistic problems as and when it is necessary to aid communication (Long, 1990). This ‘interruption’ could come from the speaker who raises questions about his own linguistic gaps, the teacher, or other learners. Researchers such as Nunan (1989) and Ellis (2003) claim that interactional feedback, negotiation of meaning, attention to form, and modified language output that occur as a result of focus on form could potentially promote L2 learning.

This study has investigated how interacting with each other while engaging in tasks could help learners develop their language skills. In particular, the study focuses on the influence of task complexity (i.e., high and low reasoning demands) and task conditions (i.e., individual versus dyadic) on the accuracy and syntactic complexity of learner written output.

## **2. Task Complexity, Task Conditions and Language Learning**

In order to understand the place of learner-learner interaction in L2 pedagogy, the theoretical underpinnings that link task-based interaction with interlanguage development need to be established. The following sections provide an overview of task-based interaction in the L2, followed by theoretical perspectives linking interaction to L2 learning in a task-based teaching- learning context.

### *2.1 Task-based Instruction and Task Complexity*

SLA researchers have recognised the potential of task-based instruction to create more opportunities to negotiate meaning in comparison to traditional instruction (e.g., Long, 1985, 1991; Robinson, 2001). DeKeyser (1998) also points out that task-based language teaching is particularly useful in facilitating the development of accuracy in language production. Pica (1992, 1994) explains that communication tasks such as information gap allow for comprehension, feedback and modified output. Robinson (2001a, 2001b, 2003, 2005, 2007) and Skehan (1998) posit that cognitive demands inherent in tasks would greatly affect learner performance.

Task complexity, in particular, has garnered attention among SLA researchers. Previous research on task complexity have addressed the notion of reasoning demands, which is the extent to which a task requires learners to reason, provide justifications, or explain causalities. Reasoning demands are often included in task complexity and task difficulty dimensions,

notably by Robinson (2001b, 2003, 2005, 2007) and Skehan (1996, 1998). Though research has shown that engaging in learner-learner interaction would provide opportunities for output, there is a lack of empirical research investigating how they affect L2 written output.

## *2.2 The Interactionist and Information-processing Perspectives of L2 Learning*

The interactionist and the information-processing perspectives are two broad perspectives that underpin the use of learner-learner interaction and tasks to assist language learning. Early second language acquisition (SLA) research emphasised the importance of comprehensible input for second language (L2) acquisition. Krashen (1985), through his Input Hypothesis, posits that language is developed through massive comprehensible input of the language. The hypothesis is based on the concept of  $i + 1$  which means that comprehensible input which is just a little beyond a learner's current level of linguistic competence is necessary and sufficient for language development. However, this view has been heavily criticised for being insufficient in explaining L2 learning. Researchers (e.g., Ellis, Basturkmen and Loewen, 2001; Leeman, 2007) have argued that input alone is insufficient for L2 learning, and engaging in interaction is essential in facilitating L2 learning and development. This is the view held by the interactionist approach. A proponent of this approach, Long (1985) claimed through the Interaction Hypothesis that comprehensible input facilitates acquisition as it provides opportunities to negotiate meaning during a communication breakdown. Later, Long (1996) extended his hypothesis to include interactional feedback and modified output as important contributors to interlanguage development. He explains that the feedback learners receive on their language production when they attempt to communicate could contribute to language acquisition, as this would push them to reformulate their productions to make them comprehensible.

Information-processing perspectives underlie the use of tasks and the importance of modified output. Through her Comprehensible Output Hypothesis, Swain (1995) claims that output serves 3 main functions. It prompts learners to test hypotheses, allows learners to notice gaps in language use and acts as a springboard for metalinguistic awareness. In other words, learners would go beyond semantic processing as they engage in syntactic processing. Swain (1995) argues that "in producing the target language, learners may notice a gap between what they want to say and what they can say [and] this may trigger cognitive processes which might generate linguistic knowledge that is new for learners, or which consolidate their existing knowledge" (Swain, 1995, p. 126). The idea of bringing learners' attention to gaps in their language use is what is termed as "noticing" (Schmidt, 1990). Through his "noticing hypothesis", Schmidt (1990) argues that noticing would facilitate learning and "those who notice most learn most, and it may be that those who notice most are those who pay attention most, as a general disposition or on particular occasions" (Schmidt 1990, p. 144).

Skehan's (2009) Trade-off Hypothesis and Robinson's (2001, 2005, 2007) Cognition Hypothesis are two constructs that dominate the current concept of task complexity. These two hypotheses were partly motivated by information-processing models such as Levelt's (1989) Speech Production Model and by earlier studies of task complexity in L2 research

(e.g., Nunan, 1989; Candlin, 1987). Candlin (1987), for instance, recognised 5 factors for task complexity, one of which is cognitive load. Similarly, Skehan (1996) identified cognitive complexity as a factor in task complexity. One of his key arguments is that attentional allocation must be prioritised because humans have limited processing capacity (Skehan, 1998). He argues that task content and task performance demands are in competition with each other. In light of this argument, high reasoning demand tasks will result in low linguistic performance, as there will be less attention available for linguistic accuracy. On the other hand, Robinson (2001, 2005, 2007), through his Cognition Hypothesis for task complexity, argues that humans possess multiple, non-competition pools of attention. He contends that certain factors of task complexity are not in competition with task performance and attention to linguistic form. Thus, he predicts that increasing task complexity along resource-directing dimensions (i.e., +/- few elements, +/- here & now, +/- no reasoning demand) will result in more accurate and linguistically complex task performance. Also, there will be more negotiation of meaning, and increased attention to linguistic forms.

### 3. The Present Study

The focus of the present study was to investigate the effects of manipulating the reasoning demand component of Robinson's (2001, 2005, 2007) task complexity dimension on L2 production. Specifically, this study sought to answer the following questions:

- a) What main and interaction effects do task reasoning demand and dyadic vs. individual task condition have on grammatical accuracy?
- b) What main and interaction effects do task reasoning demand and dyadic vs. individual task condition have on syntactic complexity?

Following the predictions of the Cognition Hypothesis (Robinson, 2001, 2005, 2007), the following hypothesis was formulated:

The +TRD task and dyadic TC will result in greater grammatical accuracy and syntactic complexity compared to the -TRD task and individual TC.

#### 3.1 Participants

The current study was carried out in a secondary school in Selangor. The participants were seventy-six Form 6 students. They were randomly assigned into four groups. Table 1 summarises the basic profile of the learners.

Table 1: Learner characteristics

|               |         |    |
|---------------|---------|----|
| Gender        | Male    | 32 |
|               | Female  | 44 |
| L1 Background | Malay   | 17 |
|               | Chinese | 48 |
|               | Tamil   | 11 |

### 3.2 Operationalisation of Task Complexity as Task Reasoning Demand (TRD)

Robinson (2001, 2005, 2007) proposed the Cognition Hypothesis which asserts that tasks should be designed and sequenced for the learners on the basis of increases in their cognitive complexity. He further stresses that these design and sequencing decisions should be the basis of the task-based syllabus. Based on these convictions, Robinson (2001, 2005, 2007) developed the Triadic Componential Framework for task design, as outlined in figure 1;

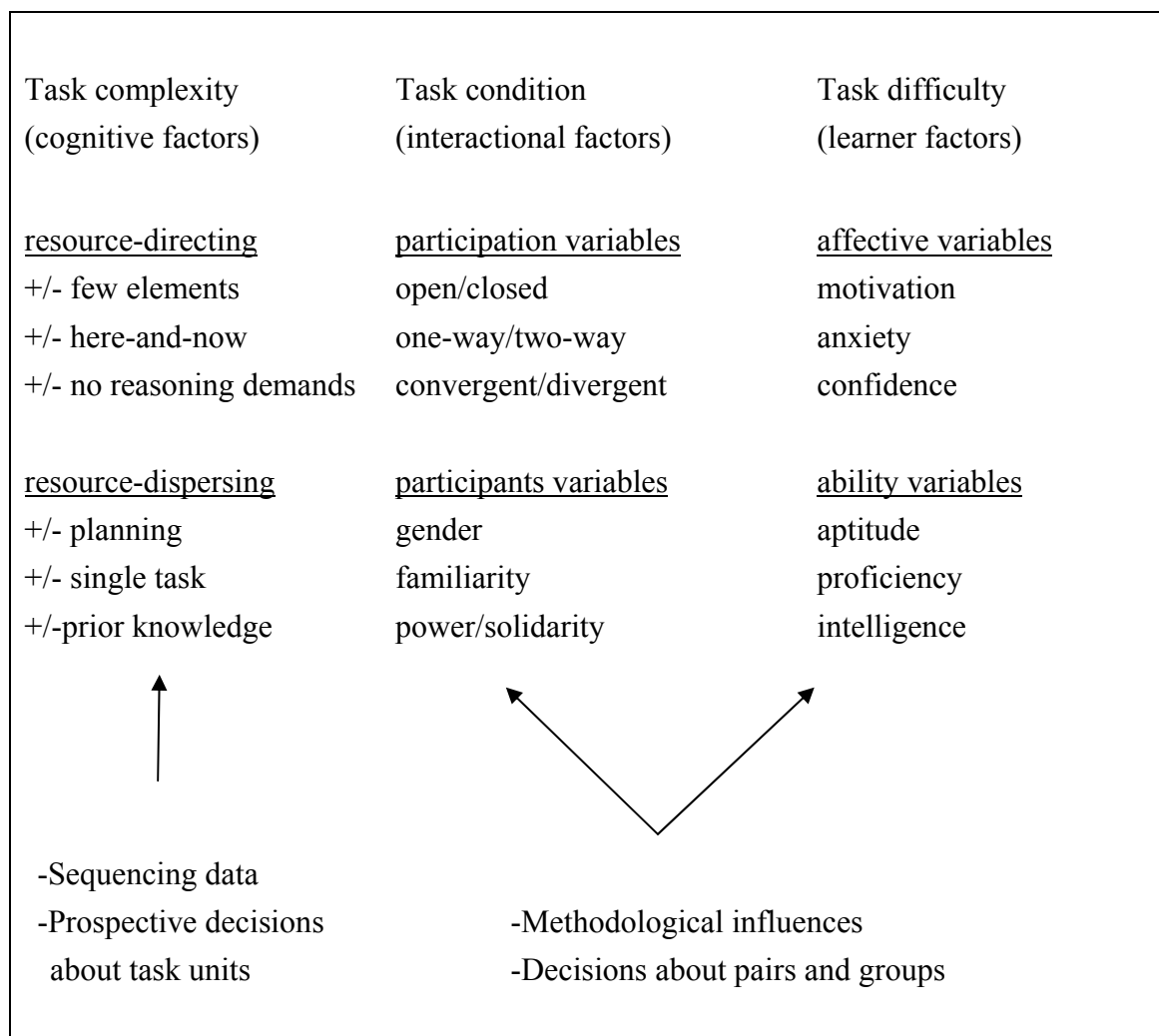


Figure 1: Task complexity, condition and difficulty (Robinson, 2001, p.30)

The framework consists of three factors; task complexity, task condition and task difficulty. The dimensions are represented by the +/- symbols which may represent presence (+) or absence (-), or a relatively greater (+) or relatively less (-) amount. In the case of the present study, '+TRD' represents relatively greater reasoning demand and '-TRD' represents relatively less reasoning demand. Task difficulty involves learner perceptions of task demands. This would be determined by (a) ability factors such as aptitude and working memory, and (b) affective factors such as anxiety and willingness to communicate. Task condition relates to participation and participant factors. Participation factors include one-way or two-way interaction, while participant factors include learner proficiency levels.

The current research is interested mainly in the task complexity aspect of the framework, in particular the resource-directing dimension. Task complexity deals with the inherent cognitive demands of the tasks. Robinson defines task complexity as follows:

Task complexity is the result of attentional memory, reasoning, and other processing demands imposed by the structure of the task on the language learner. These differences in information processing demands, resulting from design characteristics, are relatively fixed and invariant (2001b, p. 29).

Thus, task complexity is closely related to the information processing demand and cognitive load a learner deals with in the course of completing a task. Robinson (1995, 2001) suggests that task complexity variables fall into two groups: resource-dispersing dimensions and resource-directing dimensions. Resource-directing variables are those that make cognitive or conceptual demands on the learners. The subcategories in this dimension are +/- here and now, +/- few elements, and +/- reasoning. Resource-dispersing variables are variables that make performative or procedural demands on the learners. The subcategories include +/- planning time, +/- prior knowledge and +/- single task.

Within task complexity categories, reasoning demand is one of the variables in the resource-directing continuum that has garnered interests among researchers (e.g., Iwashita, McNamara and Elder, 2001; Revesz, 2009). Reasoning demands have been defined as the extent to which learners are required to reason, provide justification, or explain causalities (Candlin, 1987; Nunan, 1989; Skehan, 1996). Thus, a task with high reasoning demands would require learners to extensively engage in reasoning processes to complete the task. Robinson (2005) hypothesised that increasing the reasoning demand of a task can promote the use of syntactic complementation (i.e., learners may need to use logical connectors and cognitive verbs) to provide reason.

In the case of the current study, the low complexity task was a revised version of the dictogloss task, while the high complexity task was an opinion-gap task. The topics of the two tasks were matched. However, the provision of content support in the dictogloss task made it less demanding than the opinion-gap task. During the dictogloss task, a short paragraph was read out twice to learners, and they were allowed to jot down notes during the second reading. Thus, the learners already had a general idea of the content of the text, and they had their written notes to aid them during the text-reconstruction stage. They were then required to reconstruct the text through discussions in dyads. However, during the opinion-gap task, only a topic with two points for arguments was given to the learners. Thus they would have to access their schemata, select relevant information, build on the knowledge through interaction with other learners, express preference and feelings, argue and reason, justify their arguments, and provide cause-effect relationships. Thus, in terms of reasoning demands, the dictogloss task was considered relatively less demanding (-TRD) compared to the opinion-gap task (+TRD).

### *3.3 Task Administration*

One week was spent to introduce the tasks and eleven weeks to implement the tasks. The task administration procedures for the -TRD and +TRD tasks are explained in the following sub-sections.

#### *a) The Dictogloss Task (-TRD)*

The dictogloss or passage reconstruction task (Nabei, 1996; Wajnryb, 1990) consists of a series of stages. Each lesson started with a pre-task activity to ease learners into the topic. As a pre-task activity, learners engaged in a brief question and answer session with the teacher based on the topic of the passage that were to be read to them. After the pre-task activity, the learners listened to a short, dense passage during which they were instructed to listen only and not write anything down. Next, they listened to the passage a second time and were allowed to jot down notes, but not complete sentences. Then, working in pairs, learners pooled their notes together and attempted to reconstruct their own written version of the passage. During the sessions involving learners working individually, instead of pooling notes with a partner and reconstructing the text together, learners worked on it on their own. The written output produced after engaging in the task individually or in dyads was collected for analysis.

#### *b) The Opinion-gap Task (+TRD)*

An opinion-gap task would require students to give their opinions based on given issues in pairs. The topics for the opinion-gap tasks were matched as closely as possible to the topics for the dictogloss tasks. The three main topics were school life, social issues, and environmental issues. The topics were then broken down into a question each for learners to discuss. For example, under the topic “school life”, the question given was for the dyadic TC was “what is the best way to improve the school canteen and why?” Two options were also given “(a) improve on the quality of food served (b) reduce the price of food sold”. In each pair, learner A had to argue in support of option (a), while learner B had to argue in support of option (b). The learners were assigned to specific arguments to ensure that argumentation would take place as the options in the task created a difference in opinion. Students were then engaged in a brief question and answer session as a pre-task activity to help learners access their schemata with regards to the topic and familiarise themselves with the topic. The pre-task activity was the same as the dictogloss task. Then, students were instructed to jot down their ideas individually for about 5 minutes. After that, they discussed their opinions in pairs and wrote the outcome of their discussion in one paragraph of approximately 100 words (similar to the number of words in the dictogloss text). During the treatment involving learners working as individuals, instead of dealing with the task with a partner, learners worked on it individually. The written output produced after engaging in the task individually or in dyads was collected for analysis.

### *3.4 The Repeated-Measures Latin-Square Design*

As the experimental conditions can be re-ordered, a counterbalancing analysis was used where all learners were subjected to each TRD and TC, but in different orders. Treatments

were assigned at random within rows and columns, in order to control for variation in two directions; condition and sequence.

Table 2: Order of administration of treatments

| Group            | Session*   |            |            |            |
|------------------|------------|------------|------------|------------|
|                  | 1          | 2          | 3          | 4          |
| Group 1<br>N= 20 | +TRD, dyad | -TRD, ind. | +TRD, ind. | -TRD, dyad |
| Group 2<br>N= 20 | -TRD, ind. | +TRD, ind. | -TRD, dyad | +TRD, dyad |
| Group 3<br>N= 18 | +TRD, dyad | -TRD, ind. | -TRD, dyad | +TRD, ind. |
| Group 4<br>N= 18 | -TRD, dyad | +TRD, dyad | +TRD, ind. | -TRD, ind. |

\*Each session involved the use of three tasks at a 1 to 3 day interval, TRD= reasoning demand, ind.= individual

Using a Latin-square design, each group is exposed to all combinations of TRD and dyadic vs. individual TC, but in a different order. This helps reduce the chances that changes or outcomes are due to the learners' perception of the tasks (subject characteristic threat). Also, putting the treatments in different orders in the groups helps reduce the threat that change is a result of task-sequencing (Frankel and Wallen, 2008).

#### 4. Data Analysis

The main and interaction effects that the independent variables had on grammatical accuracy and syntactic complexity were measured using two-way repeated measures ANOVAs. The analyses would help show the following;

- the main effect of dyadic vs. individual TC on grammatical accuracy
- the main effect of high and low TRD on grammatical accuracy
- the interaction effect of dyadic vs. individual TC and high and low TRD on grammatical accuracy
- the main effect of dyadic vs. individual TC on syntactic complexity
- the main effect of high and low TRD on syntactic complexity
- the interaction effect of dyadic vs. individual TC and high and low TRD on syntactic complexity

The three aspects of L2 production are complexity, accuracy and fluency (Skehan, 1996, 1998; Ellis, 2003). As such, complexity, accuracy, and fluency have figured prominently in research involving L2 production, L2 performance, and/or L2 learning. They have been used as

indicators of proficiency underlying language performance and as measurements indicating progress in language learning (Housen and Kuiken, 2009). Evidence suggests that complexity and accuracy are primarily linked to the current state of the learner's (partly declarative, explicit and partly procedural, implicit) interlanguage knowledge, whereas fluency is primarily related to learners' control over their linguistic L2 knowledge, as reflected by the speed and ease with which they access relevant L2 information to communicate meaning (Housen and Kuiken, 2009, p.462). Skehan (1996, p.46) defines the three aspects of language production as follows;

- a) Accuracy concerns "the learner's capacity to handle whatever level of interlanguage complexity s/he has currently attained"
- b) Complexity relates to "the stage and elaboration of the underlying interlanguage system"
- c) Fluency involves "the learner's capacity to mobilise an interlanguage system to communicate meanings in real time"

As the current study deals with written data and attempts to gauge declarative knowledge gained as a result of TRD and TC, only accuracy and complexity of L2 production was assessed and compared. The written production in the form of a paragraph by each dyad or individual at the end of each treatment session was analysed in terms of their levels of grammatical accuracy and syntactic complexity.

Housen and Kuiken (2009, p.463) opine that complexity is the most complex, ambiguous, and least understood dimension of the complexity, accuracy and fluency triad. They argue that the term can be confusing as it could refer to both task complexity and L2 complexity. While task complexity refers to the inherent properties of the language task, L2 complexity refers to the properties of language performance and proficiency. The analysis of the written output is thus concerned with L2 complexity which has been interpreted in two ways; as cognitive complexity and linguistic complexity. Cognitive complexity refers to the relative difficulty with which language features are processed in L2 performance and acquisition, while linguistic complexity deals with the size, elaborateness, richness and diversity of the learner's interlanguage system (Housen and Kuiken, 2009, p.464). This study is concerned with analysing the linguistic complexity of written production, in particular, the syntactic complexity.

#### *4.1 Measuring Grammatical Accuracy*

Researchers in SLA have employed several ways to measure accuracy. They include calculating percentages of correct use of targeted linguistic items (Robinson, 1995), percentages of error free clauses or C-units (Robinson, 2007), and calculating the proportion of error free T-units to T-units (Larsen-Freeman, 2006). Although, C-units are similar to T-units, as pointed out by Skehan and Foster (1997), C-units differ in that elliptical answers to questions can also be considered a unit, and are therefore more suitable for interactive discourse. Given that there are no pre-determined linguistic targets in the current research and the data is in the form of written paragraphs and not discourse, calculating the proportion of error free T-units to T-units (Larsen-Freeman, 2006) was considered an appropriate and

effective method of measuring accuracy. The formula used for measuring grammatical accuracy in this study is:

$$\frac{\text{Number of error free T-unit}}{\text{Number of T-unit}}$$

#### *4.2 Measuring syntactic complexity*

Pertinent to this current study is measuring syntactic complexity in a valid, reliable and efficient manner. Previous studies have measured syntactic complexity using several methods. The deciding factors for the measurements used include the type of data (e.g., written or spoken discourse), the amount of data, the age group (e.g., children or adults), and whether grammar items were pre-targeted or otherwise. For the purpose of measuring L2 written or spoken production, the C-unit and T-unit as units of analysis are the most often used measures of syntactic complexity (e.g., Kuiken and Vedder, 2007; Larsen-Freeman, 2006, Robinson, 2007). For example, to measure the syntactic complexity of learners' speech production while telling stories, Robinson (2007) used measurements of clauses per C-unit (CPC). C-units or communication units would include words, phrases, onomatopoeic formulations, or abbreviations with communicative value, while a T-unit is defined as an independent clause and any subordinate clause attached to it or embedded in it (Robinson, 2007). Speech units (AS-unit) have also been used to measure syntactic complexity, particularly in analysing highly interactive spoken discourse (e.g., Foster, Tonkyn & Wigglesworth, 2000; Michel, Kuiken & Vedder, 2007).

Following Larsen-Freeman (2006), this study used the average number of clauses per T-unit as a measure of syntactic complexity. In this study the T-unit is preferred over the C-unit, since it deals with written paragraphs and not spoken discourse. The formula used in determining syntactic complexity is:

$$\frac{\text{Number of clauses}}{\text{Number of T-units}}$$

## **5. Results**

Two research questions and a hypothesis were presented in this study and they are addressed as follows.

### *5.1 Research Question 1: The Effects of TRD and TC on Grammatical Accuracy*

The grammatical accuracy of the paragraphs the learners produced after engaging in tasks in dyads or individually were measured using the proportions of error free T-units over T-units. The means of the accuracy scores in the four conditions are reported in table 3.

Table 3: Descriptive statistics: Accuracy scores (Proportion of error-free T-units to T-units)

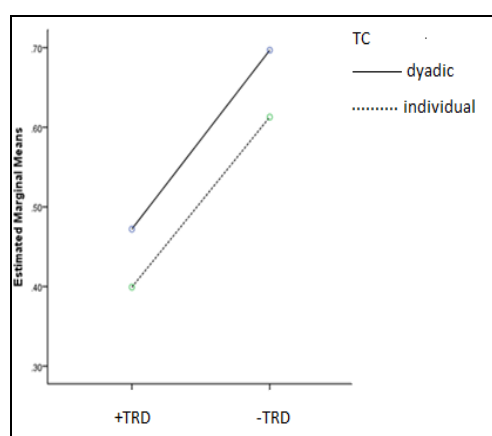
| Group |            | Mean | SD  | Skewness | Kurtosis |
|-------|------------|------|-----|----------|----------|
| +TRD  | Dyad       | .47  | .22 | .74      | -.47     |
| N= 76 | Individual | .40  | .18 | .74      | -.23     |
| -TRD  | Dyad       | .70  | .16 | -.20     | .16      |
| N= 76 | Individual | .61  | .20 | .10      | -.96     |

+TRD= high reasoning demand task, -TRD= low reasoning demand task

Table 3 shows the descriptive statistics of the accuracy scores the learners obtained in their written task while engaging in high reasoning demand tasks and low reasoning demand tasks in dyads and individually. The highest mean was recorded in the accuracy score when learners were engaged in –TRD tasks in dyads ( $M = .70$ ), while the lowest mean was recorded in the accuracy score when learners were engaged in the +TRD task individually. The significance of these results was analysed using a two-way ANOVA with repeated measures. The main and interaction effects of TRD and TC on grammatical accuracy are presented in Table 4 and Figure 2.

Table 4: Repeated-measures ANOVA: Accuracy (Proportion of error-free T-units to T-units)

| Source          | df | SS   | MS   | F      | p   | Partial eta <sup>2</sup> | Observed power |
|-----------------|----|------|------|--------|-----|--------------------------|----------------|
| Within subjects |    |      |      |        |     |                          |                |
| RD              | 1  | 3.65 | 3.65 | 117.28 | .00 | .61                      | 1.00           |
| TC              | 1  | .47  | .47  | 13.60  | .00 | .15                      | .95            |
| RD*Dyad         | 1  | .00  | .00  | .06    | .80 | .00                      | .06            |



+TRD= high task reasoning demand, -TRD= low task reasoning demand

Figure 2. Interaction effect between TRD and dyadic vs. individual TC on accuracy scores

The  $2 \times 2$  ANOVA with repeated measures was used as there were two levels of TRD and two levels of TC. Additionally, all learners were exposed to all four condition combinations. Also, there were no between-subjects variables present. As there were only two

levels for each within-subjects factor, the assumption of sphericity does not apply. Table 4 shows that level of task reasoning demand had a significant effect on the accuracy scores obtained ( $F= 117.28$ ,  $p= .00$ ). Figure 2 shows that the -TRD task resulted in significantly higher accuracy scores compared to the +TRD task.

Dealing with tasks in dyads or individually also had a significant effect on accuracy scores ( $F= 13.60$ ,  $p= .00$ ). Figure 2 shows that dealing with tasks in dyads produced higher accuracy scores compared to dealing with tasks individually in both +TRD and -TRD contexts. Conversely, the interaction between TRD and dyadic vs. individual TC did not have a significant effect on accuracy scores ( $F= .06$ ,  $p= .80$ ). The mean scores for accuracy of the two variables showed a parallel pattern in figure 2, indicating a lack of significant interaction between TRD and dyadic or individual TC. In other words, the accuracy scores due to TC were not affected by whether learners were dealing with +TRD or -TRD tasks.

## 5.2 Research Question 2: The Effects of TRD and TC on Syntactic Complexity

Written paragraphs that were analysed for grammatical accuracy using proportion of error-free T-units to T-units were also analysed for syntactic complexity using proportion of clauses per T-unit. The means of the syntactic complexity scores in the four conditions are reported in table 5.

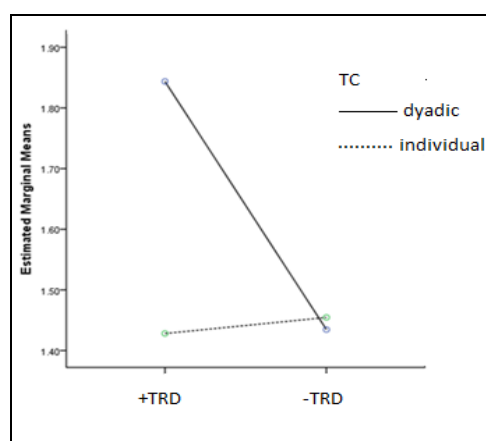
Table 5: Descriptive statistics: Syntactic complexity (Proportion of clauses per T-unit)

| Group    |            | Mean | SD  | Skewness | Kurtosis |
|----------|------------|------|-----|----------|----------|
| High TRD | Dyad       | 1.84 | .46 | -.18     | -1.27    |
|          | Individual | 1.43 | .29 | .34      | -.80     |
| Low TRD  | Dyad       | 1.43 | .29 | .36      | -.82     |
|          | Individual | 1.45 | .21 | .40      | -.81     |

Table 5 shows the descriptive statistics of the syntactic complexity scores the learners obtained in their written task while engaging in high reasoning demand tasks and low reasoning demand tasks in dyads and individually. The highest mean was recorded in the accuracy score when learners were engaged in the +TRD task in dyads ( $M= 1.84$ ), while the lowest mean was recorded when learners were engaged in the +TRD task individually and -TRD tasks in dyads ( $M= 1.43$ ). The significance of these results is analysed using a two-way ANOVA with repeated measures.

Table 6: Repeated-measures ANOVA: Syntactic complexity (Proportion of clauses per T-unit)

| Source          | df | SS   | MS   | F     | p   | Partial eta <sup>2</sup> | Observed power |
|-----------------|----|------|------|-------|-----|--------------------------|----------------|
| Within subjects |    |      |      |       |     |                          |                |
| TRD             | 1  | 2.97 | 2.97 | 39.46 | .00 | .35                      | 1.00           |
| Dyad            | 1  | 2.78 | 2.78 | 33.78 | .00 | .31                      | 1.00           |
| TRD*Dyad        | 1  | 3.60 | 3.60 | 30.08 | .00 | .29                      | 1.00           |



+TRD= high task reasoning demand, -TRD= low task reasoning demand

Figure 3: Interaction effect between TRD and dyadic vs. individual TC on syntactic complexity scores

Table 6 shows that level of reasoning demand had a significant effect on the complexity scores obtained ( $F = 39.46$ ,  $p = .00$ ). Dealing with tasks in dyads or individually also had a significant effect on complexity scores ( $F = 33.78$ ,  $p = .00$ ). Unlike the interaction for the accuracy scores, the interaction between TRD and dyadic vs. individual TC for complexity scores has a significant effect ( $F = .30.08$ ,  $p = .00$ ). The plot in Figure 3 represents a situation where the effect of reasoning demand did not make much difference when learners were engaged in tasks individually. However, when learners were working in dyads, high reasoning demand tasks helped improve syntactic complexity scores in learner production. In other words, working in dyads seemed to have had a positive effect on complexity scores when learners were engaged in the +TRD task.

*5.3 Hypothesis Testing: The +TRD task and dyadic TC will result in greater grammatical accuracy and syntactic complexity compared to the -TRD task and individual TC*

Results of the current study indicated that TRD, as predicted by Robinson's Cognition Hypothesis (2001, 2005, 2007) and Skehan's Trade-off Hypothesis (2009), can affect the extent of grammatical accuracy in learner output. Dyadic and individual learning conditions also had a significant effect on grammatical accuracy. More specifically, learners working in dyads produced a significantly higher mean score for grammatical accuracy as opposed to

learners working individually ( $p = .00$ ). Also, -TRD produced a higher mean score for accuracy as opposed to +TRD ( $p = .00$ ). Meanwhile, the interaction between TRD and dyadic vs. individual TC did not have a significant effect on grammatical accuracy scores ( $p = .80$ ). This would indicate that learner individual or dyadic performance in terms of grammatical accuracy was not affected by whether the learners were dealing with the high or low reasoning demand task.

In terms of syntactic complexity, descriptive statistics indicated that the highest mean was achieved by +TRD done in dyads. The repeated-measures ANOVAs showed that reasoning demand had a significant effect on syntactic complexity. Similarly, individual and dyadic TCs had a significant effect on syntactic complexity. Unlike grammatical accuracy, the interaction of reasoning demand and dyadic vs. individual TC had a significant effect on syntactic complexity. The results showed that +TRD resulted in a significantly higher mean of syntactic scores compared to the -TRD task. Also, working in dyads resulted in a significantly higher mean score compared to working individually. In addition, working in dyads on +TRD had a significantly positive effect on syntactic complexity scores compared to working individually on the -TRD task.

Results of the current study seem to partly support the assumptions of the Cognition Hypothesis (2007) that “more complex tasks, along dimensions which direct the cognitive/conceptual effort they require to linguistic resources that can enable them to be performed should result in more accurate and complex ... speech production” (Robinson, 2007, p.193). More specifically, the results indicated that +TRD resulted in greater syntactic complexity than the -TRD task, but significantly less grammatical accuracy compared to the -TRD task. It could be argued that accuracy gains were at the expense of linguistic complexity. This argument would be in line with Skehan’s Trade-off Hypothesis. A relatively low reasoning demand allows learners to direct their cognitive resources to the production of more accurate language. Skehan (1996, 1998) argued that by decreasing task cognitive demands, learners could channel their attention to language use.

In terms of L2 accuracy, results of the current study are comparable to the claims made by Skehan (1998) and Skehan and Foster (1997). These studies also found that tasks with less cognitive demands resulted in greater accuracy in L2 production. For example, Skehan and Foster (1997) in a study using narrative tasks found that the planned version (i.e., less complex version) resulted in higher accuracy but less complexity. They concluded that the results clearly demonstrate a trade-off between accuracy and complexity. Menhert (1998, p.104) concurs when he concluded that “any gains in accuracy and complexity are not simultaneously achieved. This is because attentional resources are limited and therefore allocated either towards achieving accuracy or complexity at any one time”. Therefore, the hypothesis of the current study is only partially supported. The dyadic TC did result in greater grammatical accuracy and syntactic complexity compared to individual TC. However, the +TRD task only resulted in greater syntactic complexity, but not greater grammatical accuracy compared to the -TRD task.

## 6. Discussion and Conclusion

Inferential statistics applied to the grammatical accuracy and syntactic complexity scores showed that the dyadic TC resulted in significantly higher mean scores compared to the individual TC. This would be in support of the Interaction Hypothesis (Long, 1990) and Swain's Output Hypothesis (1985, 1998). Long (1990) argues that a range of interactional processes that occur during the course of completing tasks such as negotiating meaning, modifying output and getting feedback could contribute towards L2 development. This could mean that interacting with their peers helped learners to work out linguistic problems, thus resulting in higher accuracy scores compared to working individually. In the same vein, Swain (1998) attributes L2 acquisition to collaborative dialogue, during which learners negotiate meaning and build knowledge on language use. This could be further explained by Kowal and Swain (1994) who point out that "if a task can be devised to have learners talk about the language they are producing, their talk may well serve the function of raising their awareness of forms, rules and their relationship to the meaning they are trying to express" (p. 5). Thus, collaborative language production tasks may prompt learners to deepen their knowledge of linguistic rules.

With the use of tasks in the classroom, teachers have an added role of critically evaluating the tasks to ensure that they could trigger negotiation and conscious reflection of language. Tasks such as the dictogloss could also provide modelling of language use. Teachers could then manipulate the texts used in the dictogloss to suit the learners' specific needs and proficiency levels. In other words, content-based lessons such as the ones that use tasks as the organising principle need to be geared towards getting students to pay attention to language use other than content. Therefore, learners must be encouraged to be precise and coherent in their production, which means they must pay attention to the content of their output as well as the language used to convey the content.

Considering the results of the current research, in light of Robinson's Cognition Hypothesis (2007) and Skehan's Trade-off Hypothesis (2009), teachers must also take into consideration the effects of task reasoning demand on accuracy and syntactic complexity on learner output, when designing or selecting tasks of varying complexity levels for use in the classroom. Nonetheless, due to the characteristics of the participants and the classroom-based nature of the current study, findings cannot be generalised to other populations or teaching contexts. Thus, replication studies in a variety of contexts would be highly desirable. There is certainly a need to study the effects of task complexity and task conditions, and the effects of their interaction on different aspects of L2 learning with different groups of learners. The current study operationalised task complexity based on one variable (i.e., task reasoning demand) and included one learner participation variable (i.e., individual vs. dyadic) as the task condition. To gain greater insights and more extensive data, there is a need for more classroom-based experiments that aim to test Robinson's Cognition Hypothesis (2007) with a variety of tasks at various degrees of task complexity and in different learning conditions. Such studies could help teachers make informed decisions about task design, task selection and task implementation in the L2 classroom.

## References

- Candlin, C. (1987) Towards task-based language learning. In C. Candlin & D. Murphy (Eds.), *Language learning tasks*. Prentice-Hall, Englewood Cliffs, NJ.
- DeKeyser, R.M. (1998). Beyond focus on form: Cognitive perspective on learning and practicing second language grammar. In C. Doughty & J. Williams, J. (Eds), *Focus on form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.
- Ellis, R. (2003). *Task-based language teaching and learning*. Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., & Loewen, S (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51, 281-318. <http://dx.doi.org/10.1111/1467-9922.00156>
- Foster, P., Tonkyn, A. and Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Applied Linguistics*, 21(3), 354-375. <http://dx.doi.org/10.1093/applin/21.3.354>
- Housen, A., & Kuiken, F. (2009) Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461-473. <http://dx.doi.org/10.1093/applin/amp048>
- Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of information processing approach to task design, *Language Learning*, 51(3), 401-436. <http://dx.doi.org/10.1111/0023-8333.00160>
- Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3(2), 73-93. <http://dx.doi.org/10.1080/09658416.1994.9959845>
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Kuiken, F., & Vedder, I. (2007) Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics*, 45, 261-284. <http://dx.doi.org/10.1515/iral.2007.012>
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619. <http://dx.doi.org/10.1093/applin/aml029>
- Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In R. DeKeyser (Ed.), *Practice in a second language*. Cambridge: Cambridge University Press.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, Mass: MIT Press.
- Long, M. (1985). Input and second acquisition theory. In S. Gass & C. Madden (Eds.), *Input in second language acquisition*. Rowley, MA: Newbury House.

- Long, M. (1990). Second language classroom research and teacher education. In Brumfit, C., & Mitchell, R. (Eds.), *Research in the language classroom*. ELT Documents 133, 161-70.
- Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective*. Amsterdam: John Benjamins.
- Long, M. (1996). The role of linguistic environment in second language acquisition. In W.C. Ritchie & B.K. Bahtia (Eds.), *Handbook of second language acquisition*. New York: Academic Press.
- Michel, M. C, Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *IRAL*, 45, 241-259. <http://dx.doi.org/10.1515/iral.2007.011>
- Nabei, T. (1996). Dictogloss: Is it an effective language learning task? *Working Papers in Educational Linguistics*, 12(1), 59-74.
- Nunan, D. (1989). *Designing tasks for a communicative classroom*. Cambridge: Cambridge University Press.
- Revesz, A.J. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, 31(3). pp. 437-470. <http://dx.doi.org/10.1017/S0272263109090366>
- Robinson, P. (2001). 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*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524780.012>
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL*, 43, 1-32. <http://dx.doi.org/10.1515/iral.2005.43.1.1>
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL*, 45, 193-213. <http://dx.doi.org/10.1515/iral.2007.009>
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 17—46. <http://dx.doi.org/10.1093/applin/11.2.129>
- Skehan, P. (1996). A framework for the implementation of task based instruction. *Applied Linguistics*, 17, 38-62. <http://dx.doi.org/10.1093/applin/17.1.38>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2009). Modelling Second Language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532.

<http://dx.doi.org/10.1093/applin/amp047>

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

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.

Wajnryb, R (1990). *Grammar dictation*. Oxford: Oxford University Press.

### **Copyright Disclaimer**

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).