Modern English Education, Vol. 15, No. 4, Winter, 2014

The Effects of Different Tasks on English Language Learners' Speaking Complexity, Accuracy, and Fluency*

Mi-Hye Kim
Hankuk University of Foreign Studies

Kim, Mi-Hye. (2014). The effects of different tasks on English language learners' speaking complexity, accuracy, and fluency. *Modern English Education, 15*(4), 99-116.

This study investigated the effects of different task types on the speaking complexity, accuracy, and fluency of Korean university students who were learning English. Seventy students taking a course in English pronunciation were asked to record and upload their verbal answers to various task prompts. The chosen task types were argumentative, descriptive, and narrative, and the researcher divided the task types into these three groups based on the characteristics of each task type. Topic familiarity was controlled as an intervening variable. The results of analysis showed that the descriptive task was influential on the level of learners' accuracy. The argumentative task was influential on the learners' fluency and complexity in speaking performance. The results suggest that different task types may influence learners' speaking performance, so teachers should be keen at choosing appropriate types of tasks for a particular teaching purpose. To check accuracy in speech production for university level EFL learners, teachers can use descriptive tasks. To measure speech complexity and fluency, argumentative tasks can be employed.

**[argumentative task/descriptive task/narrative task/
토론식 과제/묘사식 과제/설명식 과제]**

I. INTRODUCTION

Because *task-based language teaching* (TBLT) is considered an important teaching method in English as a foreign language (EFL) classrooms, various researchers have investigated tasks used in these classrooms (H. Brown, 2004; Bygate, 1996; Ortega,

* A part of the study was presented at 2014 ETAK International Conference.

1999; Thornbury, 2005; Wendel, 1997; Wigglesworth, 1997). Given that the research on speaking tasks is gaining popularity, it is critical to understand the effects of different speaking tasks on the *complexity*, *accuracy*, and *fluency* of students' speech (Robinson, 2001; Skehan, 1998). Much research has been conducted on the cognitive dimensions of educational tasks; however, although some researchers (Robinson, 2001; Skehan, 1998) have claimed that *topic familiarity* in speaking tasks also affects performance, this dimension has been mostly ignored. Therefore, this study aims to investigate task performance across topics.

Although various kinds of tasks have drawn researchers' and practitioners' attention, previous studies have been done from a curriculum development perspective but not from the perspective of the complexity, accuracy, and fluency of output. Because learners' speaking performance may vary depending on different task types, understanding the effect of a particular task-type on learners' performance is clearly important (Tavakoli & Foster, 2011). Moreover, there are few studies of university-level students in this regard. Therefore, the present study may be helpful for choosing relevant speaking task types for university-level classrooms and may help material developers to create proper textbooks for teaching speaking in EFL settings. The researcher divided the task types into three groups including argumentative, descriptive, and narrative based on the characteristics of each task type. The guidelines used by this study to categorize speaking tasks were as follows. First, *argumentative tasks* are those where learners can integrate their speaking ability with their prior knowledge and experience on the topic of the task (Duff, 1986). Second, *descriptive tasks* are those that use pictures and ask learners to speak about them. Learners can easily grasp key ideas from the pictures before starting their speaking task, and can thus lower their speaking anxiety (Skehan, 1998). Third, *narrative tasks* are those in which learners talk about their memories or previous experiences (Foster & Skehan, 1996).

II. LITERATURE REVIEW

The role of tasks in the EFL curriculum has been debated for over 20 years. Willis (1996) argues that tasks should be used as the focus of the syllabus within a supportive framework that enables language development. This is TBLT in its strictest version. However, in a *task-supported approach*, a weak version, tasks are used to enrich the syllabus or to provide additional learning opportunities, including online support and selective feedback (Brumfit, 1984; Bygate, 2000; Ellis, 2003; Nunan, 1995, 2001; Samuda & Bygate, 2008; Ur, 1982). The present study takes a task-supported

approach because the institution at which the work was carried out did not permit whole-scale changes to the curriculum.

There are several speaking task types that have been tested more directly through empirical research, for example, *information-gap tasks* (Doughty & Pica, 1986; Fotos & Ellis, 1991; Yule & Macdonald, 1990), *problem-solving tasks* (Duff, 1986; Lynch, 1991), *debate-style opinion exchange* (Duff, 1986), *decision-making tasks* (R. Brown, 1991; Foster & Skehan, 1996), and *narrative tasks* (Foster & Skehan, 1996). This study investigates learner output in argumentative, descriptive and narrative speaking tasks. *Pictures and picture story tasks* (Pattison 1987) include using pictures for communication. Learners can find differences using two pictures or make stories using pictures. According to Pattison (1987), pictures can be used for communication in the classroom. Pictures can help learners organize their ideas and lower their anxiety given that pictures can narrow down word choices for a conversation.

The idea of using complexity, accuracy, and fluency to analyze students' speaking performance is not new (Ellis & Barkhuizen, 2005). Though the definitions of these three qualities vary, these three elements remain important, as they constitute overall language proficiency and can thus represent learners' performance (Ellis, 2009).

Complexity is traditionally the most complex, most ambiguous, and least understood dimension of speaking performance. The term refers both to properties of the language task and to properties of L2 performance and proficiency, that is, to *task complexity* and *L2 complexity* (Robinson, 2001; Skehan, 2001). This study does not handle task complexity but focuses on the complexity of learner output, defined "as the capacity to use more advanced language" (Ellis, 2009, p. 475). Fluency also includes two kinds of measures: first is *temporal fluency*, represented by, for example, *number of syllables per minute* and *repairing phenomena* such as false starts, repetitions, and reformulations. According to Skehan (1992), fluency can be defined as the capacity to use language in real time, to emphasize meanings, possibly drawing on more lexicalized systems. Additionally, he defines accuracy as the ability to avoid errors in performance, possibly reflecting higher levels of control in the language as well as a conservative orientation, that is, avoidance of challenging structures that might provoke error (Skehan, 1992). Bygate (1996) claims that accuracy is based on the percentage of errors, lexical selection and use of collocations. Linguistic accuracy refers to how well the language is produced in relation to the rules of the target language (Skehan, 1996) and has been measured in previous studies by global indices such as error-free clauses (e.g., Crookes, 1989).

Reflecting these points, the research question for the present study is as follows: Are there differences in speaking performance (in terms of complexity, accuracy, and fluency) among argumentative, narrative and descriptive tasks?

III. RESEARCH METHOD

1. Participants

Seventy Korean undergraduates taking an EFL conversation class participated in this study. By gender, there were 50 females (71. 4%) and 20 males (28. 6%). In terms of academics, 37 students (52. 9%) were majoring in English, and 33 (47. 1%) in other fields. The higher ratios for English major participants as well as female students can be considered a limitation of this study. The researcher was their course instructor.

2. Procedures

This study took place from September to December 2011 at a university in Seoul. The students talked in small groups of four or five members on a previously chosen task for 20 minutes in class, after which each group leader led the class in group work. After class, each student uploaded a recorded spoken response on the task within 12 hours so that other group members could listen to it. The number of each task type used in the present study varied among students since the speaking tasks for this study were chosen by learners. The learners preferred the speaking tasks which were related to their own experience and previous memories followed by argumentative task types. The descriptive tasks were the least preferred speaking tasks, so the learners avoided choosing the descriptive speaking tasks. The different number of task types for each speaking task could be a limitation of the study. Table 1 shows four different speaking topics for this study.

TABLE 1

Topics for Speaking Tasks

Week	Topic	Task type	Week	Topic	Task type
1st	Self-introduction	Narrative	5th	Suggesting to foreigners good places to visit in Korea	Narrative
2nd	Describing impressive places	Narrative	6th	Describing five different pictures	Descriptive
3rd	Thoughts on plastic surgery	Argumentative	7th	Comparing advantages and disadvantages of cars and bicycles	Argumentative
4th	Describing daily life	Narrative	8th	Thoughts about Korean English education	Argumentative

The tasks were divided into three different types. First, narrative tasks were employed by utilizing learners' individual experiences. In other words, talking about

learners' own previous experiences was considered as a narrative task. Second, an argumentative task was used when the learners were given a chance to discuss something or express their ideas related to the topics. The topics could include pros and cons of certain issues. Therefore, learners could express their own ideas freely. Third, descriptive topics included having learners describe pictures. In this study, learners were not to use their imagination to describe the picture. The learners were only to use the objective facts that they could see in the pictures. Therefore, their words had to be related to the picture.

The number of task-types used by each student in the present study varied since the speaking tasks for this study were chosen by learners. The learners preferred the speaking tasks which were related to their own experience and previous memories followed by argumentative task types. The descriptive tasks were the least preferred speaking tasks so the learners avoided choosing the descriptive speaking tasks. The different number of task types for each speaking task could be a limitation of the study. However, it was assumed that the tendency toward narrative tasks and argumentative tasks may well reflect the current status of task portions in a typical ELT textbook for university-level learners. Therefore, the analysis was carried out to find a possibly generalizable aspect of the task effect.

3. Speaking Analysis

Students' recordings were classified and scored for complexity, accuracy, and fluency. The spoken recordings were transcribed by a native speaker of English who was trained and experienced in transcribing non-native speech. The researcher then checked each transcription for scoring.

1) Accuracy, Complexity, and Fluency Measurements

In this study, *accuracy* was measured by the ratio of error-free clauses to total clauses, following Crookes (1989) and Foster and Skehan (1996), and the percentage of verbs correct within each narrative, following Wendel (1997). Only absolute errors based on "Standard American English were counted as errors" (Ellis & Barkhuizen, 2005, p. 59). Errors in syntax, morphology, and lexical choice were coded; however, pronunciation errors were not. "When a student made an accurate self-correction, that unit was considered error-free", following Ellis and Barkhuizen (2005, p. 49).

For the *complexity measurement*, the Analysis of Speech unit (AS-unit), which is a sentence-length utterance defined for oral language (Foster, Tonkyn, & Wigglesworth, 2000) was determined to be the most appropriate base unit for the data. Following

Norris and Ortega's (2009) suggestion, each speech was measured for syntactic complexity at three different levels. Length of AS-unit and clause length were measured in words. Subordination was calculated as clauses (finite and non-finite) divided by AS-units. Following Skehan's (2009) suggestion, each speech was also measured for lexical variety. According to Foster and Skehan (1996), complexity reflects how learners can use the forms closer to the cutting edge of interlanguage development and is associated with learners' willingness to take risks to use language with which they are familiar. Syntactic complexity, syntactic variety, and lexical variety are the measures that have been widely used (Crookes, 1989) to measure how complex language is used for production.

Syntactic complexity refers to how elaborate the sentence structure is, and is calculated as the ratio of clauses to t-units used in the subjects' production. Clauses are either simple independent finite clauses, dependent finite clauses, or nonfinite clauses. A t-unit is "one main clause plus whatever subordinate clauses happen to be attached or embedded with it" (Hunt, 1966, p. 735). Traditionally, the units used to measure syntactic complexity are c-units, with a focus on the semantic aspect (e.g., Foster & Skehan, 1996) and t-units, with a focus on the syntactic aspect (e.g., Wendel, 1997). The survey conducted by Foster, Tonkyn, and Wigglesworth (2000) revealed that the t-unit is clearly the most popular unit for the analysis of both written and spoken data; however, the c-unit has been chosen by some researchers because t-units are inadequate for coding utterances such as elliptical utterances, which are common in daily interactions. Since the data collected for this study were monologues, retellings of stories that would not contain many elliptical utterances compared to dialogues, t-units seemed to be more appropriate than c-units in the present case, following Vercellotti's (2012) study.

In the transcriptions, the boundary of a t-unit is marked by an upright slash (|). A clause boundary within a t-unit is marked by a double colon (::). Repetitions, reformulations, and false starts are put inside brackets { }. This is borrowed from Foster, Tonkyn, and Wigglesworth (2000); this study also adapted their measurement for complexity. These examples are from students' descriptive task utterances in this study.

Examples:

[1] | A woman is watching very {very} beautiful scene through the window.|

[2] | He is looking at {looking for} his carrier near the place then :: where it should come out.|

[3] | He is putting his bags into the trunk :: while a taxi driver helps him and says something to him with his pleased sentiment.|

[4] | He is looking for his wallet :: while the hotel staff is waiting for his credit card.|
[5] |A driver is watching a man :: who is taking the back seat.|
[6] |I hope :: to study English in America|

Example 1 is a simple t-unit that has a subject, a verb, and a complement. Example 2 is a complex t-unit containing a main clause and a subordinating adverbial clause. Example 3 is another complex t-unit with a main clause and a conjunction clause. Example 4 is another complex t-unit with a main clause and conjunction clause. Example 5 is another complex t-unit, but containing a main clause and a relative pronoun. The last example demonstrated the classification of non-finite clauses. This sentence can be considered to constitute two clauses and one AS-unit since the non-finite clause functions like an additional clause in this sentence. However, non-finite clauses can be difficult to classify (Foster, Tonkyn, & Wigglesworth, 2000). For example, in the utterance "I like reading," "reading" can be analyzed as a minimal non-finite clause. Therefore, this sentence can be considered to be one clause. Another example of this would be the sentence "I want an opportunity to visit," which can also be considered one clause.

In the present study, *fluency* was measured by *phonation time ratio* (PTR) as a general measure (De Jong & Perfetti, 2011). Phonation time ratio (TPO) can be defined as speaking time (excluding filled pauses) divided by total time. (Vercellotti, 2012) Following De Jong and Perfetti (2011), pause length was calculated as the average length of all pauses.

2) Scoring

To measure a participant's speaking performance, a scoring method was used. The highest possible score for each criterion was five, and the lowest possible score was one. For this study, subordinate conjunctions and relative pronouns, which can link two different sentences, are included for complexity. There are many researchers who have included subordinate conjunctions and relative pronouns for complexity. However, the researcher modified the score for complexity since most Korean students are less proficient compared to other native speakers. The student who used the subordinate conjunctions and relative pronouns most frequently received the highest score of 5 for complexity, and students who used the subordinate conjunctions and relative pronouns less frequently received lower scores.

The accuracy score for this study includes grammatical errors and word choice for the topic. The students who made fewer grammatical errors received the highest score for accuracy. The score for accuracy also was modified by the researcher because the

aim for this research was to compare the participants in the same situation. Therefore, the researcher adapted the idea of accuracy and factors for accuracy measurement.

The fluency score is mainly related to total utterance length. The researcher gave all the participants the same time limit and requested the same recording method, so it was reasonable to measure the total length of the utterance for fluency. The results show that even the same students produced the different lengths of utterances depending on the topic. Therefore, the researcher assumed the topic influenced the students speaking outcomes.

The *accuracy score* for this study includes grammatical errors and word choice for the task. Students who made fewer grammatical errors got the highest score of 5 for accuracy. The score for accuracy was modified by the researcher because the aim for this research was to compare the participants in the same situation.

For the present study, subordinate conjunctions and relative pronouns that can link two different sentences are included for *complexity.* It is possible to include subordinate conjunctions and relative pronouns for measuring complexity. However, the researcher modified the score for complexity since most Korean students seemed to be less proficient compared to native speakers. Students who used subordinate conjunctions and relative pronouns more than five times for one speech were considered to have a high level of complexity and got the highest score of 5 for complexity, and students who used subordinate conjunctions and relative pronouns less frequently got lower scores.

The *fluency score* is mainly related to the total utterance length. The researcher gave all the participants the same time limit and requested the same recording method, so it was reasonable to measure the total length of the utterance for fluency.

3) Rater Reliability

Five speaking performances were randomly selected from the corpus to determine inter-rater reliability. A trained rater and the researcher coded the data using the measures described above. Whenever the scores exceeded the acceptable range of more than three points difference in scores for suggested assessment measurements, we discussed the reasons and shared explanations for giving our respective scores and added additional details for each assessment category. This discussion was performed as part of examiner training to increase the reliability of this study. After the discussion, we graded some students' speaking performances again and decreased our scoring differences. Inter-rater reliability was determined by looking at the percentage of agreement for each rating category between the raters. In each speaking performance, a high level of agreement (over 95%) was reported.

4. Topic Familiarity

For further analysis, the possible influence of individual learner's topic familiarity on speaking performance was checked. The topic familiarity was measured by a questionnaire. The questionnaire was divided into two parts. The first part covered demographic characteristics, such as gender, major, and experience living in English-speaking countries. The second part covered students' topic familiarity. A seven-point Likert-scale was employed, indicating 7 as extremely familiar. The learners were requested to rate their familiarity of each of item, such as "The topic'thoughts about plastic surgery' is familiar to me in the context of studying English." The questionnaire was administered and collected in the classroom by the teacher-investigator. For analysis, the researcher divided students into 'High' (scores 5–7) and 'Low' (scores 1–3) topic familiarity groups based on their questionnaire answers (that is, on how familiar the topic at hand was). Students with a score of 4 were disregarded from the analysis.

IV. RESULTS

This section presents the answer for the research question: Are there differences in speaking performance (in terms of complexity, accuracy, and fluency) among argumentative, narrative and descriptive tasks?

1. Argumentative, Narrative, and Descriptive Tasks

The following figure displays the results of speaking performance in terms of complexity, accuracy, and fluency among argumentative, narrative and descriptive tasks. The scores indicate the means of 70 students for three argumentative tasks, four narrative tasks and one descriptive task.

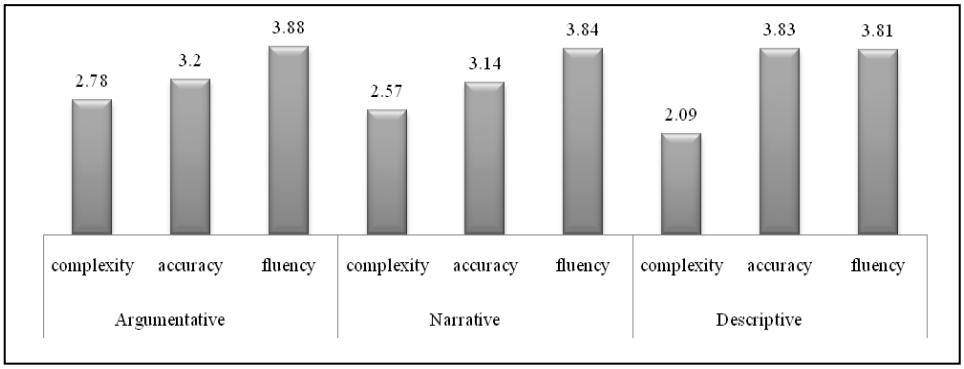

FIGURE 1 Speaking Performance of Three Different Tasks

When comparing the scores for *complexity* on three different tasks, the *argumentative task* was the highest, followed by the narrative task, and descriptive task. According to Figure 1, the mean score for complexity on the argumentative tasks was 2.78, and the mean score for complexity on the descriptive tasks was 2.09. These two tasks show the significant difference on complexity scores. For *fluency* scores, the same pattern was noticed; the *argumentative task* was the highest, followed by the narrative task, and descriptive task. The reason for the highest complexity score for the argumentative task seems to be that the learners believed it was necessary to use more complicated sentences to express their ideas on the argumentative task. The present results are somehow different from those in Skehan and Foster's (1997) study. In their study, they found that the narrative task with planning time was associated with fluency and accuracy, but not with complexity. The narrative task in their study involved describing a story represented in a cartoon strip with a clear, linear story line. In the present study, however, the argumentative task led to better speaking performance than the narrative tasks in terms of complexity, accuracy and fluency since the discussion format allows learners to freely and naturally converse with their peers (Bonk & Ockey, 2003; Fulcher, 1996).

The students got the highest score on fluency on the narrative tasks. However, they got the lowest on complexity. The graphs above show that students' accuracy scores were between those for the argumentative and descriptive tasks. Therefore, we can assume that university students are less proficient in terms of fluency on narrative tasks than argumentative tasks.

For *accuracy* scores, *the descriptive task* was the highest followed by the argumentative task and narrative task. The reason for the highest score for accuracy on the descriptive task is probably that the learners' burden to use various tenses and complicated grammatical skills were reduced since they could only describe the

suggested pictures. The descriptive and argumentative tasks showed noticeable differences in accuracy scores. The accuracy score for the argumentative task was 3.20, and that for the descriptive task was 3.83. Therefore, students showed more accuracy on the descriptive task. B. E. Cho (2004) comments that Koreans students tend to be afraid of making mistakes and may respond to an utterance only in short phrases because they may not feel confident. Descriptive tasks seemed to allow students to use short sentences and to speak accurately. This finding is partially supported by Foster and Skehan (1996) and Skehan and Foster (1997), who showed that complexity and accuracy have a trade-off effect, with more difficult tasks resulting in improvement of complexity and easy tasks resulting in improvement of accuracy.

2. Topic Familiarity

Two previous studies (Foster & Skehan, 1996; Skehan & Foster, 1997) explored the extent to which *cognitive complexity*, operationalized as familiar and unfamiliar task content, would impact task performance. The results of their study show that well-known information (e.g., oneself and one's life) engendered a more fluent and accurate performance, whereas talking about new and unfamiliar information (e.g., certain crimes and punishments) prompted a performance that was less fluent and less accurate but more complex.

To check the effect of topic familiarity on speaking performance, the researcher divided students into two groups by topic familiarity: 'High,' those with a great deal of familiarity, and 'Low,' those who had little: Those who answered with a score of 4 in the questionnaire topic-item were disregarded. The researcher then categorized speaking tasks into three different types.

The results for topic familiarity on the argumentative task are presented below in Table 2.

TABLE 2

Effects of Topic Familiarity on Speaking Performance: Argumentative Task

Argumentative task		N	Mean (M)	SD	*t*	*p*	M difference
Complexity	Low	26	2.69	0.79	-0.90	0.372	-0.18
	High	32	2.88	0.75			
Accuracy	Low	26	2.92	0.74	-2.61	0.012***	-0.48
	High	32	3.41	0.67			
Fluency	Low	26	3.92	0.80	0.09	0.357	0.02
	High	32	3.91	0.69			

***p<0.01

The results show that the learners who had high topic familiarity received statistically higher scores for accuracy. That implies that when the students are familiar with the topic, they can produce relatively more accurate speech in argumentative tasks.

When we compared the participants' complexity, accuracy, and fluency scores on the descriptive task, we found a significant difference in accuracy scores, with 2.73 for the low topic-familiarity group and 3.33 for the high. Table 3 shows the results.

TABLE 3

Effects of Topic Familiarity on Speaking Performance: Descriptive Task

Descriptive task		N	Mean (M)	SD	*t*	*p*	M difference
Complexity	Low	15	2.53	0.83	-1.64	0.108	-0.37
	High	42	2.90	0.73			
Accuracy	Low	15	2.73	0.70	-2.89	0.006***	-0.60
	High	42	3.33	0.69			
Fluency	Low	15	3.67	0.82	-1.43	0.157	-0.31
	High	42	3.98	0.68			

***$p<0.01$

The reason for this result seems to be that the participants who had high topic familiarity tried to use more accurate sentences when they spoke on the descriptive topic. We assume that learners' topic familiarity influenced their accuracy in argumentative and descriptive tasks. We can assume that the reason for the significant difference on accuracy scores was because of the grammar focused language teaching in English teaching classes in Korea.

For the narrative topic, students' topic-familiarity level showed no statistical difference in complexity, accuracy, and fluency. Table 4 shows the results.

TABLE 4

Effects of Topic Familiarity on Speaking Performance: Narrative Task

Argumentative task		N	Mean (M)	SD	*t*	*p*	M difference
Complexity	Low	22	2.82	0.80	0.19	0.848	0.04
	High	36	2.78	0.76			
Accuracy	Low	22	3.14	0.83	-0.43	0.670	-0.09
	High	36	3.22	0.68			
Fluency	Low	22	3.91	0.92	-0.04	0.970	-0.01
	High	36	3.92	0.60			

It can be inferred that the narrative task does not require significant cognitive development in order to argue on the topic just like argumentative tasks since learners can use their own prior experiences or memories to complete narrative tasks.

Additionally, narrative tasks are not as easy as descriptive tasks since learners cannot lower their speaking burden using pictures. Therefore, there is no statistical difference related to topic familiarity for narrative tasks in speaking performance in terms of complexity, accuracy, and fluency.

3. Participants' Preferences on Task-type

According to the participants' journals, the easiest and most familiar task type was the descriptive task. When interviewed, they also mentioned that the descriptive task, 'describing five different pictures,' was easy because when they saw the pictures, they could choose the right words for each picture. They indicated that the descriptive tasks would be helpful for studying English. A slightly different attitude was found on the narrative task. Prior to completing the tasks, the learners thought the narrative task would be more helpful for learning English; however, after completing the tasks, they pointed out it was more difficult to speak about the narrative task compared to the descriptive task.

As for the argumentative topic, the learners mentioned it was not so difficult for them because they had thought about the controversial argumentative topics before. Therefore, they thought it might be useful to practice speaking tasks on argumentative topics familiar to them. Their positive attitude toward the argumentative task type may have resulted in their better speaking performance in terms of complexity, accuracy and fluency in comparison with the narrative task-type.

V. CONCLUSION

This study investigated the effects of three different types of tasks including argumentative, narrative and descriptive tasks on learners' speaking performance. Vercellotti's (2012) definitions of complexity, accuracy, and fluency were adapted to measure students' speaking performance. These three aspects have also been used by Foster and Skehan (1996) and Wendel (1997), among others. The effect of topic familiarity was also checked. In order to obtain relevant data, both quantitative and qualitative methods were used. For quantitative data, learners' voice recordings and a questionnaire were employed. Qualitative data included the learners' reflective journals and interviews. The research question was: Are there differences in speaking performance (in terms of complexity, accuracy, and fluency) among argumentative, narrative and descriptive tasks? A total of seventy students who took an English pronunciation course at a university participated in the study.

The results revealed that different task types affected learners' speaking performance. To be more specific, the argumentative task produced higher scores than the narrative task in terms of complexity, accuracy and fluency. The argumentative task also produced higher scores than the descriptive task in terms of complexity and fluency. The descriptive task produced the highest scores for accuracy compared with other task types. Therefore, when teachers want to check learners' speech complexity and fluency, it may be useful to choose argumentative tasks for classroom activities.

In terms of the learners' topic familiarity, the learner group with higher topic familiarity received higher scores for accuracy, especially for argumentative and descriptive task types. This indicates that this group of learners tried to avoid longer utterances and take advantage of the pictures in the descriptive task. These results suggest that when the classroom objective is to focus on accuracy, teachers should use descriptive or argumentative tasks.

The findings of this study have implications for EFL teaching. First, different task types may influence leaners' speaking performance, so teachers should be keen at choosing appropriate types of tasks according to their teaching purpose. Some speaking tasks in ELT textbooks are not relevant for university students. Therefore, when choosing speaking tasks, teachers should consider students' speaking proficiency and their interests. Second, argumentative tasks may be specifically useful for the teacher to check university-level learners' complexity and fluency in speaking. According to students' interviews, the most preferred speaking task type was the argumentative task. They mentioned that they could integrate their prior knowledge and English proficiency through argumentative tasks. When considering the limitations of a one-semester English class, teachers should choose familiar speaking tasks depending on students' interests. Third, accuracy in speaking may be influenced by individual learners' topic-familiarity, especially on argumentative and descriptive tasks. Thus, teachers should be cautious in grading or estimating individual learners' accuracy of speaking on these types of tasks.

Although we have attempted to obtain and analyze the relevant data in a systematic manner, the study has limitations. First, the number of of each task type used in the present study was not equal. In further studies, to measure the effects of different task types, the number of each type should be more carefully controlled. Second, using topics that match learners' interests can motivate them and provide better opportunities to learn the language. In other words, 'topic' is a medium for operating the language (White, 1988). Many researchers have also suggested that topics are the most relevant for speaking when they address learners' individual interests (Cunningsworth, 1984; Skeierso, 1991). Therefore, individual learners' topic interests should be included as another intervening variable in measuring task effect on

speaking performance. We hope that this present study will instigate more related studies in the future.

REFERENCES

Bonk, W. J., & Ockey, G. J. (2003). A many-facet Rasch analysis of the second language group oral discussion task. *Language Testing, 20*(1), 89-110.

Brown, H. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education.

Brown, R. (1991). Group work, task difference, and second language acquisition. *Applied Linguistics*, *12*(1), 1–12.

Brumfit, C. J. (1984). *Communicative methodology in language teaching.* Cambridge: Cambridge University Press.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 36–46). London: Heinemann.

Bygate, M. (2000). Introduction. *Language Teaching Research*, *4*(3), 185–192.

Cho, Byung Eun. (2004). Issues concerning Korean learners of English: English education in Korea and some common difficulties of Korean students. *The East Asian Learner*, *1*(2), 31–36.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, *11*(04), 367–383.

Cunningsworth, A. (1984). *Evaluating and selecting EFL teaching materials*. London: Heinemann.

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language and Cognition*, *28*(1), 7–21.

Doughty, C., & Pica, T. (1986). "Information gap" tasks: Do they facilitate second language acquisition? *TESOL Quarterly*, *20*(2), 305–325.

Duff, P. A. (1986). Another look at interlanguage talk: Taking task to task. In R. Day (Ed.), *Talking to learn* (pp. 147–181). Rowley, MA: Newbury House.

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford: Oxford University Press.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, *30*(4), 474–509.

Ellis, R., & Barkhuizen, G. P. (2005). *Analyzing learner language*. Oxford: Oxford

University Press.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(03), 299–323.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *21*(3), 354–375.

Fotos, S., & Ellis, R. (1991). Communicating about grammar: A task-based approach. *TESOL Quarterly*, *25*(4), 605–628.

Fulcher, G. (1996). Testing tasks: Issues in task design and the group oral. *Language Testing, 13*(1), 23-51.

Hunt, K. W. (1966). Recent measures in syntactic development. *Elementary English*, *43*(7), 732–739.

Lynch, T. (1991). Questioning roles in the classroom. *ELT Journal*, *45*(3), 201–210.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, *30*(4), 555–578.

Nunan, D. (1995). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Nunan, D. (2001). *Expressions.* Boston: Heinle & Heinle.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, *21*(01), 109–148.

Pattison, P. (1987). *Developing communication skills.* Cambridge: Cambridge University Press.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning.* Basingstoke: Palgrave Macmillan.

Skehan, P. (1992). Second language acquisition strategies and task-based learning. *Thames Valley University Working Papers in English Language Teaching*, *1*, 178–208.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1)*,* 38–62.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford: Oxford University Press.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogical tasks: Second language learning, teaching and testing* (pp. 167–185). New York: Pearson Education Limited.

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

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

Skeierso, A. (1991). Textbook selection and evaluation. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (2nd ed., pp. 432–453). Boston, MA: Heinle & Heinle.

Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 61*, 37-72.

Thornbury, S. (2005). *How to teach speaking.* Harlow, Essex: Pearson Longman.

Ur, P. (1982). *Discussions that work: Task-centred fluency practice.* Cambridge: Cambridge University Press.

Vercellotti, M. L. (2012). *Complexity, accuracy, and fluency as properties of language performance: The development of the multiple subsystems over time and in relation to each other.* Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

Wendel, J. (1997). *Planning and second language narrative production*. Unpublished doctoral dissertation, Temple University, Philadelphia.

White, R. (1988). *The ELT curriculum: Design, innovation and management*. Oxford: Blackwell.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, *14*(1), 85–106.

Willis, J. (1996). *A framework for task-based learning.* Harlow, Essex: Addison Wesley Longman.

Yule, G., & McDonald, D. (1990). Resolving referential conflicts in L2 interaction: the effect of proficiency and interactive role. *Language learning, 40*(4), 539-556.

예시언어(Examples in): English
적용가능 언어(Applicable Languages): English
적용가능 수준(Applicable Levels): Tertiary

Mi-Hye Kim
Hankuk University of Foreign Studies
Imun 1-dong, Dongdaemun-gu,

Seoul, South Korea
Tel:(02) 2173-2114
Email: kim32mary@hanmail.net

Received 15 September 2014
Revised 14 November 2014
Accepted 21 November 2014