

The Role of Task Type in Foreign Language Written Production: Focusing on Fluency, Complexity, and Accuracy

Mohsen Rezazadeh (Corresponding author)

Department of English Language, Faculty of Foreign Languages
University of Isfahan, Isfahan, Iran
E-mail: M1Rezazadeh@gmail.com

Mansoor Tavakoli

Department of English Language, Faculty of Foreign Languages
University of Isfahan, Isfahan, Iran

Abbas Eslami Rasekh

Department of English Language, Faculty of Foreign Languages
University of Isfahan, Isfahan, Iran

Received: November 15, 2010 Accepted: January 11, 2011 doi:10.5539/ies.v4n2p169

This study was completed at the University of Isfahan and supported by the Office of Graduate Studies. The authors are grateful to the Office for their support.

Abstract

The aim of this study was to investigate the effects of two task types on foreign language written production. Particularly it addressed the issue of how three aspects of language production (i.e. fluency, complexity, and accuracy) vary among two different task types (i.e. argumentative writing task and instruction writing task). One hundred sixty eight intermediate learners of English were randomly selected and divided into two task type groups. Separate ANOVAs were conducted on each dependant variable to see whether there were any statistically significant differences across the groups. The findings revealed that participants in the instruction-task group performed significantly better than those in argumentative-task group in terms of fluency and accuracy. In addition, analysis of complexity measures showed that argumentative essays were produced with more complex language than the instruction essays. Possible explanations are provided and the pedagogical implications of these findings are discussed.

Keywords: Task type, Foreign language written production, Fluency, Complexity, Accuracy

1. Introduction

There was a growing interest in the role of tasks in second language teaching and learning in recent years. To date, most task-based studies have focused on oral language production. Previous research indicates that task type is an important variable both in how learners approach language production and in how it affects the fluency, accuracy, and complexity of language products. Foster and Skehan have conducted several studies on task type within the cognitive research framework (Foster and Skehan 1996, Skehan and Foster 1997, Skehan 1998, Skehan and Foster 1999, Foster and Skehan 1999, Skehan 2003). They believe that learners' attentional resources are limited and there should be a competition between the three aspects of fluency, accuracy and complexity in order to get the learners attention.

Skehan and Foster (1999) investigated the effects of structured and unstructured narrative tasks on L2 learners' speech performance. Forty-seven intermediate EFL students at an English university participated in this study. Each participant told either a structured narrative or an unstructured narrative under four experimental. The aim of the four conditions was to place different levels of cognitive demands on the participants while they told their narratives. Skehan and Foster (1999) found that the task that had more structured characteristics (i.e., lower cognitive demand) led to significantly more fluent speech than did the unstructured task (i.e., higher cognitive demand). However, the task that had more structured characteristics did not lead to speech that was significantly more complex or more

accurate than did the unstructured task. In addition, they found an interaction effect between task structure and processing condition. A combination of a structured task and a less demanding processing condition promoted speech accuracy. The results of Skehan and Foster (1999) are particularly relevant for the current study. They found that tasks with a greater degree of inherent structure resulted in increases in fluency.

Task types may be one of the crucial factors in determining if writers are able to automatize certain features of writing tasks or deal with additional cognitive load to process those aspects (Foster & Skehan, 1996; Franken & Haslett, 2002; Sweller, 1994). For example, argumentative writing requires writers to generate complex information, which makes it more demanding than descriptive writing. In contrast, descriptive writing is characterized as a task that has a clear inherent structure (Foster & Skehan, 1996), which requires writers to describe individual actions or characters. Findings by previous researchers have revealed the differential effects of task type on L2 performance.

Foster and Skehan's framework has been very influential in task-based research within the cognitive framework. But there are also a number of other empirical studies that have shown that there are important differences in the language that learners produce while performing different tasks. These studies tend to look at particular features of learner language that vary by task. Two relevant studies to this research have found differences relating to task type in syntactic complexity (Duff 1986) and nominal and verbal constructions (Bygate 1999).

Duff (1986) investigated task type among eight pairs of non-native speakers. She focused on both quantity and quality of interaction in two tasks. One task required the participants to solve a problem together; the other assigned them different viewpoints on an issue which they had to debate. The quantity of their language production was measured in c-units, defined as "a word, phrase, or sentence that in some way contributed pragmatic or semantic meaning to a conversation". The quality was measured by the number of turns, types of questions, syntactic complexity. Results revealed that the problem-solving task generated more turns (per individual and for the whole task), and more c-units per task. The debate task resulted in more words per turn, more words per c-unit, and more syntactic complexity.

In another study, Bygate (1999) investigated the effects of two different tasks on grammatical complexity. Two argumentation tasks and two narrative tasks were used. Participants were asked to prioritize a list of options in the argumentative task and in the narrative task they had to tell a story based on a series of pictures. Complexity was measured in terms of length of T-unit, subordination, number of verb arguments, and type of subordination. Findings suggested that the narrative task had more words per T-unit, but there was no significant difference considering the use of subordination. Results showed that the argumentation task presented a language that contained more 'verbal', while the narrative task had more 'nominal' output. Bygate's results offer further proof of the often-unexpected differences in learner speech on different tasks.

It is evident from the above studies that the type of task presented to learners can lead to great variability in the results. The research to date indicates that task type is a rich area for further research. Consequently, more attention needs to be paid to the relationships between task types on the performance of a written text. It was therefore decided to undertake further exploration and to collect additional evidence about the role of task type in fluency, complexity, and accuracy of EFL learners' written products.

2. Purpose and the Research Questions

This study is an attempt to examine the role that variation in task type may play in the characteristics of EFL learners' written performance. The variables that are examined in conjunction with task type are fluency, complexity, and accuracy. Two task types are argumentative writing task and instruction writing task. That is, this study is to investigate the variation that may exist in the fluency, complexity, and accuracy of argumentative and instruction essays.

Based on the problem and purpose discussed above, the following research questions were addressed:

1. Does task type have any effects on the *fluency* of EFL learners' written products?
2. Does task type have any effects on the *complexity* of EFL learners' written products?
3. Does task type have any effects on the *accuracy* of EFL learners' written products?

3. Method

3.1 Participants

One hundred sixty eight fulltime undergraduate English majors enrolled in EFL writing courses at two universities in Isfahan, Iran took part in the study voluntarily. Following an Oxford Placement Test, intermediate-level learners were selected for the study in each class. The test contained 60 multiple choice items, and it was used to enable the researchers to select a homogeneous group. This test consisted of grammar (20 items), vocabulary (20 items),

reading comprehension (20 items) and Writing. The allotted time for answering the questions was 45 minutes. After correcting the papers, 168 students were selected as the intermediate group based on the OPT manual. However, measuring proficiency was not important in this study because it intended to portray EFL classrooms where proficiency levels were not often controlled (Park, 2010). Afterwards, participants were assigned to two groups, argumentative writing task (N = 94) and instruction writing task (N = 74). Data-gathering began 5 weeks after the start of term to allow the subject population to stabilize. All data were collected during normally scheduled class times by course instructors.

3.2 Essay Writing Tasks

The two writing tasks selected for this study were instruction task and an argumentative task. The instruction task, adapted from Mehnert's (L998) study, required the participants to leave a text message for their friends who planned to visit them in Isfahan, telling them how to get from the train station to the University. The reason provided for not meeting the friend was that the participant had an important test to take at the same time that their English-speaking friend was to arrive. As a result, their friend had to proceed to their university by herself/himself. The argumentative task, on the other hand, required that participants give their opinions concerning the pernicious influence of examinations on education. The instruction task was considered to have lower cognitive and linguistic demands than the argumentative task according to Skehan's (1996) criteria for task grading: cognitive complexity, code complexity.

Skehan's (1996) first criterion, cognitive complexity, covers two areas: cognitive familiarity and cognitive processing. Regarding the first area, cognitive familiarity, the instruction task would be more familiar to the participants of the present study than the argumentative task for two reasons. First, the participants clearly knew the directions from the train station to their university. Second, the participants were more familiar with the structure of instruction tasks than they were with the structure of argumentative tasks. With respect to the second area, cognitive processing (i.e., "the amount of on-line computation that is required while doing the task" (Skehan, 1996, p.52)), the instruction task required a smaller amount of on-line computation than the argumentative task for the following three reasons. First, the instruction task did not require as many reasoning operations as the argumentative task. For example, the instruction task only required participants to give directions whereas the argumentative task required them to come up with reasons to support their positions. Second, regarding the nature of input material used in the task, when undertaking the instruction task participants simply drew on ready-made knowledge since they already knew how to go from the train station to their university. However, when performing the argumentative task the participants could not draw on existing knowledge but had to transform the knowledge in order to come up with main and supportive reasons to back-up their positions. Finally, the degree of organization of input material of the instruction task was higher than that of the argumentative task. The sequence of the content of the instruction task has higher predictability and interconnectedness while that of the argumentative task tends to be arbitrary. That is, participants had to weigh the significance of their main and supportive reasons in order to come up with the organization of the content.

Based on the above-discussed criteria, it is reasonable to claim that the instruction task was lower in its cognitive and linguistic demands than the argumentative task.

3.3 Measures

Essays were coded for a range of dependent variables. Measures of fluency, complexity, and accuracy were employed to evaluate the quality of the participants' written production. These measures were largely the same as those used in other studies (i.e., in Ellis & Yuan, 2004; Foster & Skehan, 1996; Storch, 2005; Wigglesworth & Storch, 2009). In order to undertake this analysis, all written work was coded in the first instance for T-units and clauses. A T-unit is defined by Hunt (1966) as "one main clause plus whatever subordinate clauses happen to be attached to or embedded within it" (p. 735). In order to measure complexity and accuracy, the compositions had to be analyzed for clauses, distinguishing between independent and dependent clauses. There is some disagreement among researchers as to how to code for clauses, particularly dependent clauses. In this study, a dependent clause was one which contained a finite or a non-finite verb and at least one additional clause element of the following: subject, object, complement or adverbial.

Following Wigglesworth and Storch(2009), fluency was measured in terms of the average number of words, T-units and clauses per text. Another measure was complexity which reflects the writer's willingness to engage and experiment with a range of syntactic structures, moving beyond coordination to more complex structures which include subordination and embedding (Storch & Wigglesworth, 2007). Complexity was measured through proportion of clauses to T-units, which according to Foster and Skehan (1996) is a reliable measure, correlating well with other measures of complexity. Another measure of complexity is the proportion of dependent clauses to clauses

(DC/C), which examines the degree of embedding in the text (Wolfe-Quintero, Inagaki, & Kim, 1998).

To evaluate accuracy, two measures were used: the proportion of error-free T-units to all T-units (EFT/T) and the proportion of error-free clauses of all clauses (EFC/C). Both proportions were expressed as percentages (Wigglesworth & Storch, 2009). Errors in this study included syntactical errors (e.g., errors in word order, missing elements) and morphology (e.g., verb tense, subject-verb agreement, errors in use of articles and prepositions, errors in word forms). Errors in lexis (word choice) were included only when the word used obscured meaning. All errors in spelling and punctuation were ignored. In summary, the quantitative measures shown in Table 1 were used to analyze the writings produced by the participants.

3.4 Data Analysis

A series of one-way ANOVAs were performed on all the measures and the alpha for achieving statistical significance was set at .05. Outliers were detected by means of box plots and eliminated from the calculation in order to achieve the sphericity of the data, which was confirmed by means of Mauchly's test. Inter-rater reliability coefficients were obtained on all categories identified for analysis by two raters working independently. The analysis of the written texts was carried out by the researcher and a research assistant. Inter-rater reliability was above 88% on all measures.

It was felt important to examine the dependent variables in separate ANOVAs, rather than through a more general multivariate ANOVA. The rationale for the dependent variables presented earlier indicates the distinct role that each contributes. This claim is supported by a factor analytic study of a pooled data set from the present data set and previous researches (e.g., Foster & Skehan, 1996; Skehan & Foster, 1997). This analysis generated a three-factor solution, with the three orthogonal factors clearly identifiable as fluency, complexity, and accuracy, suggesting adequate independence among them. Moreover, examining the effects of the treatments by one-way ANOVAs, minimizes the risk of a Type 1 error (Yuan & Ellis, 2003). Further, Keselman et al. (1998) argue that there is very limited empirical support for a MANOVA- univariate data analysis strategy.

4. Results

As indicated earlier, three aspects of language use were examined to see how the participants performed the two writing tasks. In this section, we look at the data to establish what differences, if any, appear in learners' written production when they were asked to perform different tasks. The results of one-way ANOVAs will be reported separately comparing the results on the measures for fluency, complexity and accuracy for instruction writing task and argumentative writing task.

4.1 Fluency

Table 2 shows that instruction essays had more T-units and clauses than the argumentative essays. That is learners made a larger number of T-units and clauses in instruction task compared to the argumentative task. In terms of the number of words vs. the total number of words per text, the results show both task types were roughly the same, although instruction essays were slightly longer.

The results of the one-way ANOVA shown in Table 2 revealed that the difference in the groups was statistically significant in two measures (T-units per text $F = 51.104$, $df 1,168$, $p = 0.000$; clauses per texts $F = 20.310$, $df 1,168$, $p = 0.000$). Regarding the first measure of fluency (words per text) the difference between the two groups was not significant (words per texts $F = 0.129$, $df 1,168$, $p = 0.720$). The results show an overall difference that is statistically significant in the case of clauses and T-units per text but not in the case of words per text although the P value (0.720) is very close to statistical significance.

Based on the above analysis, fluency was significantly higher in instruction essays indicating that in terms of fluency (i.e. length of production), the low cognitive and linguistically demanding task was significantly more effective than the high cognitive and linguistically demanding task.

4.2 Complexity

Two variables were assessed to measure the complexity of language use in the participants' written tasks. Table 3 shows that the argumentative-task group wrote more complex essays producing more clauses per t-unit.

The results of one-way ANOVAs indicate that there is a statistically significant difference for production with p values of 0.001 and 0.000. As it is shown in table 3, analysis of the two measure of complexity shows that argumentative essays were produced with more complex language than the instruction essays (ratio of clauses to T-units: $F = 10.904$, $df 1,168$, $p = 0.001$; percentage of dependent clauses: $F = 20.523$, $df 1,168$, $p = 0.000$). However, as stated in Wigglesworth and Storch (2009), the two measures used here reflect the same construct and there is the possibility that other measures of complexity such as the Mean Segmental Type/Token Ratio (Malvern &

Richards, 2002) or other measures of grammatical verb form such as modality, tense or voice (Ellis & Yuan, 2004) might elicit different results.

4.3 Accuracy

With regards to accurate language use in written tasks, four measures were used: error free T-units, error free T-units percentage, error free clauses, and error free clauses percentage. As Table 4 shows, the mean scores of the instruction-task group for all measures of accuracy are higher than the means for argumentative-task group. However, in the case of error free clauses percentage, the means for the two groups are close to each other.

The results of the one-way AVOVAs for complexity variables show that the two task type groups differed with regard to error free T-units ($P = 0.000$), error free T-units percentage ($P = 0.025$), and error free clauses ($P = 0.000$) but not with regard to error free clauses percentage, though it came close to reaching statistical significance ($P = 0.653$) (see table 4).

Based on the above analysis, the subjects in the instruction-task group performed significantly better than those in argumentative-task group in terms of accuracy. Overall, these results for accuracy indicate that argumentative essays were more accurate than instruction essays.

5. Discussion

The purpose of this study was to investigate how certain aspects of learners' written performance are affected in two different task types and to use the results of the analyses to improve learners' L2 writing process in the classroom context. In what follows, the findings obtained regarding each research question will be discussed in turn.

The first research question addressed the effects of task type on the fluency of learners' production in written tasks. Learners showed increases in fluency when working on instruction task which is lower in its cognitive and linguistic demands than the argumentative task.

This finding may be explained on the basis of writing production processing. Participants may have been able to run most of the formulator of their production processes automatically and in parallel with the conceptualizer when they performed their writing for the low cognitive and linguistically demanding task. This seems reasonable, given that this type of task required simple language (i.e. vocabulary and grammatical structures). In contrast, the participants may have had to run most of the formulator consciously and in sequence with the conceptualizer (rather than parallel to) when performing the high cognitive and linguistically demanding task (i.e. argumentation). Possibly, this is because the argumentative task required more complex language. Since the instruction task enabled most of the formulator to run automatically and to run in parallel with the conceptualizer, it led to higher fluency than did the argumentative task.

This finding is consistent with the results of previous studies conducted by Foster and Skehan (1996, 1997). However, it does not support the results of the study concluded by Mehnert (1998). Mehnert found that both tasks lead to language of similar fluency levels. Nevertheless, Mehnert's (1998) non-significant finding may have been caused by the insufficient difference between the cognitive and linguistic demands of the two tasks as assigned in her study.

The second research question concerned the effects of task type on the complexity of learners' written output. Results show that argumentative task is significantly more effective than the instruction task in promoting complexity. This Finding may be accounted for by a task-demand-based explanation, namely, that the demands of the argumentative task for elaborated content and complex language influence participants both to produce more elaborated content and to use more advanced vocabulary and grammatical structures. The finding of the present study is consistent with the results reported by most previous research (Foster & Skehan, 1996; Mehnert, 1998; Skehan and Foster, 1997, 1999).

The third research question addressed the effects of task type on the accuracy of learners' written production. The subjects in the instruction-task group performed significantly better than those in argumentative-task group in terms of accuracy. This may be explained on the basis of the participants' L2 grammatical knowledge. The instruction task requires simpler grammatical structures, whereas the argumentative task requires more complex grammatical structures. Since the participants may have had higher control over the simpler grammatical structures required for the instruction task, than for the more complex structures required for the argumentative task, they produced language with greater accuracy for the instruction task.

This result is consistent with findings of most previous research. Foster and Skehan (1996,1997), and Mehnert (1998) reported that low cognitive and linguistically demanding tasks are more effective than high cognitive and linguistically demanding tasks in promoting accuracy. The explanation provided in this previous research for the

effects of assigned low demanding tasks in promoting language accuracy is that this type of task enables participants to allocate more attentional resources toward achieving speech accuracy than do the high demanding tasks. Foster and Skehan (1997) put forward that participants performing a low demanding task were not influenced to use complex language. However, the explanation provided by Foster and Skehan (1997) may be inadequate. Based on the results of the analysis in the present study, one could argue that participants may direct more attentional resources toward monitoring the grammatical accuracy when performing high demanding tasks. Despite the fact that the participants directed smaller amounts of their attentional resources toward monitoring the grammatical accuracy for the instruction task than toward monitoring the grammatical accuracy for the argumentative task, the participants still produced language for the instruction task with higher accuracy levels than for the argumentative task. This result suggests that the participants' knowledge of grammatical structures may be a stronger factor in determining speech accuracy than the amount of attentional resources the participants direct toward monitoring the grammatical accuracy of their written products.

6. Conclusion

The purpose of the present study was to provide insight into the effects of task type i.e. cognitive and linguistic demands of tasks (instruction and argumentation) on the fluency, complexity, and accuracy of learners' written performance. Findings indicate that participants in the instruction-task group performed significantly better than those in argumentative-task group in terms of fluency and accuracy. In contrast, analysis of complexity measures revealed that argumentative essays were produced with more complex language than instruction essays.

The findings of the present study yield the following pedagogical implications:

With regard to fluency, for low demanding tasks like instruction task in this study, teachers should instruct learners to plan the language of their essays. Teachers should also provide learners with the vocabulary they may need for their writings. For high demanding tasks like argumentation in this study, teachers should instruct learners to plan the language of their essays. In addition, teachers should provide learners with both; the vocabulary the learners may need and information about grammatical structures that are relevant to the assigned tasks.

To promote complexity for low demanding tasks, teachers should instruct L2/FL learners to how plan only the content of their essays. To promote complexity for high demanding tasks, teachers should instruct L2/FL learners how to plan both the content and language.

Regarding accuracy, for both low and high demanding tasks, teachers should remind learners to direct some of their attention toward monitoring the grammatical accuracy. Additionally, for high demanding tasks, teachers should provide learners with information about the grammatical structures that are relevant to the assigned tasks to help learners apply the grammatical structures accurately.

The following recommendations for future research are based upon the results in this study. First, the present study needs to go beyond the cognitive domains over more social/cultural domains in order to make it more comprehensive. Writing cannot be conducted in a social vacuum and social/cultural contexts where writing takes place cannot be ignored as well. Thus, incorporating learners' cognitive process of writing with their social/cultural underpinnings of written academic genres will enrich the understanding of the L2 writing process. Second, more research needs to be conducted that assigns tasks comparable to those used in this and previous studies in order to further verify the interaction effects of different levels of cognitive and Linguistic demands of tasks. Third, based on literature, this study recruited students with intermediate English proficiency. Therefore, the findings of the present study cannot be generalized to apply to L2/FL learners with low or high proficiency levels. In order to examine this issue, a future study could include participants with different proficiency levels.

References

- Bygate, M. (1999). Task as context for the framing, reframing and unframing of language. *System*, 27(1), 33-48
- Duff, P. (1986). Another look at interlanguage talk: Taking task to task. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition*, (pp. 147-181). Rowley, MA: Newbury House.
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26(01), 59-84
- Foster, P., & Skehan, P. (1996). The Influence of Planning and Task Type on Second Language Performance. *Studies in Second Language Acquisition*, 18(3), 299-323
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3(3), 215
- Franken, M., & Haslett, S. (2002). When and why talking can make writing harder. In S. Ransdell & M. L. Barbier

- (Eds.), *New directions for research in L2 writing* (pp. 208-229). Dordrecht, The Netherlands: Kluwer Academic.
- Hunt, K. (1966). Recent measures in syntactic development. *Elementary English*, 43, 732-739
- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., et al. (1998). Statistical Practices of Educational Researchers: An Analysis of Their ANOVA, MANOVA, and ANCOVA Analyses. *Review of Educational Research*, 68(3), 350-386
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85-104
- Park, S. (2010). The influence of pretask instructions and pretask planning on focus on form during Korean EFL task-based interaction. *Language Teaching Research*, 14(1), 9-26
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38-62
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford University Press, USA.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(01), 1-14.
- Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task-based learning. *Language Teaching Research*, 1(3), 16-33
- Skehan, P., & Foster, P. (1999). The Influence of Task Structure and Processing Conditions on Narrative Retellings. *Language Learning*, 49(1), 93-120
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(3), 153-173
- Storch, N., & Wigglesworth, G. (2007). Writing Tasks: The Effects of Collaboration. In M. d. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157-177). London: Multilingual Matters.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295-312
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445-466
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. Honolulu, Hawaii: University of Hawaii Press.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1-27

Table 1. Measures used in analysis of written productions

Fluency	Complexity	Accuracy
average number of words per text	proportion of clauses to T-units	percentage of error-free
average number of T-units per	percentage of dependent clauses to all	percentage of error-free
average number of clauses per		

Table 2. Fluency Measures ANOVA

Fluency	Task Type	N	Mean	Std. Deviation	F	Sig.
Average words per text	Instruction	74	162.97	39.354	.129	.720
	Argumentation	94	161.00	31.954		
	Total	168	161.87	35.307		
Average T-units per text	Instruction	74	16.05	4.141	51.104	.000
	Argumentation	94	12.43	2.362		
	Total	168	14.02	3.724		
Average clauses per text	Instruction	74	27.68	6.728	20.310	.000
	Argumentation	94	23.68	4.748		
	Total	168	25.44	6.025		

Table 3. Complexity Measures ANOVA

Complexity	Task Type	N	Mean	Std. Deviation	F	Sig.
Clauses per T-unit	Instruction	74	1.7608	.35601	10.904	.001
	Argumentation	94	1.9237	.28349		
	Total	168	1.8519	.32670		
Dependant Clauses Percentage	Instruction	74	34.5161	11.01015	20.523	.000
	Argumentation	94	41.1907	8.07915		
	Total	168	38.2507	10.01921		

Table 4. Accuracy Measures ANOVA

Accuracy	Task Type	N	Mean	Std. Deviation	F	Sig.
Error free T-units	Instruction	74	9.84	3.698	39.238	.000
	Argumentation	94	6.74	2.700		
	Total	168	8.11	3.522		
Error free T-units percentage	Instruction	74	60.8847	16.12866	5.134	.025
	Argumentation	94	54.4879	19.61602		
	Total	168	57.3055	18.38856		
Error free clauses	Instruction	74	20.68	5.871	16.107	.000
	Argumentation	94	17.40	4.696		
	Total	168	18.85	5.477		
Error free clauses percentage	Instruction	74	74.3572	8.66368	.203	.653
	Argumentation	94	73.5074	14.30174		
	Total	168	73.8817	12.12002		