Title of Project:
Mapping the Relationships among the Cognitive Complexity of Independent Writing Tasks, L2 Writing Quality, and the Complexity, Accuracy, and Fluency of L2 Writing

Researcher:
Weiwei Yang  
Georgia State University  
yang-weiwei@hotmail.com

Research Supervisor:
Dr. Sara Cushing Weigle

Project Summary:

Task equivalence is one of the main concerns with respect to validity in second language writing assessment. Task equivalence can manifest itself in equivalence in 1) writing scores, 2) language performance in the writing produced, and 3) how language production features predict writing scores. Determining the cognitive complexity of tasks is one way to classify tasks and to study task equivalence. Cognitive complexity has been defined as "the level of thinking skills or intellectual functioning required to accomplish certain tasks" (Hale et al., 1996, p. 12) or “the extent to which task characteristics can affect the allocation of an individual’s attention, memory, reasoning and other processing resources” (Robinson, 2007a, p. 17). The construct of cognitive complexity has been studied in both the task-based language literature and the first (L1) and second (L2) writing literature. However, researchers in the two areas rarely cite each other’s work, and the cognitive complexity of tasks has not been given adequate attention in L2 writing assessment, particularly as it relates to task equivalence.

The current study is designed to bridge the task-based language literature and the L1 and L2 writing literature. It examines the effect of the cognitive complexity of tasks on L2 writing scores, the effects of the cognitive complexity of tasks on the complexity (lexical and syntactic complexity), accuracy and fluency (CAF) of L2 production, and the predictive power of CAF features on writing scores for tasks of different cognitive complexity. Two cognitive complexity dimensions were investigated in this study: rhetorical task, which varies in reasoning demand, and topic familiarity, which varies in the amount of direct knowledge of topics. Four levels of rhetorical tasks were studied: narrative, expository, expositional-argumentative, and argumentative tasks. Three levels of topic familiarity were examined: personal-familiar, impersonal-familiar, and impersonal-less familiar tasks. Six writing prompts were used to study these cognitive complexity dimensions and levels. The subject matter for all the writing tasks was controlled; when one dimension was studied, the other dimension was also controlled.
A total of 375 EFL undergraduate students at a university in Southeast China participated in the study, with each student writing on one of the six tasks and with approximately a total of 60 students writing on each task. The writing task was timed and completed within 30 minutes in each case. The essays were rated by five experienced raters who all had ESL teaching experience using the TOEFL iBT Test Independent Writing Rubrics, with half-point ratings added. The essays were also rated on task fulfillment by an experienced ESL teacher, writer, and the researcher. Thirteen CAF measures were carefully selected and used, and the measures were all automated through appropriate computer tools. Analyses of variance were conducted to examine the effects of the cognitive complexity dimensions on L2 writing scores and on CAF features. All-possible subsets regression analyses were conducted to investigate the predictive power of the CAF features on L2 writing scores for each task. Comparisons of the best regression models were then made among the tasks for each of the cognitive complexity dimensions.

The one-way analysis of variance (ANOVA) tests revealed that neither rhetorical task nor topic familiarity had an effect on the L2 writing scores of the participants. The one-way multivariate analysis of variance (MANOVA) tests and follow-up univariate analyses showed that neither rhetorical task nor topic familiarity had an effect on the accuracy and the fluency of L2 writing, but that the argumentative essays were significantly more complex in global syntactic complexity features than the essays on the other rhetorical tasks. In addition, the essays on the less familiar topic were significantly less complex in lexical features than the essays on the more familiar topics. All-possible subsets regression analyses revealed that the CAF features explained approximately half of the variance in the writing scores across the tasks and that writing fluency, as measured by essay length, was the most important CAF predictor for five of the six tasks. It is important to note that lexical sophistication was however the most important CAF predictor for the argumentative task.

The regression analyses further showed that the “best” regression model for the narrative task was distinct from the ones for the expository and argumentative types of tasks, with the former consisting of the fluency, lexical diversity, and global syntactic complexity measures while the latter primarily consisting of the fluency, lexical sophistication, and accuracy measures. The “best” regression model for the personal-familiar task was distinct from the ones for the impersonal tasks, with the former including the fluency and the accuracy measures and the latter including the fluency, accuracy, lexical sophistication, and overall clause-level complexity measures. The results of the multiple regression analyses for the topic familiarity dimension need to be interpreted with the understanding that the findings might be different depending on whether the writers fulfilled the tasks as instructed. Specifically, a number of the writers in this study did not produce personal essays for the personal task while others approached the less familiar topic by making it a more familiar one. The correlation results for the on-task samples only suggest that the predictive power of the CAF features on scores for on-task samples is likely to be different from that for off-task samples.

The study findings have implications for L2 writing assessment, in terms of task selection, rating rubric development, rater training, and automated essay scoring. First, in regards to task selection, the scores did not differ across the tasks; however, the study suggests that
argumentative tasks can be viable writing tasks in an assessment of general L2 writing ability, as they are likely to elicit syntactically more complex language and have higher demands on the use of sophisticated and more advanced vocabulary. The study also suggests that impersonal-less familiar tasks may not be good choices for L2 writing assessments because they are likely to elicit lexically less complex language. Furthermore, a good number of writers seem to have difficulty in fulfilling the requirements for both personal-familiar and impersonal-less familiar tasks, making these tasks not the best choices for writing assessment.

The study also has implications for rating rubric development, rater training, and automated essay scoring. The findings from the regression analyses suggest that the rating rubrics for different rhetorical tasks should reflect the differential importance of the CAF predictors for the different task types. Similarly, such rating criteria should be made explicit to novice raters during rater training and be considered in the logarithms for automated essay scoring for the different task types. The study findings related to the different sub-constructs of lexical complexity and syntactic complexity also point to the ambiguity and the inadequacy of the rating criteria of “range of vocabulary” and “complex constructions” in rating rubrics (e.g., Jacobs, et al., 1981) and the necessity to specify such criteria in relation to task types.
References


Genung, F. (1900). *The working principles of rhetoric: Examined in their literary relations and illustrated with examples*. Boston, MA: Ginn.


Yang, W. (2009). Topic effect on writing fluency and linguistic complexity of ESL writers and predictive values of writing fluency and linguistic complexity of ESL writers on writing scores (Unpublished course paper). Georgia State University, Atlanta, GA.


