



Cross-cultural Differences in Writing Styles of L2 Students of English

Roxanne Wong

City University of Hong Kong (Hong Kong)

University of Jyväskylä (Finland)

elroxann@cityu.edu.hk

Abstract

This paper will give details of a cross-cultural study conducted between students in the United States and Hong Kong. The author looks at the differences between language use, sentence structure and writing style between the two groups of students involved in the study. This study has more global implications in that the students come from a variety of ethnic backgrounds and are not a homogeneous group. As part of my PhD studies and in collaboration with Edsphere, a study was undertaken in Hong Kong and Mississippi, to develop a Diagnostic Writing system that automatically score essays. A pilot was conducted with 1400 students in Hong Kong. These students were asked to write a 500-word essay. A further 540 students from the USA wrote essays using the same writing prompts. The two sets of essays from both Hong Kong and the United States were subsequently double scored using the same rubric. The essays were analyzed for the differences in the writing styles, vocabulary and grammar. The author will show some of the major differences between the two groups of L2 learners. The implications of the findings to a more global language testing community will also be presented.

1. Introduction

In 2011 a research project was begun in Hong Kong to design an automated essay scoring system for a diagnostic test that has been in use since 2008. This study included the redesign of a scoring rubric, item writing and piloting of both the items and rubric with human raters. The study included provisions to trial online scoring systems to determine the effectiveness of these with the in house developed materials. In 2014 the project came to an end.

The current study was undertaken subsequent to the original project. The Metametrics EdSphere system was used as the capture system and the researcher has since gone on to work with Metametrics to continue to develop an online system. As Metametrics is a United States based company and has been extensively used at schools throughout the United States, it was decided to pilot the same test items on secondary school students there.

The main focus of the study was to determine if there was a significant difference in the vocabulary, semantic, and syntactical abilities of the students in the United States and Hong Kong. In addition, the scoring instrument was being tested for ease of use by human raters and for comprehensiveness of explanation of the items in each of the scoring domains.

It is important that the scoring instrument be useful for the task as well as for the stakeholders. (Knoch, 2009) reported that when a rating scale with descriptors based on discourse-analytic measures is used the results are more valid and useful for giving students feedback on diagnostic assessment. Tierney and Simon (2004) report that most rubrics contain flaws that affect not only the usefulness, but also the validity of the results. This is particularly important when developing a scoring instrument that will be used by both human and machine raters.

The descriptors being used must be clear to all users of the instrument. They must be reliable for both student users and raters (Arter & McTighe, 2001; Moskal, 2003; Popham, 1999; Wiggins, 1998).

When designing a scoring instrument that is to be used for automated scoring, it is important that the instrument is able to differentiate the levels given to the students (Moskal, 2003; Wiggins, 1998). Finally, a balance must be struck between accessible wording for usability and detailed descriptors and consistency in performance description must be maintained across achievement levels (Popham, 1997; Simon & Forgette-Giroux, 2001; Wiggins, 1998).

2. Population

The study consisted of 1491 participants from 5 universities in Hong Kong and 539 participants from two school districts in the United States. The Hong Kong students range in age from 17 to 19. All were incoming freshmen that had not previously attended university courses. The students from the United States were aged 14 to 18. There were a total of 4 expert raters involved. Two were university instructors and two were primary school teachers who have had previous experience teaching



secondary school students in Hong Kong and abroad. The variety of age/grade levels of the students and different expert raters was used to get a broad spectrum of language abilities of the students that will be the subsequent users of the program.

Methodology

The students wrote a 450 word argumentative essay based on one of two reading passages. They entered the data online in the EdSphere Lexile writing system. 2501 papers from Hong Kong and 1079 papers from the United States were double marked.

The study results were tabulated for students between the United States and Hong Kong. The papers were divided into three basic levels, elementary, intermediate, and advanced. The levels were devised as they already hold face validity for the students and require limited explanation. Elementary is students with a Lexile writing score below 900, Intermediate has a Lexile score between 901 and 1300, and advanced has a Lexile score of above 1301.

30 papers (five from each band level and each country) of each topic were then randomly selected for detailed analysis of vocabulary use, to semantic complexity (the level of the words written) and syntactic sophistication (how the words are combined) (Schwartz, 2008). Each group of students' papers was first analyzed for density of errors in the writing. They were then analyzed for the types of grammatical errors made, and finally for the language that was used within the texts.

The final part of the study was to interview the users of the scoring instrument to determine the usefulness and the ease of use. This was done on an individual basis with open ended questions.

3. Discussion

The first area of investigation was general grammatical accuracy. Students were scored on a range of errors including but not limited to: subject verb agreement, pronoun referencing, and tense.

When looking at the writing of the elementary and intermediate level students, the first most prominent error made by students in both groups was with subject verb agreement. Both groups of students had a tendency to use a plural noun with a singular form of the verb. This was especially noticeable in the writings from the Hong Kong students. A sampling of the errors can be seen in table 1.1.

Corinth (USA)	Hong Kong
Lecturing are not for everyone.	The traditional lecture are totally...
...lecturing show if your really learning...	There are no more lecture to see...
When the teacher lecture and some likes...	They has a little ways to find...
Lectures has not been replaced...	The two source have their...
MOCOs is going to fail...	The lecture are only one...

Table 1.1

The second common error type found among this group of students was one of preposition usage. Students do not understand prepositional collocations. Some excerpts can be seen in table 1.2.

Corinth (USA)	Hong Kong
...hands at activities...	...efficient ways of delivery knowledges.
...to get in the Internet.	Past of the years...
...self gradifacation to all of my dediction.	...methods on transferring knowledge.
It's put on the mind with you...	...share things to students.
So in the moment it is way...	...technology on the lecture.

Table 1.2



With the higher-level students, these types of errors were much less frequent or non-existent. Some minor errors were found: “classes on your on time”, “there is usually twenty to twenty-five students in a class”, and “The later regards to an interest-based learning” are a few examples. The more common errors found at this level were with plural forms of nouns: “no professor are needed anymore”, “interactive education methods is”, and “Student that thrive on”.

Syntax was another area that was analyzed and the most common error found in the lower two levels of students was that of run-on sentences. Every one of the papers evaluated had at least one example,

“In my opion i belive that the only reason nobody really learned of it because , you cant just keep doind the same lecturing process everytime i think you can mix it up sometimes such as go on trips so they can see things and you can make fun doing work , i dont think you should stay in a classs all day and listen to your teacher talk all day about one subject.”

The other predominant error found was in the use of comma splices. The students frequently combined sentences with commas, or they would insert commas which were not needed,

“It lies it's superiority ,it alive and, real, It's put on the mind with you.”

“But I think this also has a good effect that many students in the future, won't just sitting down in the classroom listen to the lecture ,the mode's change can help student learn by themselves ,and the lecture just help them when the need help.”

Finally, students in all three levels had a number of sentence fragments,

“Compare to study at home or use other methods.”

“The polyester shirts and the winebox.”

“Because everything can be done at home.”

The largest noticeable difference between the writing styles of the students of the three levels was in the use of overall sentence structure. The students’ from Corinth had a tendency to use simple sentences (over 80%) with very few instances of compound structures. The students’ from the Hong Kong cohort used a variety of sentence structures, but with a lower degree of accuracy when these were attempted. Simple sentence usage in this group was no higher than 30% per paper.

The final area of writing that was analyzed was that of vocabulary. The students’ papers were examined for lexical chains as well as use of the Academic Word List. It was discovered that the students from Corinth used lexis that was lifted from the texts given. Any chains that were used were already within the text and very little task appropriate vocabulary was added. The students from Hong Kong had a wider range of vocabulary and were able to add additional vocabulary not found within the texts provided.

When looking at the use of the Academic Word List, it can be seen that the students’ In Hong Kong had a higher usage of the AWL. The vocabulary in the off word list was not discipline specific as would be expected, but rather it was a series of spelling errors that were made. It can be seen in table 2.1 that the students in Corinth had more errors in spelling than those in Hong Kong.

	Corinth				HK			
Elementary/ Intermediate	1-1000	1000- 2000	AWL	Off List	1-1000	1000- 2000	AWL	Off List
	83.8	2.18	4.09	9.83	81.97	5.35	4.78	7.88
	79.88	3.39	8.21	8.49	88.03	3.04	6.49	2.43
	82.87	4.79	3.42	8.9	72.72	5.37	12.8	9.09
	79.73	4.98	3.98	11.29	83.07	2.69	8.07	6.15

Advanced	88.32	2.82	4.17	4.54	81.59	2.48	9.95	5.97
	77.94	3.8	9.31	8.93	81.15	3.62	8.45	6.51
	83.5	3.05	4.68	8.75	74.51	7.34	9.93	8.21
	83.89	3.28	4.3	8.51	82.29	4.66	9.53	3.5

Table 2.1

Upon interviewing the expert raters, they believed that the instrument was not fine-tuned enough to be able to determine the difference between levels. They believed a more detailed instrument would enable them to more accurately score students' writing. In some instances they believed there needed to be more accessible language used in order for it to be understood by the students.

4. Conclusion and implications for further study

The preliminary results have shown that there are still areas for improvement in the scoring rubric. It was discovered that the scoring instrument was not detailed enough to capture all of the error types the students' were making in these three categories. Additional improvements need to be made to the scoring instrument before it will be ready for use by a greater audience, or be able to be incorporated into an automated essay scoring system. The scoring instrument will go through further refinements in the near future. There is potential for a more refined scoring scale, and it would be of great interest to further investigate the usefulness of the rubric with a larger sample size.

As this study consisted of a small population of students' from only two countries, it is not generalizable to students' from other countries. Further investigations will be carried out in the coming years on students from other Asian and European countries. As the study is a work in progress the results should be viewed with that in mind. At the time of writing, a second phase of the research is being carried out to continue improvement of the scoring instrument.

One area for additional investigation could be the teaching methods that are employed in both cities, as the error types the students are making are similar in nature. It would be useful to the linguistic development of the students to see if there are similarities in teaching methodology that leads to the errors or if they are independent of classroom input.

References

- [1] Arter, J. A. (2001). Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance. Thousand Oaks , CA: Corwin Press/Sage Publications.
- [2] Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing* , 26 (2).
- [3] Moskal, B. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. Retrieved December 14, 2014, from Practical Assessment Research and Evaluation: <http://pareonline.net/getvn.asp?v=8&n=14>
- [4] Popham, W. J. (1998). Classroom assessment: What teachers need to know. Needham Heights, MA: Allyn and Bacon.
- [5] Popham, W. (1997). What's wrong - and what's right - with rubrics. *Educational Leadership* , 55, 72-75.
- [6] Schwartz, C. & Sanford-Moore, E. (2008). Implications of the Lexile Framework for Writing for North Carolina's General Model of Writing Assessment. Chapel Hill: Metametrics and The University of North Carolina.
- [7] Simon, M. & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. Retrieved December 15, 2014, from Practical Assessment, Research & Evaluation: <http://pareonline.net/getvn.asp?v=7&n=18>
- [8] Tierney, R. & Simon, M.(2004). What's still wrong with rubrics: focusing on the consistency of performance criteria across scale levels. Retrieved December 18, 2014, from Practical Assessment, Research & Evaluation,: <http://PAREonline.net/getvn.asp?v=9&n=2>



International Conference The Future of Education



- [9] Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.