Analyzing Features of Writing by Chinese College English Learners Based on Juku Correcting Net

王　萍

[齐鲁工业大学（山东省科学院）　济南　250353]

Abstract: The paper examined the writing features of Chinese college English learners under Juku Correcting Net (www.Pigai.org), an automated essay scoring system in English writing. Based on the evaluation results from Juku Correcting Net and close reading, the paper analyzed the writing features mainly from three levels: lexical level such as lexical diversity; syntactical level such as the syntactic complexity; textual level such as organization of the text and cohesive devices employed among sentences and paragraphs. The writing samples were obtained from a writing homework with prompts for ninety-two Year-one law students from a financial university in Shandong Province. The result indicated that with the help of Juku Correcting Net, all students could achieve high spelling accuracy, and most of the writings had similar text structure, but high-scored writings performed much better on lexical diversity, syntactic complexity and cohesive devices. Common error types of all essays were analyzed on different levels respectively. Implications on writing teaching were discussed.

Keywords: Juku correcting Net; automated essay scoring; writing; ESL

Introduction

For ESL (English as a second language) learners, writing is always regarded as the most challenging part, which depends upon both writing ability and second language ability. In China, there are two national English level tests for non-English major college students: College English Test Band 4 (CET4) and College English Test Band 6 (CET6), in which the writing part requests that students should be able to finish a short composition of more than 120 words within half an hour according to a certain topic or outline and the composition ought to have correct writing format, coherent logic content and accurate language use. Previous test results showed that the average score of the writing part was the lowest among listening, reading and translation section. Therefore, developing a high-efficiency way in improving writing skills is an ever-more important concern.

There is a general consensus that the best way to improve one's writing skills is to write, receive feedback from an instructor, revise based on the feedback and then repeat the whole process as often as possible. (Burstein,

Chodorow & Leacock, 2004) In the light of the real teaching situation, an instructor usually has to make feedback to many students in a very short period and the time-consuming work not only puts an enormous load on teachers but also constrains the quantity and quality of feedback to students.

The automated essay scoring (AES) system, also referred to as automated writing evaluation (AWE), was born out of this predicament. AES is defined as "the provision of automated scores derived from mathematical models built on organizational, syntactic, and mechanical aspects of writing" (Ware, 2011) which is primarily intended for large large-scale tests. Coming after it, automated feedback, or "computer tools for writing assistance rather than for writing assessment" is primarily intended for instructional use. (Weigle, 2013). Yet, these two functions are usually combined in recent AES systems.

Starting early in the 1960s, work in this field is productive. Project Essay Grade (PEG), the first version of it being created by Ellis Page in 1973, is one of the earliest and longest-lived implementations of automated

作者简介：王萍，女，汉族，齐鲁工业大学（山东省科学院）外国语学院（国际教育学院）讲师。研究方向：读写能力研究、语言测试。

essay scoring and mainly focuses on the surface linguistic features of a text. Since the 1980s, with the rapid development of artificial intelligence, various commercial automated writing evaluation programs supported by sophisticated language processing technologies have flourished. Valenti ed al. (2003) reviewed some popular ones like Project Essay Grade (PEG), Intelligent Essay Assessor (IEA), Educational Testing service I, Electronic Essay Rater (E-Rater), C-Rater, BETSY, Intelligent Essay Marking System, SEAR, Paperless School free text Marking Engine and Automark. More detailed descriptions of most of the systems are illustrated in Shermis and Burstein (2003). These automated essay evaluation systems are widely applied to the writing instruction and assessment in K-16 education in the United States, which to some extent implicitly assumes that the writers being assessed have native or native-like command of English. However, for non-native language learners, their efficiency remains in further discussion.

Juku Correcting Net (www.Pigai.net), designed by Beijing Speech Technology in China, is one of the family members of automated essay scoring tools, as it were, a representative of Chinese localization in the AES field. Based on corpus and cloud computing technology, it is a free online automated essay scoring system, aiming at reducing the burden of English teachers in marking students' compositions and enabling students to get more practice in writing. After coming online in 2011, it has become the largest English writing platform in China, and more than 1000 universities in China using it as a supplement for English teaching, including Tsinghua University, Fudan University, and Sun Yet-sen University. (Tang et al., 2016) Similar to most AES systems, it carries great advantages like "diagnostic feedback and holistic scores on student writing, immediacy of online comments and positive effects on student autonomy" (Du, 2017). Besides, it allows students to submit multiple times after revising with the feedback from Juku Correcting Net each time before the deadline set by teachers. In addition, Juku Correcting Net also generates a similarity parameter related to plagiarism, which is a result of a comparison of students' compositions with Juku's corpus. A high similarity index indicates plagiarism. After all the students submit their essays, Juku Correcting Net will also demonstrate some good essays written by their peers, thus providing opportunities for students to learn from each other.

In the interface of Juku Correcting Net, sentence-level based feedback provides comments in various aspects falling into different categories: (1) lexical level: verb form, noun form, preposition, collocation, article, incorrect word choice, etc.; (2) syntactic level: tense, subject-verb disagreement, undesirable stylistic features like redundancy and repetition, etc.; (3) textual level: mechanics; (4) formation: spelling, punctuation; (5) learning tips: discrimination of synonyms, high-frequency collocations, etc.. (6) remarks: highlights in sentences and phrases. In the interface of the holistic score, the scale evaluates four aspects: vocabulary, sentence, discourse, and content, and demonstrates a bar graph of the proportion of each aspect accounting for. At the same time, there is a detailed statistical analysis report for different dimensions which can be seen by both teachers and students. In addition, from the teacher port, an overall report can be exported about the information on the same topic, including students' ID, class, submission time, title, content, number of submissions, score, word count, and similarity.

Research about Juku Correcting Net started in the year of 2012. Empirical studies were conducted and showed that the writing quality of the experimental group which used Juku Correcting Net was improved higher than the control group, so that Juku Correcting Net may account for the improvement of students' writing skills. (e.g., Chen, 2013; Gu & Wang, 2012; Hu, 2015; Liao, 2016; Liu & Shen, 2015). When it comes to reliability and validity, He (2013) claimed that the system was very reliable, but the scores given by it were higher than those by teachers. It can provide sufficient information on the vocabulary use and grammatical points but cannot provide much evaluation on their organizations, styles, coherence, and cohesion. The result is consistent with the argument of Chung and Baker that "high reliability or agreement between automated and human scoring is a necessary, but insufficient condition for validity" (Chung & Baker, 2003).

In respect of attitude toward Juku Correcting Net, Chen (2014) conducted a questionnaire survey and follow-up interviews, and the results showed that most students gave favorable remarks to the new form of focused feedback. Motivated by the score, they would like to revise again and again with feedback, which has a positive effect on their control of specific structures. Some students

mentioned they also wanted and needed more feedback on content and organization. Tang et al. (2016) interviewed the instructor in their study, and the instructor stated that in the process of teaching English writing, appropriate human feedback was needed for the human factor still played an important role in the teaching process.

At present, more and more universities join the army of Juku Correcting Net, which indicates that the traditional way of helping students to improve writing skills may have been changed with the great help of such an automated essay evaluation tool. Thus, the purpose of this study is to investigate the features of writing by Chinese College English learners based on Juku Correcting Net. There are four research questions to be addressed for this study:

1.What kind of features does the final submission version of college students' essays take on in terms of lexical level, syntactical level, and textual level under the help of Juku Correcting Net?

2.What are the common errors that students commit?

3.What kind of feedback do students need from teachers?

4.What are the implications for classroom teaching in English writing?

1.Methodology

1.1 Research Subjects

The study took the data from one writing prompt in April 2017 of year one law student from a finance university in China. There are 92 compositions in total coming from four different classes of the same English teacher. Students are required to write around 200 words on a topic: Are We More Connected or More Alone. They can submit their writings as many times as possible, which means they can revise their writings according to the feedback of Juku Correcting Net until they feel that they get a satisfactory score. In this case, students usually revise it several times for the score given by Juku Correcting Net will be their final score for this assignment, except for similarity problems. Students with high similarity can still obtain a quite good score, for the score given by Juku Correcting Net is based on language features, so a high similarity essay may obtain a high score but have an issue of plagiarism. If the similarity index surpasses 30%, the teacher of these four classes will just fail the student (different teachers may deal with scores given by Juku Correcting Net in different ways). Of the 92 essays, most of the essays have a similarity index under 20%, but there

is one case where the similarity index was 85.44%, which means this essay might be a copy of others and cannot reflect the real level of the student. Thus, this case was taken as an outlier and was removed from the database.

1.2 Data collection

The data was collected by logging on to the teacher's account with the teacher's permission. Firstly, from the interface of the teacher port, every student's writing is available with detailed sentence-level comments. Second, from the holistic score interface, it can be seen that the full mark is 100. What can also be obtained is the statistical analysis report about word count, the average length of words, the proportion of academic words, coherent vocabulary, number of paragraphs, and number of subordinate clauses of each student. What's more, there is a diagnostic interface showing the summary of error types of this assignment. Finally, there is an overall report exported about this assignment, including students' ID, class, submission time, title, content, the number of submissions, writing score, word count, and similarity.

1.3 Data Analysis

With the statistics provided by the system, both the quantitative method - statistical analysis by SPSS 25.0 and qualitative method - close-reading to analyze the writing features of these college students were applied.

Firstly, a multiple regression was conducted and examined the relationship between the number of submissions, word count, similarity, and writing score. Second, according to the overall report given by Juku Correcting Net, two groups of compositions were selected—the writing score above 90 (including 90) as Group A and writing score below 80 (not including 80) as Group B. Detailed analysis will be performed from lexical level, syntactic level, and textual level respectively and common types of errors are also discussed.

2.Result and Discussions

2.1 The Influence of Word Count, Number of Submissions, and Similarity on Writing Score

Multiple linear regression was calculated to predict writing scores based on word count, number of submissions, and similarity. In Table 1, the result showed that only word count was statistically significantly and positively related to writing score(p=.000)with a standardized coefficient of .632. This indicates that for every standard deviation increase in word count, the writing score goes up by .632 standard deviations. It is not

hard to understand that word count plays such an important role, because it carries substantial information on linguistic features which is related to writing skills.

Table 1. Hierarchical Regression of Writing Score on Word Count, Number of Submissions, and Similarity

	b	SE B	β	p
Step 1				
Constant	74.428	1.531		p= .000
Word Count	0.057	0.007	.673**	p= .000
Step 2				
Constant	72.787	1.542		p= .000
Word Count	0.054	0.07	.639**	p= .000
Number of Submissions	0.021	0.015	.117**	p= .154
Step 3				
Constant	73.011	1.696		p= .000
Word Count	0.053	0.007	.632**	p= .000
Number of Submissions	0.021	0.015	.116**	p= .160
Similairity	-0.023	0.072	-.027**	p= .747

Note. R^2=.453 for Step 1; ∇R^2=.013 for Step 2 (ps>.05); ∇R^2=.001 for Sept 3 (ps>.05) (two-tailed).

The number of submissions was found to be not statistically significant to the writing score. It is in consistency with the result of Hu(2015). Sometimes students only revise very few times which does little influence on the final score; sometimes the number of submissions can be very large, like a student who submitted 71 times in Hu's study. In later interview, the student said he was just fooling the system, and curious about the change of score. In the current study, there is also one student who was found to submit 170 times and she got the highest score. Despite this, the number of submissions and writing scores are not or should not be significantly related. Maybe at the initial stage, revising times can predicate the increase in writing score, but the final score ultimately depended on the writing ability. When the writing ability reaches its maximum value for a student, the number of submissions has nothing to do with the writing score.

Similarity was also not statistically significant to the writing score. As is mentioned before, the score given by Juku Correcting Net is based on language features, so even an essay with a high similarity index may obtain a good score. In other words, Juku Correcting Net is a system to evaluate the language quality, but not a judge to give a sort of fairness or justice.

2.2 Linguistic Features on Lexical Level
2.2.1 Lexical Features of Group A and Group B

According to the diagnostic report given by Juku Correcting Net, a comparison between Group A(the writing scores are all above 90)and Group B(the writing scores are all below 80)based on their performance on the lexical level was made, as shown in Table 2 and Table 3, and the mean score of each index is illustrated in Table 4.

Table 2. Group A Performance on Lexical Level

Name	Writing Score	Words Count	Type Token Ration	Spelling Accuracy	Average Word Length	High-frequency Vocabulary	Academic Words	Verb Phrase
A1	93.5	295	7.47	1	4.78	64%	16%	17
A2	91.5	292	6.76	1	4.65	63%	16%	17
A3	91	291	8	1	4.99	72%	10%	28
A4	90.5	289	7.3	1	4.7	66%	10%	23
A5	90.5	292	6.33	1	4.73	70%	13%	25
A6	90	251	6.58	1	4.85	65%	10%	12
A7	90	275	6.68	1	4.95	63%	12%	19

Table 3. Group B Performance on Lexical Level

Name	Writing Score	Words Count	Type Token Ration	Spelling Accuracy	Average Word Length	High-frequency Vocabulary	Academic Words	Verb Phrase
B1	77	195	5.4	0.97	4.76	74%	11%	11
B2	77	157	5.66	0.98	4.3	82%	7%	9
B3	77	161	4.79	1	4.15	79%	3%	10
B4	76	242	6.07	0.99	4.36	72%	4%	18
B5	74.5	132	4.81	1	4.38	89%	4%	5
B6	72.5	128	5.06	1	4.34	84%	5%	9

Table 4. Mean Score in Different Indice of Group A and Group B

Group	Writing Score	Words Count	Type Token Ration	Spelling Accuracy	Average Word Length	High-frequency Vocabulary	Academic Words	Verb Phrase
A	91	284	7.01	1	4.81	66.14%	12.43%	20
B	75.67	169	5.30	0.99	4.38	80%	5.67%	10

As the above tables show, there is a significant discrepancy between the two groups in different indexes. It has been mentioned before that in the multiple regression analysis, word count is statistically significantly related to the writing score, for the number of words carries different information on vocabulary and sentence. It is not the number that matters, but the internal quality of the essay counts.

Type Token Ration (TTR) is an index that can reflect lexical diversity which refers to "the range of different words used in a text, with a greater range indicating a higher diversity" (McCarthy & Jarvis, 2010). Lexical diversity deals with the vocabulary range or even more

than that. It is thought to be an important measure of text difficulty. Put it in another way, a text with a higher index of lexical diversity is likely to be more complex, more advanced, and more difficult (if other things are equal). It can be seen that students in Group A have a much better command of vocabulary range than in Group B. Meanwhile, the average word length and the index of verb phrases shown in the table are also strong proofs. It is true indeed by a close reading of Group A's compositions. Students with high writing scores are more capable of using various noun phrases, preposition phrases, adjectives, and adverbs.

In terms of spelling accuracy, the two groups are almost the same. With the help of Juku Correcting Net, most spelling errors can be identified and corrected. Different from Microsoft Word, no mistake is automated corrected by Juku Correcting Net and it means that students have to make corrections by themselves. Thus, it would be no wonder that there are still some students not achieving perfect spelling accuracy. For example, in student B1's composition, the sentence-level feedback for one sentence is like this:

The left part is the sentence written by student B1 and the right part is the comments from Juku Correcting Net. Below is a translation version for the comments here:

✗[Wrong words]: Please check the spelling of "arevconnected".

📖[Learning tips]: Confusing vocabulary: "little", "few", and" several" all have a meaning of "not many".

It can be seen that if the student followed the instructions and proofread them, he might achieve perfect spelling accuracy. It is also true of students B2 and B4. Thus, with the help of Juku Correcting Net, if students are conscientious enough, teachers do not have to worry about spelling issues necessarily.

In terms of academic writing, the proportion of high-frequency vocabulary and academic words deserves attention as well. The result showed that Group A has a lower proportion of high-frequency vocabulary but a higher proportion of academic words than Group B. A high-frequency word is usually one that people will encounter regularly in reading; it is one of a small number of common words that make up a majority of any English text, such as "he, she, you, I, ask, is, but, the, have, and

good". However, in writing, the repeated appearance of high-frequency words is not qualified especially for argumentative writing. On the contrary, academic words are encouraged to show up as much as possible. Group B shows an obvious deficiency in academic writing.

2.2.2 Common Error Types on Lexical Level

Based on a close reading of students' compositions, the main lexical error types are summarized.

Table 5. Error Types on Lexical Level

Error Types	Examples
Noun Plural	Example 1: I hold that if we can hold some individuals in our heart and have a face to face party, we will cope with the problem of connecting. Example 2: The blessing of the 100 network can not compare with our personal greetings, the network teaching at home will never have the atmosphere of the classroom. Example 3: Many person holds the view that we are more connected now on the Internet. Example 4: These softwares have a very large benefit which is the elderly and children can use it too.
Preposition	Example 5: Secondly, it is universal for us to play tablet PC or phone on the bus, on the subway and other public occasions. Example 6: The reason of aloneness that they aren't making use of the Internet properly.
Part of Speech	Example 7: Nowadays' society, a variety of electronic equipment becomes more and more population, which is the reason why we be used to chat online by electronic equipment. Example 8: To restore balance to the life, to relief the sense of isolation, you are supposed to force yourself back to the real world.
Article	Example 9: Little by little, the condition has aroused the public concern that needs rethinking it's benefit or not. Example 10: The first of all, speak out our mind directly with no worry about the embarrassment which may happen in face to face talk.
Collocation	Example 11: From another point of view, we should combine "more connection" and "more loneliness". Example 12: In the future, we will become more and more lonely, because we lost in the virtual network in Los Angeles, the constant reduction of contact between people, life slowly become virtual, untrue.

Noun plural is the first grammar point for most Chinese English learners, which is also one of the biggest language differences between English and Chinese. Although the principle and rules are easy to understand, it is often neglected by students. Errors in Example 1 "in our heart (hearts)" and Example 2 "100 network (networks)" might stem from students' carelessness, for they sometimes are unconscious of or not sensitive enough to the plural form in English. These errors can be identified by Juku Correcting Net or can be avoided with a double check by students themselves. While Example 3 "many person (people) holds" and Example 4 "These softwares (software)" illustrate another situation that students are aware of noun plural form of English words but maybe not clear about the plural form of certain words.

As for preposition errors, this study classifies the cause together with that of collocation. For one reason, students might not well master the natural combinations of words in their English learning process, like Example 5 "it is universal for (to) us" and Example 11 "combine 'more connection' and (with) 'more loneliness'"; for another, when dealing with expressing complicated ideas, they choose to do some literal translation, which might be influenced by their mother tongue - Chinese, like Example 6 "the reason of (for) loneliness" and Example 12 "constant reduction of contact".

Further, for the errors in part of speech, the misuse of nouns to replace verbs and adjectives is pervasive in low score writings, as shown in Example 7 "becomes more and more population (popular)" and Example 8 "to relief". While as for the use of article like "a", "an" and "the", it is usually difficult to achieve perfection for both high-score writings and low-score writings.

2.3 Linguistic Features on Syntactic Level

2.3.1 Syntactic Features of Group A and Group B

For college students, various English sentence structures have been taught during high school, such as infinitive phrases, gerund phrases, participle phrases, adverbial clauses, complementizer clauses, and relative clauses. Based on a close reading, the total number of different types of clauses in the compositions of Group A and Group B are counted, as shown in Table 6, Table 7, and Table 8.

Table 6. Group A Performance on Syntactical Level

Name	Sentence Length	Number of Sentences	Simple Sentences	Compound Sentences	Subordinate Clauses	Complex Phrase Structures
A1	25.31	13	2	0	12	15
A2	28.08	14	2	2	12	9
A3	23.57	14	3	2	11	11
A4	23.57	14	5	3	8	10
A5	23.57	14	5	2	12	11
A6	26.27	11	2	6	6	2
A7	22.71	13	5	1	9	18

Note. Subordinate Clauses include adverbial clause, relative clause, and complementizer clause; complex phrase structures include infinitive phrases, gerund phrases and participle phrases.

Table 7. Group B Performance on Syntactical Level

Name	Sentence Length	Number of Sentences	Simple Sentences	Compound Sentences	Subordinate Clauses	Complex Phrase Structures
B1	16.69	12	6	5	3	4
B2	15	12	6	0	4	2
B3	20.78	9	4	2	7	4
B4	19.07	14	5	1	13	3
B5	16.56	9	1	0	11	1
B6	12.33	12	10	0	1	2

Note. Subordinate Clauses include adverbial clause, relative clause, and complementizer clause; complex phrase structures include infinitive phrases, gerund phrases and participle phrases.

Table 8. Mean Score of Different Indice of Group A and Group B

Group	Sentence Length	Number of Sentences	Simple Sentences	Compound Sentences	Subordinate Clauses	Complex Phrase Structures
A	24.73	13.29	3.43	2.286	10	10.86
B	16.74	11.33	5.33	1.33	6.5	2.67

It is obvious that in high-score writing, more complex syntactic structures are used, such as infinitive phrases, participle phrases, and gerund phrases. Various strands of research have indicated that syntactic complexity is positively related to the quality of the written composition. (Marion, 1983). There are a variety of ways to measure the syntactic complexity - sentence length, clause length, ratio of subordinate clauses to all clauses, and T-unit. Here, the average sentence length was given by Juku Correcting Net. In addition, other data related to syntactic complexity was counted and calculated by the researcher. It can be noted that high score writings have done a much better performance on the average sentence

length, which is consistent with former research studies. Besides, the amount of complex structures in Group A is far more than that in Group B, which makes compositions in group A more difficult than those in Group B.

What's more, through the comparison of two groups in Table 8, in terms of the choice of sentence type, two groups show a discrepancy. We can see that students in Group A prefer to adopt various subordinate clauses which increased the richness of sentence type and difficulty of composition, while students in Group B are inclined to use simple sentences. It should be noted that in Table 7, student B4 and student B5 also used quite a lot of subordinated clauses. Yet with a detailed reading, it was found that the types of subordinate clauses clustered in complementizer clauses. The extent to the grasp of clause types indicates distinctive proficiency in English writing.

2.3.2 Common Error Types on Syntactic Level

Unlike error types in the lexical level, Juku Correcting Net has relatively limited capability in identifying syntactic error types. The report of Juku Correcting Net concluded that ill-organized sentences accounted for the largest proportion of all types of error, yet it does not give detailed comments on different types except subject-verb disagreement and run-on sentences. Through reading in-depth, the common error types on the syntactic level were summarized as shown in Table 9.

Table 9. Error Types on Syntactic Level

Error Types	Examples
Subject-Verb Disagreement	Example 13: Furthermore, Internet make our communication with educators and students all over the world, which has broken through distance and makes our more connect.
Fragments	Example 14: For example, if we have a friend far away.
Run-on Sentences	Example 15: From my perspective, we are hooked on the virtual world and overlook the true life, the real world disappears, and the nerds, contemptible person who lack social skills are emerging.
Garbled Sentences	Example 16: Second, exposure privacy, people a lot of privacy and sensitive information leaking out through the network.

Among these error types, subject-verb disagreement accounts for the largest proportion: it is found to appear 24 times in a total of 66 ill-organized sentences of all 91

compositions. Although in Chinese there are no inflections on verbs, most of the students know the rule of subject-verb agreements in English, as is shown in Example 13 and Example 14. There are two clauses in each of the sentences, but students are half correct in two subject-verb agreements, which indicates that their error is due to carelessness. However, errors in respect of fragments, run-on sentences, and garbled sentences reveal students' weakness in syntactic rules. In Example 14, there is a lack of the main clause; in Example 15, there is no conjunction between clauses; in Example 16, it seems that the student does not know the basics of syntactic rules to make a meaningful sentence, so he just put many words there without grammatical rules.

It is also a pity that Juku Correcting Net does not offer many suggestions for revising, such as

Example 17
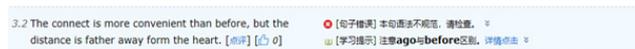
Translation for the comment on the right side:

✗ [Wrong sentence]: This sentence is not grammatically correct. Please check again.

📖 [Learning tip]: Please pay attention to the difference between "ago" and "before". To see detail.

In this case, even though students know that the sentence needs to be revised, there is no direction he can follow.

One more deficiency of Juku Correcting Net is that it cannot identify all the ill-organized sentence structures, for instance,

Example 18: By comparison with the outdated interpersonal communication as setting apart the weekend for shopping, having appointments or eating with friends and co-workers.

Example 19: To be honest, connected with other people online is actually convenient.

Example 18 was written by student A7 who got a high score of 90, but the sentence-based comment by Juku Correcting Net did not point out anything wrong with this sentence. The same situation is true of Example 18. On this occasion, Juku Correcting Net is not as valid as we expected.

2.4 Linguistic Features on Textual Level

2.4.1 Textual Features of Group A and Group B

On a textual level, Juku Correcting Net only gives the paragraph of each comparison and the inaccurate number of coherent words. To investigate the features of these texts,

the cohesive devices applied in Group A and Group B were counted. As is shown in Table 10, Table 11and Table 12.

Table 10. Group A Performance on Textural Level

Name	A1	A2	A3	A4	A5	A6	A7
Number of Paragraph	3	4	4	4	3	3	5
Cohesive Devices	8	5	7	4	8	4	7

Table11. Group B Performance on Textual Level

Name	B1	B2	B3	B4	B5	B6
Number of Paragraph	3	2	3	4	3	3
Cohesive Devices	5	4	6	2	2	5

Table12. Cohesive Devices in Group A and Group B

Name	Cohesive Devices
Group A	currently; firstly; secondly; hence; besides; thus; what's more; therefore; in a word; on the contrary; nevertheless; whereas; briefly; in the main; nonetheless; admittedly; from another perspective; accordingly; furthermore; generally speaking; moreover; frankly; last but not least; all in all; believe it or not; meanwhile; however; it has to say; additionally; ultimately
Group B	to be honest; for example; what is more; however; firstly; on the other hand; lastly; what's more; on one hand; all in all; first and foremost; additionally; of course; most of all

From Table 10 and Table 11, we can see that all of the compositions are composed of three to five paragraphs. Based on the detailed reading, it has been found that most discourse structures are following the classical argumentative writing pattern: claim, data, and conclusion. The reason for this can be explained by the standard training model in wring classes and writing templates designed for CET-4 (college English Test Band 4) and CET-6 (college English Test Band 6) from the Internet and various practice books. From Table 12, concerning cohesive devices, it is obvious that Group A is better at employing cohesive devices than Group B, which can be attributed to the vocabulary usd by each Group, or the extent of attention paid by different groups to the cohesive devices.

2.4.2 Common Error Types on Textual Level

Juku Correcting Net does not provide any discourse guidance for students, but there are some problems necessary to be addressed on the textual level or discourse organization indeed, such as one-sided structure and overuse of cohesive devices.

For English argumentative writing, the British philosopher Toulmin once proposed a model of argument structure, which is known as Toulmin model and is widely used in teaching and argumentative writing research. It includes six elements: claims, data, warrants, qualifiers, rebuttal, and backing. Qualifiers refer to statements that limit the strength of the argument or statements that propose the conditions under which the argument is true, and rebuttals refer to counter-arguments or statements indicating circumstances when the general argument does not true. It would be more persuasive to employ counterarguments or opposite views in argumentative writing. (Du, 2017)

However, it seems that most of the ninety-one compositions are organized in one sided-structure, which means that students focus only on how to state their main claim and provide data, but neglect the use of counterargument structure, or in other words, opposing viewpoints in their argumentative writing. Here counterargument is coded as "the possible opposing views than can challenge the validity of a writer's claim" (Du, 2017). Below are showing two examples of students' writing for each structure respectively. Example 20 is for one-sided structure and Example 21 is for counterargument structure.

Example 20:

...From my view, we are more connected. The first of all, speak out our mind directly with no worry about the embarrassment which may happen in face to face talk. Secondly, we can communicate with our family, friends or anyone else with no payment. The most of all, we do not have to worry about the regional restrictions. For example, if we have a friend far away. It is impossible to meet frequently. Our ability to connect with other people online makes us feel like a part of something bigger…

Example 21

…Frankly, the booming of the virtual life mirrors the progress of our society. There is no denying the fact that we could keep in touch with others conveniently however the distance is, which is owned by the virtual life. Nevertheless, in my standpoint, it's the booming virtual life that people feel more alone in real life. Many reasons can account for it. Firstly…

In Example 20, the student just stated his claim and provided data for why he believed "we are more

connected", while in Example 21, the student used the counterargument first and then demonstrated her claim and data in the following, which made her argument very persuasive. It should be noted that only one-third of students are aware of the employment of counterarguments in their composition. However, it has been found that "the overall quality of English argumentative essays were not correlated with fundamental elements, claim and data, but was positively correlated with the uses of counterargument and rebuttal" (Du, 2017).

Another prominent problem is the overuse of cohesive devices. Cohesive devices can indeed make clear the relationships between different clauses, sentences, and paragraphs, but overusing them not only destroys the fluency and coherence of writing but also causes logic confusion. Example 22 is an extract from a student's writing.

Example 22

At present, social network platform has become a prevailing trend gradually, such as QQ, we-chat, face-book and so on. Moreover, people are occupied with their work and they don't spend some time in making friends. In consequence, individuals become addicted to the virtual world and feel lonely in mind.

In this short paragraph, there are three sentences in total, but cohesive devices also appeared three times. It should be noted that "moreover" here actually does not mean anything, because there is no parallel or progressive relationship with the first sentence. Meanwhile, "in consequence" here does not show any causality between this sentence and before. Overuse of cohesive devices also appears in high-score writing, for instance:

Example 23

Firstly, the Internet, which handles an astonishing amount of information, is an imperative in our daily life. Hence it goes without saying that it does give rise to the lifting of our work efficiency. Besides, with the development of science and technology, it serves as an appropriate platform where people have access to exchanging standpoints on diverse issues, producing one part of public opinion. Thus, given the fact above, it is easy to draw a conclusion that it enables us to obtain a comprehensive viewpoint from a variety of angles. What's more, chatting online is expected to be an exceptional option on account of its convenience in communicating with our old friends and making new friends. Therefore, there is no doubt that it is social networking websites that

get folks mutually involved and make a big difference. Last but not least, donating money via the Internet to high-poverty areas and regions where there is a natural disaster, especially a severe catastrophe, is a feasible way to contribute to our society.

The extract above is taken from the composition of student A1, the highest score writing among ninety-one compositions. Similarly, every sentence begins with a cohesive device, and the most prevailing logic is parallel which is indicated by "firstly, besides, what's more, last but not least" and causality which is indicated by "hence, thus, therefore". On one hand, these various cohesive devices seem to make the logic between lines clear; on the other hand, it interrupts the reader's thinking constantly and destroys the fluency of the whole paragraph in part.

3. Conclusions

With the data report of Juku Correcting Net, it has been found that among word count, number of submissions, and similarity, only word count is statistically significantly related to writing score. Through a comparative analysis of high score writing group and low score writing group, linguistic features vary between the two groups on different levels. On the lexical level, all the students can achieve high spelling accuracy, which can be explained by constant revising with feedback from Juku Correcting Net, while the high score group has a better performance on lexical diversity and they are more capable in academic word use. On the syntactic level, as far as syntactical complexity is concerned, various types of clauses and phrase structures increase the richness of sentence type. Yet sentence types in the low score writing group are relatively monotonous. On the textual level, both groups have a similar text structure, but the high score group is better at employing various cohesive devices.

A comprehensive analysis is conducted to find out common error types on the three levels respectively in the ninety-one compositions, which can give some implications on what kind of feedback should be provided by teachers. For students at lower levels of language proficiency, the focus of feedback can be generally on linguistic issues; that is, the degree to which writers have control over basic vocabulary, syntax, and text structure. For students at higher levels of language proficiency, feedback from the teacher can shift to higher order concerns such as development, the strength of argument, and precision in language use.

Juku Correcting Net works well in evaluating and commenting on surface linguistic features such as lexical characteristics or some simple syntactic structures, but its capability is quite limited for deep linguistic features especially the intrinsic logic within sentences and text. Language teachers are still playing significant roles in the process of students' writing. The idea hybrid may be one in which using the score given by Juku Correcting Net as an evaluation of vocabulary and syntax performance, leaving the teacher free to comment on higher-order concerns.

Bibliography

[1]Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online writing service. *Ai Magazine*, 25(3), 27-36.

[2]Chen, W. (2014). *A study on feedback in college English writing based on automated scoring system-In case study of Juku* [Master's thesis, Chongqing University https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201401&filename=1013044819.nh

[3]Chung, G. K., & Baker, E. L. (2003). Issues in the reliability and validity of automated scoring of constructed responses. In M.D. Shermis & J.Bustein(Eds.), *Automated essay scoring: A cross disciplinary approach* (pp. 23-40). Lawrence Erlbaum Associates Publishers.

[4]Crowhurst, M. (1983). Syntactic Complexity and Writing Quality: A Review. Canadian Journal of Education / Revue Canadienne De L'éducation, 8(1), 1-16.

[5]Du, F.Y. (2017). The analysis of argument-counterargument structure in Chinese EFL learners' argumentative writing. *Journal of Studies in Education*, 7 (3), 121-129.

[6]Gu, C. H. & Wang, L. (2012). An empirical study of college English writing teaching based on Juku correcting network. *Journal of Yangzhou University (Higher Education Study Edition)*, 16(4), 92-96.

[7]He, X. L. (2013). Reliability and validity of the assessment by the Pigaiwang on college students' writings. *Modern Educational Technology*, 23(5), 64-67.

[8]Hu, X. W. (2015). Effects of online self-correction on EFL students' writing quality. *Computer-assisted Foreign Language Education*, 163, 45-49.

[9]Liao, H. C. (2016). Using automated writing evaluation to reduce grammar errors in writing. *ELT Journal*, 70(3), 308-319.

[10]Liu ,J. & Shen, Y. (2015, January). A case study of text and discourse based on Juku, Coh-Metrix, and linguistic inquiry word cunt (LIWC). International Conference on Social Science and Technology Education (ICSSTE 2015). Atlantis Press, 1046-1050. https://doi.org/10.2991/icsste-15.2015.261

[11]Tang, Z.Y., Cheng, L., & Chen, C. (2016). Improving engineering students' writing skills in a computer based course using an automated writing evaluation system. In Z.Lu, Y Lu & G.M. Davis (Eds), *Tertiary education: Issues and perspectives from Asian Contexts*. Hong Kong Polytechnic University.

[12]McCarthy, P.M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2): 381-392.

[13]Shermis, M. D.,& Burstein, J. C.(2003). *Automated essay scoring: A cross-disciplinary perspective*. Routledge.

[14]Valenti, S., Neri, F., & Cucchiarelli, A. (2003). An overview of current research on automated essay grading. *Journal of Information Technology Education*, (2), 319-330.

[15]Ware, P. (2011). Computer-generated feedback on student writing. *TESOL Quarterly*, 45(4), 769-774.

[16]Weigle, S. C. (2013). English as a second language writing and automated essay evaluation. In Sherims, M. & Burstein, J. (Eds.). *Handbook of automated essay evaluation: Current applications and new directions*(pp.36-54). Routledge.

[17]Zuo, Y.J. & Lei, F. (2015). A study on dimensions of analytic rating scales in college English writing-Based on a comparative study of writingroadmap and Pigai. *Modern Educational Technology*, 25(8), 60-66.