ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 5, No. 9, pp. 1929-1934, September 2015
DOI: http://dx.doi.org/10.17507/tpls.0509.23

A Coh-metrix Study of Writings by Majors of Mechanic Engineering in the Vocational College

Ronggen Zhang
Shanghai Publishing and Printing College, Shanghai, China

***Abstract*—Th*rough Co-Metrix analyses of the writings of the mechanic vocational students, some findings are as follows: First, the students prefer to use more concrete and familiar content words. Second, they seldom use pronouns and parallel structures, as a result, the writing is not central to the main idea and not coherent enough. Third, those in favor of simple sentences tend to prefer similar sentence structures and repetitive tenses and aspects. Fourth, they prefer to use overlapping pronouns and content words. Finally, the students use hypernymous verbs can make the writing more diverse in sentence structure. In summarizing the findings, the paper also gives some suggestions on how to improve the writing proficiency of the students.**

***Index Terms*—English exposition, corpus, Co-metrix, writing proficiency**

I. Introduction

Coh-Metrix is a computational tool that produces indices of the linguistic and discourse representations of a text. Coh-Metrix was primarily used to improve reading comprehension in classrooms. Then it is used to calculate the coherence texts on many different measures and to study vocabulary (Graesser et al., 2004, 2011, 2013).

Duncan (2008) uses Coh-Metrix to provide a computational linguistic analysis of the English-language biomedical research abstracts and posit how these conventions differ among native and nonnatice speakers of English. Chen (2009) finds that, when writing on tasks with richer contextual features, students tended to better display their writing ability; test-takers who were more proficient tended to write more fluently, to use more complex syntactic structures, to demonstrate greater linguistic accuracy, to adapt to audience traits and expectations more sensffively, and to achieve their writing purpose more effectively; and more proficient students tended to be aware of the genre and the style of their writing.

Qin. & Gu (2011) finds that topic familiarity exerts no significant effect on the overall use of connectives, yet causes a significant difference in both referential indices measure and lexical repetition measure. Wang (2011) finds that textual features determines L2 writing professiocny with a percentile of 25.6 coefficient of determination.

Chen & Du (2012) finds postgraduates improved significantly in lexical difficulty degree, lexical chunks and complexity of sentences, but improved little in cohesion in their writings.

Li (2012) finds there is a relationship between discourse coherence and writing level for the Chinese university students, but there is no significant difference between high level group and low level group. Du & Cai (2013) reports a Coh-Metrix-based study of the linguistic features that influence the argumentative writing quality of English majors in China, and finds that readability lexical frequency and cohesion are correlated with the writing score to some degree.

He (2013) finds that high frequency words are used more often in TEM-8 than in IELTS, while low frequency words are less densely used in TEM-8 than in IELTS.

The literature review above shows that Coh-Metrix study in China is mainly focused on university undergraduates' and postgraduates' reading and writing. With regard to such a case, this study aims at China's vocational college students' writing, especially that of the mechanic engineering majors:

a) What linguistic features the English writings by China's vocational college students may display.

b) How those linguistic features contribute to the writing professiocny of the students.

c) Pedagogical implications of the findings.

II. Methodology

A. Sampling: The corpora concerned are based on the 91 pieces of students' writings, randomly sampled among the 600 pieces of writings from the placement test for freshmen majors of printing mechanic engineering in Shanghai Publishing & Printing College in September of 2013. The topic of the writing is an expository composition on 'No Smoking in Public Places', with each student given a picture in which many people are smoking at a restaurant at the time. And the writing is required to be finished within half an hour.

B. Data processing: including scoring, concordancing, and editing, by using the software such as Coh-Metrix 3.0, SPSS 19, etc. Each piece of writing is scored through the scoring system provided by http://pigai.org/guest.php, just for reference.

C . Concepts concerned in data processing:

DESSC - Sentence count, number of sentences
DESWC - Word count, number of words
DESSL - Sentence length, number of words, mean
DESWLlt -Word length, number of letters, mean
DESWLltd - Word length, number of letters, standard deviation
PCNARp - Text Easability PC Narrativity, percentile
PCSYNp - Text Easability PC Syntactic simplicity, percentile
PCCNCp - Text Easability PC Word concreteness, percentile
CRFAO1 - Argument overlap, adjacent sentences, binary, mean
CRFAOa -Argument overlap, all sentences, binary, mean
CRFCWO1 - Content word overlap, adjacent sentences, proportional, mean
CRFANP1 - Anaphor overlap, adjacent sentences
CRFANPa - Anaphor overlap, all sentences
LDTTRc - Lexical diversity, type-token ratio, content word
LDTTRa - Lexical diversity, type-token ratio, all words
SMTEMP - Temporal cohesion, tense and aspect repetition, mean
SYNNP - Number of modifiers per noun phrase, mean
SYNSTRUTa - Sentence syntax similarity, adjacent sentences, mean
SYNSTRUTt - Sentence syntax similarity, all combinations, across paragraphs, mean
WRDFAMc - Familiarity for content words, mean
WRDCNCc - Concreteness for content words, mean
WRDHYPn - Hypernymy for nouns, mean
WRDHYPv - Hypernymy for verbs, mean
RDFRE - Flesch Reading Ease
RDFKGL - Flesch-Kincaid Grade Level
RDL2 –Coh-Metrix L2 Readability

III. Results and Analysis

A. Descriptive Statistics

TABLE I.
DESCRIPTIVE STATISTICS

	Minimum	Maximum	Mean	Std.Deviation		Minimum	Maximum	Mean	Std.Deviation
CRFCWO1	0.00	0.42	0.13	0.08	RDFKGL	2.34	33.89	5.37	3.38
SYNSTRUTt	0.00	0.27	0.15	0.05	WRDHYPn	3.01	6.63	5.68	0.58
SYNSTRUTa	0.00	0.29	0.16	0.05	DESSC	1.00	17.00	9.41	2.80
CRFANPa	0.00	0.67	0.17	0.14	DESSL	5.31	86.00	12.32	8.54
CRFANP1	0.00	1.00	0.32	0.21	RDL2	12.16	45.18	27.99	6.81
CRFAOa	0.00	0.82	0.41	0.18	PCCNCp	0.06	99.95	42.35	27.46
CRFAO1	0.00	1.00	0.44	0.23	Score	12.00	70.00	50.29	13.26
SYNNP	0.19	1.00	0.57	0.16	PCSYNp	0.00	99.51	75.61	23.54
LDTTRa	0.41	0.75	0.61	0.07	RDFRE	1.70	94.32	78.09	10.84
LDTTRc	0.48	0.89	0.73	0.09	PCNARp	15.87	99.63	79.87	16.47
SMTEMP	-2.00	1.00	0.77	0.31	DESWC	46.00	163.00	100.98	21.98
WRDHYPv	0.87	2.16	1.36	0.22	WRDCNCc	317.95	433.91	366.84	24.09
DESWLltd	1.67	2.78	2.10	0.20	WRDFAMc	573.35	603.38	592.13	5.56
DESWLlt	3.74	5.31	4.22	0.23					

Table 1 shows, the overall writing professiocy of the students is far from satisfactory, with a mean of 50.29, and a Std.Deviation of 13.26; that is, they are poor in English, especially at uneven levels (the lowest scoring 12 while the highest scoring 70). It may be a common tendency in China that the quality of college students, especially that of the vocational students is worsening. Table 1 further shows, for one thing, the 5 top linguistic components of the students' writings are: Familiarity for content words, Concreteness for content words, number of words, Text Easability PC Narrativity (percentile), and Flesch Reading Ease. This means that they prefer to use more concrete and familiar content words, which leads to the easiness of the writings. And the average number of words in a writing is 100.98, which is far below the lower standard of CET-4, college English test in China, with a 120-to-150-word writing within half an hour. For another, the 5 lowest linguistic components of the students' writings are: Content word overlap, Sentence syntax similarity both across paragraphs and in adjacent sentences, Anaphor overlap, and Argument overlap. This means they seldom use pronouns and parallel structures in their writing, thus the writing is not central to the main idea and not coherent enough.

B. Pearson Correlation

TABLE II
PEARSON CORRELATION FOR LINGUISTIC FEATURES AND WRITING PROFESSIONCY

DESSC	0.17	CRFAOa	0.06	SYNSTRUTt	-.241*
DESWC	.678**	CRFCWO1	-0.15	WRDFAMc	-.367**
DESSL	0.00	CRFANP1	0.18	WRDCNCc	0.10
DESWLlt	0.01	CRFANPa	0.02	WRDHYPn	.358**
DESWLltd	0.17	LDTTRc	0.17	WRDHYPv	-0.09
PCNARp	-0.08	LDTTRa	-0.03	RDFRE	0.00
PCSYNp	-0.13	SMTEMP	0.21	RDFKGL	0.00
PCCNCp	0.14	SYNNP	.230*	RDL2	-.210*
CRFAO1	0.06	SYNSTRUTa	-0.17		

*. Correlation is significant at the 0.05 level (2-tailed); **. Correlation is significant at the 0.01 level (2-tailed)

Table 2 shows the correlation between the score and the linguistic components of the students' writing. On the one hand, there are three components positively correlated with score: DESWC (with a Pearson correlational coefficient of.678**), WRDHYPn (.358**), and SYNNP (.230*). This means, the better English student writer can write greater number of words within a limited period of time, with more hypernymic nouns, and more noun phrases with modifiers. On the other hand, there are three components negatively correlated with score: WRDFAMc (-.367**), SYNSTRUTt (-.241*), and RDL2 (-.210*). This means, the worse English student writer tends to use more familiar content words and more similar sentence structures, which leads to lower readability of the writing.

TABLE III
PEARSON CORRELATION BETWEEN LINGUISTIC FEATURES

	DESSC	DESWC	DESSL	DESWLlt	DESWLltd	PCNARp	PCSYNp	PCCNCp	CRFAO1	CRFAOa	CRFCWO1	CRFANP1	CRFANPa
DESSC	1.00	.468**	-.543**	-0.02	0.07	0.17	.583**	-.413**	-0.17	-.248*	-0.16	0.00	-0.16
DESWC	.468**	1.00	0.03	-.225*	-0.09	0.11	-0.11	-0.08	0.11	0.06	-0.08	0.19	0.02
DESSL	-.543**	0.03	1.00	0.00	0.00	-.364**	-.596**	.338**	-0.05	-0.07	-0.08	-0.09	-0.04
DESWLlt	-0.02	-.225*	0.00	1.00	.569**	-.420**	0.16	0.17	-0.12	-0.01	-0.15	-.386**	-.297**
DESWLltd	0.07	-0.09	0.00	.569**	1.00	-.334**	0.06	-0.15	-0.18	-0.14	-.238*	-0.10	-0.13
PCNARp	0.17	0.11	-.364**	-.420**	-.334**	1.00	0.20	-0.16	.440**	.478**	.503**	.522**	.505**
PCSYNp	.583**	-0.11	-.596**	0.16	0.06	0.20	1.00	-0.20	-.285**	-.234*	-0.11	-0.05	-0.16
PCCNCp	-.413**	-0.08	.338**	0.17	-0.15	-0.16	-0.20	1.00	0.02	0.11	-0.03	-0.13	-0.07
CRFAO1	-0.17	0.11	-0.05	-0.12	-0.18	.440**	-.285**	0.02	1.00	.782**	.790**	.352**	.481**
CRFAOa	-.248*	0.06	-0.07	-0.01	-0.14	.478**	-.234*	0.11	.782**	1.00	.645**	.358**	.523**
CRFCWO1	-0.16	-0.08	-0.08	-0.15	-.238*	.503**	-0.11	-0.03	.790**	.645**	1.00	.329**	.489**
CRFANP1	0.00	0.19	-0.09	-.386**	-0.10	.522**	-0.05	-0.13	.352**	.358**	.329**	1.00	.789**
CRFANPa	-0.16	0.02	-0.04	-.297**	-0.13	.505**	-0.16	-0.07	.481**	.523**	.489**	.789**	1.00
LDTTRc	-0.17	-0.12	0.05	0.06	0.21	-.432**	-0.14	0.19	-.403**	-.455**	-.576**	-0.02	-0.04
LDTTRa	-.249*	-.382**	0.05	.299**	.228*	-.552**	0.00	.222*	-.386**	-.447**	-.525**	-.251*	-.290**
SMTEMP	.345**	0.15	-.861**	-0.03	0.00	.352**	.249*	-0.21	.210*	.243*	0.16	0.14	0.09
SYNNP	-0.15	0.16	0.09	0.03	0.06	-.308**	-0.17	0.16	-0.11	-0.02	-0.19	-0.13	-0.17
SYNSTRUTa	.479**	-0.15	-.515**	0.03	0.03	0.20	.527**	-.299**	0.07	0.00	0.14	0.14	0.06
SYNSTRUTt	.568**	-0.17	-.571**	0.09	0.07	.223*	.669**	-.358**	-0.08	-0.08	0.02	0.05	-0.03
WRDFAMc	0.02	-0.09	0.00	-0.16	-.282**	.327**	0.10	-0.01	0.11	0.13	.310**	-0.01	-0.09
WRDCNCc	-0.14	-0.01	0.03	0.11	-0.09	-0.02	0.02	.742**	-0.06	0.03	-0.04	-0.20	-0.15
WRDHYPn	-0.09	0.16	0.00	-0.06	-0.08	-0.13	-0.02	0.18	0.10	0.07	0.09	-0.08	-0.07
WRDHYPv	-0.16	-0.07	0.08	0.03	0.01	0.11	-0.08	.212*	0.12	0.18	0.12	0.08	0.15
RDFRE	.472**	0.14	-.774**	-.462**	-.420**	.549**	.393**	-0.20	0.12	0.08	0.17	.227*	.208*
RDFKGL	-.546**	-0.04	.959**	.208*	0.19	-.473**	-.542**	.297**	-0.09	-0.08	-0.13	-0.16	-0.12
RDL2	0.15	-0.08	-.288**	-.251*	-.291**	.497**	0.13	-.240*	.551**	.390**	.746**	.321**	.356**

(TO BE CONTINUED)

	LDTTRc	LDTTRa	SMTEMP	SYNNP	SYNSTRUTa	SYNSTRUTt	WRDFAMc	WRDCNCc	WRDHYPn	WRDHYPv	RDFRE	RDFKGL	RDL2
DESSC	-0.17	-.25*	.345**	-0.15	.479**	.568**	0.02	-0.14	-0.09	-0.16	.472**	-.546**	0.15
DESWC	-0.12	-.38**	0.15	0.16	-0.15	-0.17	-0.09	-0.01	0.16	-0.07	0.14	-0.04	-0.08
DESSL	0.05	0.05	-.86**	0.09	-.515**	-.571**	0.00	0.03	0.00	0.08	-.774**	.959**	-.288**
DESWLlt	0.06	.299**	-0.03	0.03	0.03	0.09	-0.16	0.11	-0.06	0.03	-.462**	.208*	-.251*
DESWLltd	0.21	.228*	0.00	0.06	0.03	0.07	-.28**	-0.09	-0.08	0.01	-.420**	0.19	-.291**
PCNARp	-.43**	-.55**	.352**	-.31**	0.20	.223*	.327**	-0.02	-0.13	0.11	.549**	-.473**	.497**
PCSYNp	-0.14	0.00	.249*	-0.17	.527**	.669**	0.10	0.02	-0.02	-0.08	.393**	-.542**	0.13
PCCNCp	0.19	.222*	-0.21	0.16	-.299**	-.358**	-0.01	.742**	0.18	.212*	-0.20	.297**	-.240*
CRFAO1	-.40**	-.39**	.210*	-0.11	0.07	-0.08	0.11	-0.06	0.10	0.12	0.12	-0.09	.551**
CRFAOa	-.46**	-.45**	.243*	-0.02	0.00	-0.08	0.13	0.03	0.07	0.18	0.08	-0.08	.390**
CRFCW01	-.576**	-.525**	0.16	-0.19	0.14	0.02	.310**	-0.04	0.09	0.12	0.17	-0.13	.746**
CRFANP1	-0.02	-.251*	0.14	-0.13	0.14	0.05	-0.01	-0.20	-0.08	0.08	.227*	-0.16	.321**
CRFANPa	-0.04	-.29**	0.09	-0.17	0.06	-0.03	-0.09	-0.15	-0.07	0.15	.208*	-0.12	.356**
LDTTRc	1.00	.789**	-0.05	0.17	-0.12	-0.15	-.58**	0.06	-0.07	0.02	-0.10	0.08	-.480**
LDTTRa	.789**	1.00	-0.12	0.10	-0.08	-0.09	-.33**	0.01	-0.06	-0.04	-.212*	0.13	-.411**
SMTEMP	-0.05	-0.12	1.00	0.02	.294**	.273**	-0.07	0.04	0.07	-0.02	.687**	-.834**	0.21
SYNNP	0.17	0.10	0.02	1.00	-0.12	-0.13	-0.17	0.19	.362**	-0.19	-0.12	0.11	-0.17
SYNSTRUTa	-0.12	-0.08	.294**	-0.12	1.00	.881**	0.13	-0.10	-0.03	-.215*	.362**	-.479**	.603**
SYNSTRUTt	-0.15	-0.09	.273**	-0.13	.881**	1.00	0.17	-0.09	-0.12	-.226*	.388**	-.526**	.438**
WRDFAMc	-.58**	-.33**	-0.07	-0.17	0.13	0.17	1.00	0.01	-0.07	-0.09	0.11	-0.05	.434**
WRDCNCc	0.06	0.01	0.04	0.19	-0.10	-0.09	0.01	1.00	.314**	0.16	0.06	-0.01	-0.19
WRDHYPn	-0.07	-0.06	0.07	.362**	-0.03	-0.12	-0.07	.314**	1.00	-0.10	0.00	0.00	0.01
WRDHYPv	0.02	-0.04	-0.02	-0.19	-.215*	-.226*	-0.09	0.16	-0.10	1.00	-0.05	0.07	-0.09
RDFRE	-0.10	-.212*	.687**	-0.12	.362**	.388**	0.11	0.06	0.00	-0.05	1.00	-.922**	.399**
RDFKGL	0.08	0.13	-.83**	0.11	-.479**	-.526**	-0.05	-0.01	0.00	0.07	-.922**	1.00	-.357**
RDL2	-.48**	-.41**	0.21	-0.17	.603**	.438**	.434**	-0.19	0.01	-0.09	.399**	-.357**	1.00

*****. Correlation is significant at the 0.05 level (2-tailed); ******. Correlation is significant at the 0.01 level (2-tailed)

Table 3 displays the pearson correlation between different linguistic features.

First, word length is positively correlated with LDTTRa (.299**) and CRFANPa (.208*), while it is negatively correlated with PCNARp(-.420**), RDFRE(-.462**),and RDFKGLRDL2(-.251*). This means, the student using longer words can use more pronouns and various lexemes, which makes the writing less easy to understand.

Second, sentence length is positively correlated with PCCNCp(.338**) and RDFKGL(.959**), while it is negatively correlated with PCNARp(-.364**), PCSYNp(-.596**), SMTEMP(-.861**), SYNSTRUTa(-.515**), SYNSTRUTt(-.571**), RDFRE(-.774**), and RDL2(-.288**). This indicates, the student using longer sentences tend to use more concrete words can get a higher score, but such a writing seems to be less narrative, more syntactically complex, less temorally cohesive, and more difficult to understand.

Third, Syntactic simplicity is positively correlated with SMTEMP(.249*), SYNSTRUTt(.669**), and SYNSTRUTa (.527**), and RDFRE (.393**), while negatively correlated with CRFAO1 (-.285**), CRFAOa (-.234*), and RDFKGL (-.542**). This suggests, the student who is in favor of simple sentences tends to prefer similar sentence structures and repetitive tenses and aspects; while doing so causes the writing' argument to be less overlapping and the writing to be of lower quality.

Fourth, word concreteness is positively correlated with CRFAO1(.440**), CRFAOa(.478**), CRFCWO1(.503**), CRFANP1(.522**), CRFANPa(.505**), SMTEMP(.352**), and SYNSTRUTt(.669**); while negatively correlated with DESSL(-.364**), LDTTRc(-.432**), LDTTRa(-.552**), and SYNNP(-.308**). This shows, on one hand, the student favors concrete words also prefer to use pronouns, tenses and aspects repetitively, which contributes to an overlapping argument. One the other hand, such a student seldom uses long sentences, diverse lexemes, and complex noun phrase structures.

Fifth, Argument overlap is positively correlated with PCNARp(.478**), CRFAO1(.782**), CRFCWO1(.645**), CRFANP1(.358**), and CRFANPa(.523**), while negatively correlated with PCSYNp(-.234*), LDTTRc(-.455**), and LDTTRa(-.447**). This means, the student prefers to use overlapping pronouns and content words can make the theme more focused. However, such a student tends to use less simple sentence structures, and less diverse lexemes.

Sixth, content word overlap is positively correlated with PCNARp(.503**), CRFAO1(.790**), CRFAOa(.645**), CRFANP1(.329**), CRFANPa(.489**), and WRDFAMc(.310**), while negatively correlated with DESWLltd(-.238*), LDTTRc(-.576**), and LDTTRa(-.525**). This indicates, those who prefer content words also like pronouns and concrete words better; neverthless, they seldom use diverse lexemes.

Seventh, anaphor overlap is positively correlated with PCNARp(.505**), CRFAO1(.481**), CRFAOa(.523**), CRFCWO1(.489**), CRFANP1(.789**), RDL2(.356**), and RDFRE(.208*), while negatively correlated with DESWLlt(-.297**), and LDTTRa(-.290**). This suggests, those who prefer pronouns can make the writings easier to read, but with less diverse lexemes.

Eighth, Lexical diversity are all negatively correlated with PCNARp(-.43**), CRFAO1(-.40**), CRFAOa(-.46**), CRFCWO1(-.576**), WRDFAMc(-.58**), and RDL2(-.48**). This shows, those who prefer diverse lexemes seldom use overlapping pronouns, and familiar content words, which makes the writings more difficult to understand.

Ninth, temporal cohesion is positively correlated with PCCNCp (.222*), and LDTTRc(.789**), while negatively correlated with PCNARp(-.55**), CRFAO1(-.39**), CRFAOa(-.45**), CRFCWO1(-.525**), CRFANP1(-.251*), CRFANPa(-.29**), WRDFAMc(-.33**), and RDL2(-.41**). This means, those who prefer the same tenses and aspects use diverse content words more often. However, they use pronouns and familiar content words less often, which makes the writings more difficult to understand.

Finally, sentence syntax similarity is positively correlated with PCNARp(.223*), PCSYNp(.669**), SMTEMP (.273**), and RDL2 (438**), while negatively correlated with WRDHYPv (-.226*), RDFKGL (-.526**), and PCCNCp (-.358**). This indicates, those who prefer the same sentence structures use the same tenses and apects more often, which makes the writings easier to understand; whereas, they use hypernymous nouns and concrete words less often. In addition, those who prefer to use hypernymous verbs can make the writing more diverse in sentence structure, SYNSTRUTt (-.226*).

IV. Discussions and Conclusions

From the above Co-Metrix analyses, some findings are concluded as follows:

First, the overall writing professiony of the mechanic vocational students is far from satisfactory. Since the writing proficiency of the students in CET seems to be lowering in recent years, with an average score of 40 out of 100 for each student's writing (li, 2012), that of vocational students is no exception. And that of science majors of vocational students is even worse, for in general, the English professiony of arts students is better than their science counterparts'. Therefore, it is a long way for the vocational college English teachers as for how to improve their students' English. Maybe, how to enhance their motivation to study is the primary issue.

Second, the students prefer to use more concrete and familiar content words, which leads to the easiness of the writings. As is often the case, some vocational college students tend to use some daily conversational words in their expository writings, which is not corresponding to the style of writing. As to this point, they should be taught with more knowledge on stylistics in class. In addition, it is essential to enlarge their vocabulary by urging them to do more readings.

Third, they seldom use pronouns and parallel structures in their writings, thus the writing is not central to the main idea and not coherent enough. In light of this aspect, the Chinese students may be influenced by their mother tongue, which is one negative transfer of mother language. As we know, Chinese is a language focusing more on meaning rather than on form, where it prefers repetions of nouns rather than pronouns. Therefore, coherence of a Chinese writing does not wholy depend upon its forms, but on its inner meaning. Hence, the Chinese students should be taught more knowledge on English theory of coherence.

Coherence is the quality of meaning unity and purpose perceived in discourse. It is not a property of linguistics forms in the text and their denotations, but of these forms and meanings interpreted by a receiver through knowledge and reasoning. As such, coherence is not an absolute quality of a text, but always relative to a particular receiver and context. Coherence is usually concerned with the links inferred between sentences or utterances. It is often contrasted with cohesion, which is the linguistic realization of such links (Halliday and Hasan, 1976).

Fourth, the student who is in favor of simple sentences tends to prefer similar sentence structures and repetitive tenses and aspects; while doing so causes the writing' argument to be less overlapping and the writing to be of lower quality. This may result from two points. One is also the negative transfer of Chinese, which lacks sheer restrictions of some grammatical categories such as tense and aspect; the other is due to the poorness of their English, in which they can only write simple sentences rather than more complex ones.

Fifth, the student prefers to use overlapping pronouns and content words can make the theme more focused. However, such a student tends to use less simple sentence structures, and less diverse lexemes. This seems to be rather conflicting, for generally, the top English writer should be skillful in using pronouns and complex sentence structures, but also in using diverse lexemes. Maybe, the answer lies in the command of too small a glossary of the students. Try to enlarge their vocabulary by pushing them to do more reading.

Finally, In addition, those students who prefer to use hypernymous verbs can make the writing more diverse in sentence structure. This is a good sign of bettering writing, for many Chinese students often feel perplexed at using English verbs properly. Maybe, English verbs are far more intricate than their Chinese counterparts in their changing forms.

Acknowledgement

This paper is indebted to Prof. McNamara Crossley and the Graduate Student Nia Dowell in University of Memphis, for their help to process all the data by Coh-Metric 3.0.

References

[1] Chen, X., Du. G. (2012), Multidimensional analysis on the writing by non-English major postgraduates. *Journal of Northeast Normal University*, 259 (5): 137-139.

[2] Duncan. (2008). Ph.D. Thesis. A computational linguistic analysis of biomedical abstracts, the University of Memphis.

[3] Graesser, A. C., Dowell, N., & Moldovan, C. (2011). A computers understanding of literature. *Scientific Study of Literature*, 1(1), 24–33. doi:10.1075/ssol.1.1.03gra
[4] Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher,* 40(5), 223–234.
[5] Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-metrix: analysis of text on cohesion and language. *Behavior research methods, instruments, & computers: a journal of the Psychonomic Society*, Inc, 36(2), 193–202.
[6] Halliday, M.A.K., Hasan, R. (1976). Cohesion in English. London: Longman.
[7] He, L. (2013). MA Thesis, A Comparative study on text reading complexity of TEM-8 (Test for English Majors, Grade 8) and IELTS. Southwestern University of China.
[8] Li, X., L. (2012). Using Coh-Metrix to Automatically Access Discourse Coherence in EFL Learners' Writing: A Quantitative Analysis, *Overseas English*, 9: 117-121.
[9] Li, Y. J. (2012). Negative echoic effects of the writing topic of college English test band 4 - some afterthought of a CET4 writing rater. *College English Academic Edition*, 99, 273-277.
[10] McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (in press). Automated evaluation of text and discourse withCoh-Metrix. Cambridge, M.A.: Cambridge University Press.
[11] McNamara, D.S., Louwerse, M.M., Cai, Z., & Graesser, A. (2013). Coh-Metrix version 3.0. Retrieved [01/14/2014], from http//:cohmetrix.com.
[12] Qin, Z. & Gu, Q. (2011). Writing topic familiarity and cohesion means of the domestic learners' written discourse- A comparative study on automatic measurement method, *Journal of Xi'an international studies university*, 19(1): 95-99.

Ronggen Zhang, Associate professor, holds an MA degree in ESP from University of Shanghai for Science and Technology in China. His research interest lies in corpus linguistics and investigation of language teaching strategies. He is now teaching in Shanghai Publishing and Printing College. He has recently published in JCIT (Journal of Convergence Information Technology).