

Computer-assisted Lexical Analysis of ESL Writing Based on Cognitive Model

Qin Wang

Shaanxi Normal University, Xi'an, China

Abstract: *We present a lexical analysis of ESL writing samples from a cognitive perspective. The small writing database adopted in the research collects 528 writing samples of 88 students for one semester. Then, the lexical features of ESL writing are analyzed for word co-occurrence, extraction, repetition, and connector. Based on the lexical analysis result, we discuss the lexical features from cognitive understanding, probing into ESL learners' organization in writing, vocabulary acquisition, general thinking pattern, and logical thinking. The college students with a vocabulary of over 3,000 are capable of transferring their points into English writing. At the same time, their vocabulary acquisition pattern negatively influences their word extraction, and their lexical feature mirrors deficiency in logical thinking. We provide a short reflection on the limitations of the research.*

Keywords: *lexical analysis, cognitive model, word co-occurrence, word extraction, connector*

1. Introduction

Writing is regarded as an indispensable indicator in ESL learners' language proficiency and evaluation. Contrary to this assumption, teaching and learning ESL writing has been a tough and weak issue for a long time [1]. The possible reasons come from both the educators and the learners. The educators are struggling with oversize classes, insufficient time allocation, and evaluating difficulties, while the students tend to emphasize reading and ignore writing [2].

Many researchers devote their time and efforts to studying teaching and learning ESL writing. They have published many papers on methodology, influencing factors, testing, and evaluation. Among which papers concerning methodology accounts for the largest proportion [3]. ESL writing research, to some extent, ignores the learners themselves and the learners' thinking process.

The study is triggered by the researcher's observation in writing classes. The research focuses on learners by analyzing writing products from the point of understanding thinking in writing. The findings and reflections on ESL writers' organization, vocabulary acquisition, and thinking patterns are investigated in this study.

2. Research on ESL Writing

The proliferation of writing research probes into ESL writing. Their studies fall into three categories: product-oriented, reader-oriented, and process-oriented.

The product-oriented view is inherited from structuralism and behaviorism. It regards oral as the primary medium while writing a secondary concern which reinforces oral habits [4]. This approach encourages instructing learners to imitate, copy and transform models from textbooks and educators [5-6]. Due to its focus on the surface structure of writing sentences and discourse level, emphases are put on language forms, like grammar, syntax, cohesion, and writing samples. A large number of studies try to explain how ESL writing is influenced by those necessary writing skills [7-8]. These researches make comparisons and clarify misunderstandings in grammar, organization, rhetoric skills, etc.

The reader-oriented approach includes readers' expectations and social context in teaching writing [9]. This approach focuses on the given purpose and functional communication of writing. It has been criticized for being too prescriptive. However, in recent years, it has strong move towards ESP(English for Special Purposes) and EOP(English for Occupational Purposes).

The above notions ignore thinking while it is commonly accepted that thinking has a great influence on writing, and writing shows the way of thinking. Many researchers on ESL writing emphasize on

language skills while ignoring thinking [2]. Writing is undoubtedly representing what we are thinking and how we organize our thinking. We have to examine the intangible process and understand how the product came into being [10].

Here, comes the process-oriented approach which emerged with L1 writers [11]. The approach cares about what the writer does during the writing process [5-6]. It is most famous for its cognitive model of the writing process [12-13]. By adopting psychological thought, methods like think-aloud are deployed. The research of Kobayashi and Rinnert (1992) was based on the students' estimation of how much L1 thinking they have in writing. They find that the students of lower L2 proficiency rely more on L1 thinking [14]. By considering thinking in writing research, the researchers and educators are probing into the author's writing processes and making writing skills of skilled writers explicit and replicable by unskilled writers.

For ESL writing research in China, CNKI data reveals visible characteristics of writing research after 2000. Domestic researches focus on feedback, strategy, motivation, and ability. Other staged keywords are "transfer" and "corpus" from 2002-2007; "chunks" and "feedback" from 2010-2015, and "teaching modes" from 2016-2019 [15].

The correlation between L1 thinking and English/Western thinking and its influence on English writing are investigated as well. Many researchers focus on the existence and general influence of L1 thinking in ESL writing, especially for students of lower language proficiency. They analyze the difference between Chinese thinking and Western thinking as logic and intuition, subjectivity and objectivity, linear thinking, and spiral thinking. They discuss the related influence on word choices, sentence structure, and text organization [16-20]. Lay (1982) analyzes four Chinese students' English writing and found positive effects of native thinking in content, structure, and details. Friedlander (1991) found native thinking could help students structure Chinese-related topics. Guo and Wen (1998) found some negative effects of L1 thinking in writing [16-20].

Research has adopted psychological methodology. They introduce and evaluate major tests of critical thinking skills [21]. This research takes thinking as a complicated psychological issue that cannot be seen directly. Thus, the well-designed variable-controlled activity and necessary facilities are adopted to show the thinking process in second language writing. Wang and Wen adopt a thinking-aloud method for observation and interviews to reveal the second language writing process is bilingual. The amount of L1 thinking was measured and a correlation was made between L1 thinking and L2 writing quality. Both negative and positive correlation was found between L1 use and L2 writing. Their research presented that L1 thinking is not just an issue of "positive transfer" or "negative transfer". They also show the proportion of L1 thinking decreased with the writer's L2 development [22-23].

Other researches focus on the output and text analysis of English writing. By studying paragraph and essay organization, vocabulary, and connecting, it is summarized that the main influencing factors of second language writing are organization and vocabulary [24]. Wen and Liu reported on features and problems of abstract thinking in second language writing [2]. After collecting writing samples of argumentation from four grades, they evaluated the samples on relevance, explicitness, coherence, and sufficiency. The result shows the evolving features and problems in presupposed four thinking stages. Liu checks how the students' writing quality is affected by exterior factors, like writing task, writing environment, and writing product [25].

For Chinese learners' ESL writing and thinking, many results come from the general comparison and contrasts of thinking patterns. They care about the correlations of L2 proficiency and writing quality, trying to find replicable writing processes and writing skills. There is research on imitation or application of L1 writing and achievements, especially for the writing process. The ignored points are as follows: L1 writing and thinking are different from ESL writing. The correlation between English proficiency and ESL writing quality is still general and arguable. If we assume there is a great influence of L1 thinking on ESL writing, is there evidence to prove the detailed aspects of the influence, positive or negative? If there is a certain correlation between English proficiency and ESL writing, how about students of a certain level? What will be the influencing factors? If the general English writing pattern mentioned in previous literature has been introduced to the students, will that greatly improve their ESL writing quality?

This study is based on teaching experience and reflection on these questions. Based on the basic understanding of cognition and ESL writing, this study focus on accessible writing database and its cognitive analysis. All the writing samples applied in the study are collected from the writing classes. The lexical analysis of the writing samples may answer to the above questions.

3. Methodology and Research Process

Previous literature has proved that Chinese writing is different from ESL writing in the organization, sentence structure, and so on. In order to probe into the more detailed aspects of Chinese thinking on ESL writing, the researcher intentionally introduced the text organization of English writing and asked the students to finish several related in-class and after-class assignments. The assignments showed that they have a clear understanding of the general structure and their essays have well-organized point-supporting structure as required by English writing.

For the English proficiency test, the students of the two classes were all first-year students in college. They all finished middle-school and high-school English learning, having a vocabulary of around 3,000 (as required by High School English Curriculum Standard 2017).

From October 2020 to June 2021, three classes' writing assignments were collected 6 times for 528 writing samples. Each class has around 30 students with similar English backgrounds and English proficiency. All the writing assignments were chosen from CET-4 writing tasks. CET-4 is an authoritative English test in China for evaluating college students' English proficiency. All the students were familiar with CET-4 writing pattern. Each assignment has a word requirement of 120–150 words, which is the same as CET-4 requirement. All the writings were accomplished in class with a time limit of 30 min so that the students have no possibility of partner consulting or reference checking. Since the research does not involve spelling, all the spellings were auto-corrected by computer and rechecked by the researcher herself. Then, all the after-checking writing samples were transformed into required texts and stored as a small-scale writing database.

KH coding was used for lexical analysis. Though the text mining tools, Citespace and KH Coder are both for visualizing and analyzing trends and patterns, KH Coder is used in the research due to its accessibility and the need for processing Chinese. KH Coder is designed to achieve data from quantitative analysis of the text by Higuchi Koichi. Extracted words and clusters from the database were statistically analyzed to obtain a visual picture. The objective exploration of the data was performed to avoid subjectivity. For most English educators, it could be easily applied without unreachable professional requirements.

4. Findings and Discussion

All the collected writing samples have a basic English writing structure of point and support, which is required by college English learning and the students are familiar with it. When all the writing samples were stored in the database, the research uses KH Coder to make a lexical analysis. The following are the findings.

4.1. Organization and Point Introduction

In order to compare the students' Chinese thinking and English thinking in organizing and introducing points, the students were required to write two essays on the same topic "why I study", one in English, one in Chinese. The Chinese essay was assigned two months after the English essay to minimize the influence of English essays on Chinese essays. The two essays were both finished in class independently without any references. The time limit was 30 minutes for each.

Figure 1 shows the word co-occurrence of Chinese samples. All the English words in Figure 1 are translated from analysis result of Chinese writing samples. Figure 1 reflects that the Chinese writing samples have two main sections: the Blue one and the Orange one. The blue one has the central word "person" and the orange one has the central word "society". The word "person" is surrounded by person related words like dream, progress, development. The word "society" is surrounded by society-related words like knowledge, study, choice. Figure 1 shows the identical pattern to the personal observation of the students' writing performance. The students' writing samples have two general points: society and person. There are subpoints surrounding the general points.

The researcher's observation shows that the students' English writing has a similar organization. The English samples have two general aspects for "why": one is society, the other is person. However, English writing samples use a large number of pronouns like "we" and "I". The word "be" also has a high frequency. Due to this, the visual picture of word-occurrence does not show the organization like what we have in Chinese samples. As shown in Fig 2., the Pink section has the central word "we" and "be", which are surrounded by words like "I", "my", "study", "not", and "have". The students' word preference

are followed by job and knowledge clusters, life clusters, and future clusters. The reasons for study mainly fall in two categories: society and person, which is the same as in Chinese writing.

From the above analysis, it is clear that in Chinese writing and English writing, the student used similar organization and introduced the same or similar points. The keywords in organizing the essay and introducing points were the topic-related words like social development, society, job, future, and life. Their L1 thinking is directly transferred into L2 writing. The analysis shows that college students with a vocabulary of over 3,000 do not have difficulty in transferring their Chinese thinking into English writing in terms of organizing and introducing points. The interview of the students reveals the same. When introducing their points, they feel that they generally have enough words to use. For the students at this vocabulary level, vocabulary is not a barrier in introducing points in ESL writing. However, their word preference has a great influence on their actual writing performance. In addition, introducing points and organizing essays in ESL writing may not be an issue of positive or negative, it is more like an issue of transferring.

4.2. Word extraction

Table 3: WORDS in Brainstorming and Writing

	Words in Brainstorming	Words in Writing Samples
adj.	ancient, amazing, bright, classic, colourful, cultural, difficult, eastern, famous, fantastic, glorious, great, interesting, long, magical, modern, open, profound, special, traditional, various, lasting, intelligent, splendid, graceful, peaceful, integrative, influential, proud, inspiring, encouraging, modest, changing	traditional274, Chinese216, long42, many42, ancient38, new32, more30, different28, good20, important20, better12, colorful10
v.	achieve, celebrate, dance, gather, hold, hope, learn, memorize, worship, protect, change, spread, inherit, introduce, insist, connect, conflict, love, integrate, explore, stabilize, admire, build, learn, influence	be446, have186, make46, know32, develop24, inherit24, take22, protect20, Promote18, use18, learn16, eat14, spread14, pass12, show12,
n.	bamboo, Confucius, dragon, dumpling, great wall, herb, hero, Kungfu, lantern, opera, paper-cutting, poem, tradition, variety, tea, china, couplet, classics, red-envelope, dumpling, calligraphy, painting, silk, influence	culture279, people108, history102, life57, day39, development33, influence27, country24, poetry16, spring festival16, family14, calligraphy12, future12, innovation12

Since all the students have finished high school English learning and passed College Entrance Examination (CEE), they have the required vocabulary of over 3,000. The writing sample analysis shows how they extract and use their words.

Before each writing practice, the researcher asked the students to brainstorm without references. They were required to write down the adjectives, nouns, or verbs that they know about the topic. Then, the students finish their writing in class. When the writing samples were collected, the lexical features were analyzed. Table 3 shows the lexical differences between brainstorming and actual writing. For the brainstorming, the students write down many different words about the topic, which shows they have a large vocabulary and they do not have much difficulty in extracting words. Their words were related to the topic reflect lexical diversity. Their vocabulary and lexical diversity well represented this activity. On the contrary, Words in Writing Samples show they are not using the words in the brainstorming part. They tend to use certain general words and other topic-related general words in all the writing assignments. The words they use in writing reveal the lack of diversity and give us some understanding of word extraction in writing.

Table 3 reveals word extraction and use in writing. The students have a moderate vocabulary of over 3,000 and have the possibility of word extraction in a brainstorming activity. However, when they use the words in writing, that is different. They only use commonly used general words and topic-related general words. The possible explanation comes from the students' vocabulary acquisition. The interview of the students shows that most students pick up words by mechanically memorizing single words and

lexical bundles. Most of them do not have the experience of picking up new words in context. The lack of context understanding may lead to what is happening here in writing practice. When they store those words in their memory by Chinese meaning, each word is a single and isolated item. They could easily recall and extract many words by thinking about the topic or topic-related words in their native language because those are the one-to-one pairs they have in their memory. However, they cannot use their vocabulary to clearly explain something, support their points, and talk about more specific and accurate details in writing. Thus, they tend to use general and easy words in Table 3. In addition, the analyzed writing samples of one topic have 293 paragraphs on average. Those paragraphs have around 900 sentences/16,000 words, among which non-repeated words are around 2,000. It means several words are repeatedly used in the students' samples. Table 4 reveals the repetition. In a writing sample, the word "study" is repeatedly used in each sentence. In addition, the lexical bundles come from their high school textbook or native expressions. The structures "S. + should/have to/must/can + study (V.)" and "study (N.) makes me/us + V." are most common in the sample database. Extracted from the database are also bundles like "deep study", "study knowledge" and "study hard", which are influenced by students' native language.

Table 4: Repetition

Study	Study makes me grow up to be a better self. Through studying poems, I gradually find happiness in daily life. After studying maths, I feel my thinking becomes more flexible. Study brings me both material and spiritual treasure.
--------------	---

4.3. Repetition of Topic-related Words and General Words

Table 5: Word Frequency

	Essay 1	Essay 2
Adj.	more138, only 69, better60, good54, so48, many45, hard36, also33, important33, different27, old27, just24, always18	traditional274, Chinese216, long42, many42, ancient38, new32, more30, different28, good20, important20, better12, colorful10
v.	study301, learn216, have159, do102, make75, get66, become48, improve46, want46, know39, keep36, need36, say36, give31	be446, have186, make46, know32, develop24, inherit24, take22, protect20, promote18, use18, learn16, eat14, spread14, pass12, show12,
	Essay 1	Essay 2
n.	study183, life108, knowledge106, society72, way54, reason55, people49, time49, school192, world44, job39, thing38, future37, development34, person27, student27, horizon25	culture279, people108, history102, life57, day39, development33, influence27, country24, poetry16, spring festival16, family14, calligraphy12, future12, innovation12

Table 5 shows the word frequency of writing samples. The students tend repeatedly to use topic-related words in their ESL writing. In Essay 1, which is related to study, "study" has the frequency of 484 in 88 writing samples, both as verb and noun. The word "learn", which has a similar meaning as "study", also appears over 210 times. In Essay 2, which is about culture, the word "culture" has the highest frequency as a noun. The word "culture" appears 279 times in 88 writing samples, 3 times for each. This repetition of topic-related words may also be related to the students' vocabulary acquisition.

The students have a vocabulary of over 3,000. The vocabulary includes both general and specific words. Table 5 shows that they are repeatedly using the general words they have. In Essay 1, simple words, like "study" and "learn", are repeatedly used in the students' writing samples. In the two essays, general verbs like "have", "do", "make", "get", "become" could be seen in many sentences. If we take adjectives as an example, "more", "good" and "important" are common in each essay. They are also commonly used in students' various writing samples. The word frequency in Table 5 reflects the students' way of thinking in ESL writing: general instead of specific. The researcher's observation and interview show that the students tend to think about any topic generally in ESL writing. They understand the general writing pattern of topic-and-supporting, but they tend to avoid the fact-to-abstraction process. When they have a topic in writing, they think around the topic generally and have general points in mind. They do not have difficulty in forming a general organization. When the organization is done, the supporting part in their writing is usually from general aspects. That is why they repeatedly use topic-related words, and

there are so many general words repeated in their writing.

4.4. Connectors

Table 6 shows connectors in English writing samples. The most common connectors in Table 6 include the category of time-and-order (first). The possible reason is the requirement of CET-4 writing and the emphasis on orderly linking in college English teaching. Since the students are not clear about logical relationships, they tend to use simple linking - “first, second, third” to list reasons, suggestions, and so on. Table 6 also reveals the frequency and ranking of connectors, which represents a difference in connector use. Considering the small size of the self-built database, connectors, like “and (900)”, tends to be overused. In contrast to this overuse, Table 6 also illustrates the missing connectors. For example, “in fact” and “for instance” both have zero frequency. In Table 6., the connector “so” is noticeable for its high frequency (171), which is not appropriate for academic writing. This phenomenon may, to some extent, mirror the students’ misunderstanding of connectors’ contextual connotations. In addition, the most widely used connectors in Table 6 are simple connectors, like “and, but”, account for the largest proportion of connector uses in college English writing. By checking the lexical feature of connectors, there is one thing worth noting. The students are using connectors not because they understand the logical relations in their writing. The observation and interview show that they do not have clear logical thinking and understanding. That is why they overuse connectors. They are using first, second, third because the CET-4 writing samples are listing factors/reasons like that. They are using connectors like “because” just for the Chinese cause-and-effect word translation, which does not show logical cause-and-effect relation in English writing. The students’ misunderstanding of logic relations leads to overuse and repetition of simple connectors.

Table 6: Connectors

Connector	Frequency	Connector	Frequency
and	900	at least	0
but	333	in fact	0
or	18	in addition	18
for example	9	however	72
because	81	for instance	0
therefore	27	first(second,third)	63
so	171	moreover	3
actually	0	furthermore	0

5. Conclusion

The research discusses cognitive understanding in ESL writing by analyzing lexical features of students’ writing samples. By doing this, we have a different perspective on understanding ESL writing. Writing class has cognition-related activities, which leads to students’ better performance and more achievements. All the data and samples in the research are from students’ writing samples collected by the researcher herself in the 2020 Fall Semester.

The organization and point introduction show that the student has a similar organization in their English writing and Chinese writing. They also introduce similar points. With a vocabulary of over 3,000, college students are capable of transferring Chinese points into English writing. The research into this transfer process is probably more valuable than the perspective of positive or negative.

The Word Extraction illustrates that the students are capable of extracting single words related to the writing topic while they have great difficulty in using them in the actual writing. On the contrary, they tend to use simple and common words repeatedly. The lexical bundles reflect the negative influence of ESL learners’ native language. This may result from their vocabulary acquisition. The lexical analysis also reflects the repetition of topic-related words, which probably comes from the students’ general thinking patterns. They tend to ignore the process of facts to abstraction. From the point of hierarchy, they tend to use general words.

The connector analysis mirrors the student's understanding of connectors’ contextual connotations and logical relations. The analysis illustrates that the students do not have clear logical thinking in ESL writing. They use connectors simply because connector use is adopted in writing evaluation and it is recommended in writing samples. That is why overuse or misuse of connectors are common in their

writing samples.

There are the following limitations. The research is originally triggered by the researcher's observation and understanding in teaching, which still lacks theoretical and systematic understanding. The analysis is only based on the writing samples of one semester, which still needs further experiments and discussions in the future.

Acknowledgements

Shaanxi 2022 Education and Science 14th Five-year Plan Program(SGH22Y1261).

Fundamental Research Funds for Central Universities(Shaanxi Normal University2022JYYB15).

References

- [1] Han Han, Hong Wang, "The influence of of Ch-En thinking and speaking on college English writing", *Journal of Northwest University(Social Science Edition)*, Vol 25, pp 111-116, August, 1995.
- [2] Qiufang Wen, Runqing Liu, "An exploratory on features in English majors' abstract thinking in English argumentative compositions", *Journal of Foreign Languages*, vol 162, pp 49-58, April, 2006.
- [3] Zhixue Li, Shaoshan Li, "Reflection on domestic research on English writing", *Foreign Language World*, vol 98, pp 55-60, 2003.
- [4] Richards, J. C., & Rodgers, T. S. *Approaches and methods in language teaching: a description and analysis*. Cambridge: Cambridge University Press, 1986.
- [5] Nunan, D. *Second language teaching and learning*. Boston, Mass: Heinle & Heinle Publishers, 1999.
- [6] Tribble, C. *Writing*. Oxford: Oxford University Press, 1996.
- [7] Jinying Chen, "Analysis of English writing teaching", *Foreign Language Teaching and Research*, vol 98, pp 12-18, April, 1994.
- [8] Jin Wu & Zaixin Zhang, "Research on English writing teaching", *Foreign Language Teaching and Research*, vol 32, pp 213-218, May, 2000.
- [9] Hyland, K. *Teaching and researchign writing*. London: Longman. 2002.
- [10] Hairston, Maxine. "The winds of change: Thomas Kuhn and the revolution in the teaching of writing", *College Composition and Communication*, vol 33, pp 76-88, 1982.
- [11] Grabe, W. and Kaplan, R.B. *Theory and Practice of Writing: An Applied Linguistic Perspective*. Longman, New York.1996.
- [12] Hayes, J. R., & Flower, L. *Identifying the Organization of Writing Processes*. In L. W. Gregg, & E. R. Steinberg (Eds.), "Cognitive Processes in Writing: An Interdisciplinary Approach", pp. 3-30, Hillsdale, NJ: Lawrence Erlbaum.1980.
- [13] Bereiter, C., & Scardamalia, M. *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.1987.
- [14] Kobayashi Hiroe, & Rennert Carol, "Factors affecting composition evaluation in an EFL context: Cultural rhetorical pattern and readers' background", *Language Learning*. Vol 46, pp 397-437, 1996.
- [15] Wenjun Lu, "Englsih writing research in China(2000-2019): a scientometric analysis with Citespace", *Journal of Southwest Jiaotong University,(Social Sciences)*, vol 22, pp 65-73, 2021.
- [16] Lay N, "Composing processes of adult ESL learners :a case study", *TESOL Quarterly*, vol 16, pp 406-407, 1982.
- [17] Friedlander A, "Composing in English effects of a first langauge on writing in English as a second language, Kroll B., *Second Language Writing Research Insights for the Classroom*. Cambridge: CUP, 1991, pp109-125.
- [18] Guo Chunjie, Liu Fang, "A dynamic research into L1 influence on L2 writing", *Modern Foreign Languages*,vol 78, pp 30-38, 1997.
- [19] Guo Chunjie, Wen Qiufang, "The relationship between thinking in L1 and L2 writing ability:a study of the process of English picture composition by senior middle school students". *Modern Foreign Languages*, vol 82, pp 44-56, 1998.
- [20] Guohua Tu, Yuanling Wang, "On thinking modes and English academic writing", *Journal of Central South University(Social Science)*, vol 19, pp 238-243, 2013.
- [21] Jianqing Wang, Qiufang Wen, "Introduction of Critical Thinking Instruments from Abroad and Some Inspiration", *Journal of Jiangsu Teachers University of Technology*, vol 17, pp 38-42, July, 2011.
- [22] Wenyu Wang, Qiufang Wen, "An investigation into L1 use in the L2 writing process of tertiary-level English learners in China", *Journal of PLA University of Foreign Languages*, vol 25, pp 64-67, July, 2002.
- [23] Wenyu Wang, Qiufang Wen, "Correlation between L1 thinking and rating of L2 writing", *Foreign*

Languages and Their Teaching, vol 163. pp 17-20, October, 2002.

[24] Songtao Li, "The influence of Chinese and Western thinking mode on textual organization of college English writing", *Foreign Language Education*, vol 26, pp 52-56, March, 2005.

[25] Hongwei Liu, "Out factors and influence on English writing process and quality", *Journal of Tianjin Foreign Studies Univeristy*, vol 12, pp 75-78. 2005.

[26] Qingbin Chen, "Reconstruction of the teaching model of college English writing in the big data era", *Foreign Language Research*, vol 190, pp 129-132, 2016.