A Case Study on Errors in Engineering College Students' English Writings Based on the Correcting Network Corpus

XIE Mengzhou[1] and LIU Meiyan[2]

[1]School of Foreign Studies, Northwestern Polytechnical University,
City: Xi'an, China, Postal Code: 710129
E-mail: xie_mengzhou@163.com

[2]School of Foreign Studies, Northwestern Polytechnical University,
City: Xi'an, China, Postal Code: 710129
E-mail: myliu@nwpu.edu.cn

Corresponding author:
XIE Mengzhou
School of Foreign Studies, Northwestern Polytechnical University,
City: Xi'an, China, Postal Code: 710129
E-mail: xie_mengzhou@163.com

Published: 21 July 2021

Abstract

As English learners, Chinese college students inevitably make errors in the English writings. Based on Coder's Error Analysis Theory, this paper focuses on the different errors in 88 English writing texts of science and engineering students in grade one with the title of "Reduce Waste on Campus". The results show that there are six types of errors that students are easily to make i.e. the collocational errors, misspellings, the errors of subject-verb agreement, the errors of not capitalized initials, the errors of sentence structure, the misuse of articles, the errors of sentence components and the misuse of nouns. After reorganizing such errors based on the revised Carl James's classification of errors, it is found that students are easily to make errors on the text level, especially in lexis compared with the grammatical errors. Among them, the main influencing factors of making errors are interlingual factors and intralingual factors. Meanwhile, some feasible suggestions are put forward for English writing teaching.

Keywords: Correcting Network, engineering college students, English writings, error analysis

1. Introduction

As one of the language output skills, writing is an important part of the comprehensive ability training, and also a weak link in many students' English learning process. With the development of the corpus linguistics, it offers a new way for scholars to conduct the research on language errors. Correcting network, designed as a new type of automatic online marking system of English writing with the technological support of corpus and cloud computing, is helpful for teachers and English learners in colleges to analyze and correct errors in English writing. With this function, the Correcting Network is a very helpful tool to do the error analysis.

Error Analysis Theory was put forward by Pit Coder in 1960s. It has contributed a lot to the research on English writing teaching. Liu (2008) has discussed the intralingual and interlingual errors in a Chinese non-English major student's writing. Zheng and Li (2008) make an analysis of verb errors which are regarded as the most frequently appearing errors students make in TEM writings. Moreover, some scholars are mainly focused on the influence of language transfer. Ge (2003) has summarized the phenomena of language transfer appeared in the English writings of colleges students. Besides, corpus is widely used in the study of error analysis. Zhang (2020) and Sun (2020) both conduct the research based on the self-built corpus. And with the help of the correcting network, it facilitates the process of identifying errors. Wand and Xu (2015) take use of the function of the correcting network to analyze the different types of errors in 127 college students English writings. Cao and Xiang (2020) particularly pay attention to the collocational errors marked by the correcting network. It can be seen that by conducting this kind of research, new cases can be added to the relevant research of the error analysis theory especially with the help of the corpus, and more practical suggestions can be put forward to benefit the English writing teaching.

Therefore, based on the error analysis theory, according to Carl James' classification of errors, this study tries to conduct quantitative and qualitative research to summarize the various types of errors and find out the errors with high-frequency of appearance in 88 English writings submitted by

the engineering college students in NPU on the correcting network. Moreover, this study tends to analyze the possible causes of such errors and also provides reference for the English writing teaching.

2. Theoretical Basis

This study is based on Coder's Error Analysis Theory which originated from his work "The Significance of Learners' Errors". The essence of this theory is to analyze and compare the differences between the language generated by the learners in the process of second language acquisition and the language generated by the native speakers, and to analyze the causes of errors so as to put forward corresponding solutions. Coder holds that there are five stages involved in the process of error analysis i.e. collection of a sample, diagnosis of learners' errors, description and classification of errors, explanation of errors and evaluation of errors.

Besides, there are different scholars' classification of errors. Richards (1971) has divided the errors into interlingual errors and intralingual errors according to their sources. Intralingual errors are caused by learners' lack of proficiency in the rules of the target language, while interlingual errors are mainly caused by the negative transfer of mother tongue. Carl James (2001) classifies the errors from three levels i.e. substance, text and discourse. Substance errors includes misspelling, mispronunciation, etc. Text errors includes lexical errors and grammatical errors. According to Carl James, in addition to the formal errors of lexis, the main lexical errors lie on the semantic level including the confusion of sense relations as well as the collocational errors, while the grammatical errors include the morphology errors and the syntax errors. The morphology refers to the the morphological characteristics and changing rules of words, while syntax refers to sentence components and sentence types (Hao, 2019). To make it clear and easy to be analyzed, the author categorizes the morphology errors into the lexical errors. Besides, discourse errors mainly include the coherence errors and the pragmatic errors. In this paper, the author makes a little adjustment of Carl James's classification and regards the revised classification as the guidance.

3. Research Methodology
3.1 Research Questions and Research Methods

This study tries to answer two questions:
① What are the types and distribution of errors in these college students' English writings?
② What are the causes of such errors in these college students' English writings?

To solve these questions, this study uses quantitative analysis to identify the types and distribution of the errors, and uses qualitative analysis to give examples of some errors to analyze them specially. Meanwhile, this study tends to use the English native speakers' corpus i.e. COCA to help test if the errors marked by the network are correct. For example, in the process of corpus-assisted analysis, collocation frequency can be used to describe the phenomenon of negative transfer of mother tongue in the collocational errors. Through the retrieval in the native speakers' corpus, the commonly used collocations of certain words can be found so that the errors can be identified. This paper adopts these concepts of corpus as the assistance for the error analysis. As for the analysis of the causes of the errors, this paper analyzes them from the interlingual perspective and

the intralingual perspective.

3.2 Research Subject and Research Tool

The research subjects are 88 freshmen in NPU from different engineering majors. They were asked to write a composition with the title of "Reduce Waste on Campus" according to the directions of CET-4 writing. According to the results, finally, 88 English writings were submitted. The average word length is 4.4, and the average sentence length is 17.9. After the students submit their writings on the correcting network, the network will automatically mark the errors and categorize them into different types. In this study, the author reorganizes the errors based on Carl James's classification.

4. Results and Discussion

4.1 The Distribution of Errors

4.1.1 The Overall Distribution of Errors
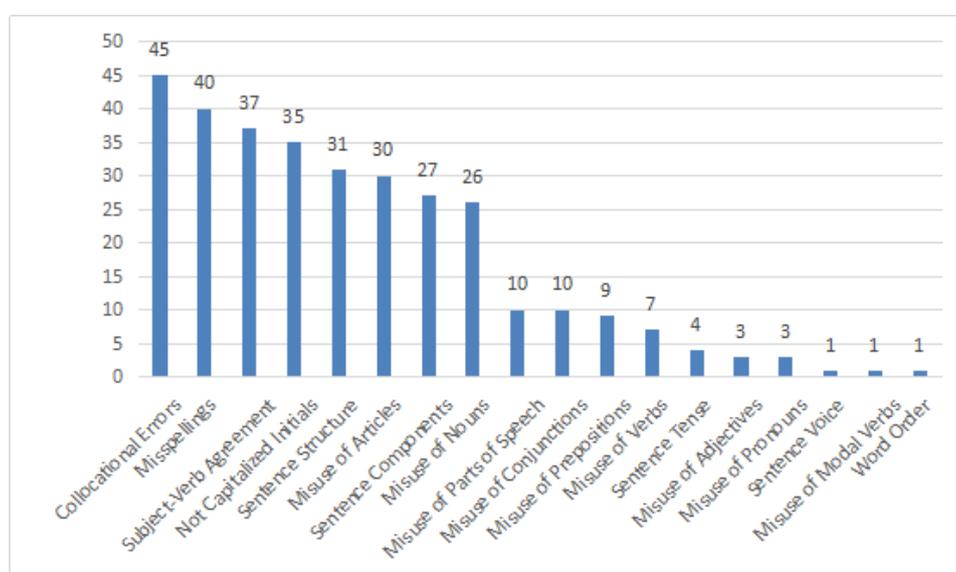
Figure 1: The Distribution of Errors On Correcting Network

According to the statistics in Figure 1, there are 18 types of errors, for a total of 320 errors marked by the correcting network. It can be seen from the Figure 1 that the collocational errors are the errors with the highest frequency of appearance i.e. 45 times. Misspellings have appeared 40 times which is also the errors that the students often make. Besides, the errors of subject-verb agreement, the errors of not capitalized initials, the errors of sentence structure, the misuse of articles, the errors of sentence components and the misuse of nouns also commonly appear in these students' English writings.

To make it clear in which aspect these students are weak, the author reorganizes these marked errors according to the revised Carl James's classification from three levels i.e. substance errors, text errors and discourse errors (The morphology errors are categorized into the lexical errors). The results are shown in Table 1.

Table 1: The Distribution of Errors Based on the Revised Carl James's Classification of Errors

	Substance Errors		**Text Errors**		**Discourse Errors**
	Misspellings	Not Capitalized Initials	Lexical Errors	Grammatical Errors	-
Quantity	40	35	134	111	-
Percent	12.5%	10.9%	41.9%	34.7%	-
Total	**75**		**245**		**-**

According to Table 1, it can be seen that students are more likely to make errors on the text level. The total number of the text errors are 245. Then the number of substance errors is 75, while there are no discourse errors in the 88 English writings. Among the errors marked by the correcting network, misspellings and not capitalized initials belong to the substance errors. Such errors are usually caused by students' carelessness and their weak memory of words. For example, in the 88 English writings, there are 40 errors of misspellings, like "prople (people)", "fefuse (refuse)", "imorove (improve)", etc. And there are 35 not capitalized initials. It is obvious that substance errors are easy to be corrected as long as students become more careful and serious when doing writings. As for the discourse errors, according to Carl James (2001: 161), it involves the coherence errors and the pragmatic errors. For the coherence errors, Carl has clearly differentiates between "cohesion and "coherence". In terms of "coherence", it emphasizes on the level of meaning and is more focused on the communicative function, which involves "the writer's intention and the reader's interpretation (James, 2001:161)." Because the discourse errors are closely related to the topic and theme of the writing, it's hard for the correcting network to identify them. In terms of "cohesion", it emphasizes the connections between sentences in form. It stressed on the cohesive links like reference, substitution, ellipsis, conjunctions and lexical cohesion (James, 2001: 159). Therefore, among the 18 types of errors, the errors of misuse of conjunctions belong to the text errors. In all the 320 errors, text errors accounts for 73.5% , which needs to be paid much attentions to. Therefore, in the following, the author discusses the text errors specially. The substance errors and discourse errors are not be discussed.

4.1.2 The Distribution of Text Errors

According to Carl James (2001: 142), text errors "arise from ignorance and misapplication of the lexico-grammatical rules of the language." Therefore, in this category, the errors can be divided into the lexical errors and the grammatical errors. The detailed classification of the 134 lexical errors and the 111 grammatical errors are shown in Table 2 and Table 3.

Table 2: The Distribution of Lexical Errors Based on the Revised Carl James's Classification of Text Errors

Lexical errors	Quantity	Percent
Collocational Errors	45	33.6%
Misuse of Articles	30	22.4%
Misuse of Nouns	26	19.4%
Misuse of Parts of Speech	10	7.5%
Misuse of Prepositions	9	6.7%
Misuse of Verbs	7	5.2%
Misuse of Adjectives	3	2.2%
Misuse of Pronouns	3	2.2%
Misuse of Modal Verbs	1	0.7%
Total	**134**	**100%**

Table 3: The Distribution of Grammatical Errors Based on the Revised Carl James's Classification of Text Errors

Grammatical Errors	Quantity	Percent
Subject-Verb Agreement	37	33.3%
Sentence Structure	31	27.9%
Sentence Components	27	24.3%
Misuse of Conjunctions	10	9.0%
Sentence Tense	4	3.6%
Sentence Voice	1	0.9%
Word Order	1	0.9%
Total	**111**	**100%**

Carl James (2001) holds the view that "lexical errors are the most frequent category of error" for some learner groups and "lexical errors out number other types by three or four to one (Meara, 1984, James, 2001)." Based on the statistics in Table 2 and Table 3, it is obvious that there are more categories of lexical errors than those of grammatical errors. According to Table 2, there are nine categories of errors in the lexical errors. Collocational errors appear 45 times in the 88 writings, accounting for 33.3% of the lexical errors. The errors of misuse of articles and errors of misuse of

nouns are also the errors students often make, and they respectively account for 22.4% and 19.4% in the lexical errors. The rest categories of the lexical errors have lower rate which is below 8 %. That means these students have a good master of the use of parts of speech, prepositions, verbs, adjectives, pronouns, and modal verbs. According to Table 3, there are seven categories in the grammatical errors. Students often mistaken the use of subject-verb agreement which appear 37 times in these writings. This kind of error is easy to be corrected, so the author does not discuss it in the following. Then it is the errors of sentences structure which accounts for 27.9% of the grammatical errors, while the errors of sentence components account for 24.3%. The rest of the grammatical errors are not the errors with high rate which is below 10%.

Therefore, in the following, the author mainly gives some examples to analyze the errors with high frequency of appearance i.e. collocational errors, misuse of articles, misuse of nouns, sentence structure and sentence components.

(1) Collocational Errors

Collocations are the other words any particular word normally keeps company with (James, 2001:152). Such errors can be intralingual or interlingual. After organizing the 45 collocational errors. There are four errors that appear for many times. They are the collocations with "waste".

Table 4: Collocations with "waste"

Collocations with "waste"
the wasting behaviors, the waste phenomenon, the waste appearance

Through searching the collocations with waste in COCA, there are not matching records of the above collocations, which means these are totally Chi-English. And when "waste"is used to modify nouns, the collocation is usually the "waste problems". This collocation appears in COCA corpus for 50 times. Actually when "waste" is used as a noun, it includes the meanings of "behaviors"and "phenomenon". For example, there are many collocations like "a waste of", "a huge waste", "combat wast" appearing frequently in COCA. When they are translated to Chinese, the word "waste" can be added the meaning of "behaviors"and "phenomenon" to make the translation more nature. But in English, "the wasting behaviors", "the waste phenomenon", etc. are not used by the native speakers. Besides, there are also some other collocational errors influenced by Chinese, such as "personal healthy growth", "the rest food", "the living qualities" which are the literal translation version of Chinese. Take "the living qualities" as an example. This collocation only appears three times in COCA. The natural expression with the same meaning is "the quality of life" which appears 1855 in COCA.

Besides, there are also some collocational errors caused by the students' inaccurate understanding of the word meaning. For example, "it (waste) is just a insignificant affair to them." In this sentence, the collocation "insignificant affair" does not fit the context. According to the Collins English Dictionary Online, the word "affair" refers to "a vaguely specified social event" or "a usually secretive

or illicit sexual relationship". The implicit meanings are not suitable to equal to "waste". In such context, it's more suitable to use the word "issue."

(2) Misuse of Articles

In terms of the misuse of articles, usually, the article is missing or unnecessary. For example, "reduce the waste on". In this phrase, the article "the" is unnecessary. Based on Juku Corpus, the corpus of the correcting network, there is no matching record of "reduce the waste on", but the phrase "reduce waste on" appears above 400 times. Moreover, some students forget to add the article when they particular refer to a noun. For example, "Campus is where we students study peacefully." In this sentence, it should be "The campus is where...". Besides, some students cannot distinguish the use of "a" and "an". For example, "a insignificant affair". "Insignificant" begins with a vowel, so it should be "an insignificant affair". In addition, students ignore the articles in some fixed collocations, such as "in canteen", "in restaurant", "as student", "as saying goes".

(3) Misuse of Nouns

The errors of nouns are about the singular or plural forms of nouns. These errors are usually caused by students' carelessness and their lack of the morphological knowledge . The below table shows some of the errors.

Table 5: Misuse of Nouns

Misuse of Nouns
few attention (attentions), the duty of minority (minorities), those resource (resources), meet our demand (demands), wastages (wastage), undertake more obligation (obligations), this things (thing), etc.

(4) Sentence Structure

The errors of sentence structure are usually the explicit grammatical errors. The examples are as follows.

For example, "Young people often buy things needless and throw away them quickly." In this sentence, there are two mistakes. First, the position of the word "needless". "Needless" should be put before the noun "things" as an adjective. Second, the word "them" should be put between "throw" and "away". According to the grammatical rules, the pronouns should be put between the two words in phrases like "throw away", "take away", "take in", etc.

There is another example——"In my opinion, undergraduates are supposed to be responsible and farsighted, instead of self-centered." In this sentence, a non finite verb is missing in the last half "instead of self-centered". It should be "instead of being self-centered". And in another example——"Reduce waste on campus requires the joint effort of all.", it is obvious that "reduce" is supposed to be "reducing" so that the clause "reducing waste on campus" can become the subject otherwise there is no subject in this sentence. Besides, some sentences are not complete in structure.

For example, "Because what we do can protect our resource and environment." According to the grammatical rules, the "because" clause cannot be a sentence itself.

(5) Sentence Components

In terms of the errors of sentences components, they are usually caused by the absence of certain components, the wrong use of some components or the repetition of some components.

For example, "As is known to all, we need spend a lot of money and time to deal with these disturbing things." In this sentence, the word "need" is a notional verb, so it should be followed by the verb infinitive "to spend". Another example is "Such as, if you took a plastic bag when you go shopping..." In this sentence, the mistake is that "such as" is used as an adverbial independently, but "such as" cannot lead a sentence. In addition, in the following example, there exist two predicates. In the sentence "There are many reasons cause this phenomenon", the clause "cause this phenomenon" is the other predicate component, which is inconsistent with the grammatical rules. Such a mistake is caused by students' literal translation of Chinese. The natural expression should be "There are many reasons for this phenomenon".

4.2 The Reasons for Making Errors

From the above analysis, it can be seen that among all the errors, there are two main causes: interlingual errors and intralingual errors. First of all, from the interlingual perspective, the negative transfer of mother tongue has the deepest influence on students. Influenced by Chinese thinking patterns and Chinese linguistic habits, students often use literal translation to choose the corresponding English words with the similar meanings to Chinese, which usually causes the semantic losses, the semantic repetition or leads to expressions that do not exist in English.

Secondly, from the intralingual perspective, students' inaccurate mastery of the target language's grammatical rules is also an important reason for making errors. Through the above analysis, it can be seen that students neglect some of the rules when forming a sentence. These are the important factors that interfere with the correct use of lexis and grammar.

5.   Conclusion

Based on the analysis of the errors in the 88 English writings of the science and engineering's freshmen, this paper finds that collocational errors, misspellings, the errors of subject-verb agreement, the errors of not capitalized initials, the errors of sentence structure, the misuse of articles, the errors of sentence components and the misuse of nouns are errors easily made by these students based on the statistics on correcting network. To make it clear in which aspect the students are weak, the author reorganizes the 18 types of errors based on the revised Carl James's classification of errors. It is found that, students are easily to make errors on the text level, especially in lexis compared with the grammatical errors. For the lexical errors, collocations are particularly tough for students to master well. For the grammatical errors, it reflects that the freshmen still lack proficient mastery of grammatical rules. The causes of such errors can be explained from the interlingual perspective and the intralingual perspective, among which negative transfer of mother tongue is the most important

influencing factor.

In view of this, some feasible suggestions can be put forward for English writing teaching. First of all, in the process of teaching, teachers should strengthen the explanation of the related knowledge of vocabulary, and strengthen the input of the vocabulary's related phrases. What is the most important thing is that students should not be limited to only memorize the surface meaning of words. Secondly, in order to reduce the interference of mother tongue in the English learning process, teachers should use the lexical context as a teaching assistance. For example, by searching in the large-scale native language corpus such as COCA and BNC, some collocations used by the native speakers can be found, so that students can understand the meaning of words and the usage of collocations in the context of the concordance lines. At the same time, teachers can also summarize a large number of idiomatic chunks commonly used by the native speakers, and continues to do the inputs in the teaching process, so that helps students internalize the collocation knowledge. In addition, teachers can collect native speakers' writing texts or native English reading materials. By reading these materials, students can fully contact with the native language context and acquire relevant knowledge in English thinking patterns.

References

[1] Carl James. (2001). *Errors in Language Learning and Use*. Foreign Language Teaching and Research Press.

[2] Richards J. C. (1971). A non-contrastive approach to error analysis. *English Language Teaching Journal, XXV*(3), 204-219.

[3] Cao Jing & Xiang Yunhua. (2020). An analysis of the errors in collocations in English majors' compositions. *Journal of Educational Institute of Jilin Province, 36*(06), 98-101.

[4] Ge Lilian. (2003). The influence of "mother tongue transfer" on Chinese students' English expressions —— An analysis of typical examples of College English Writing. *Journal of Capital Normal University (Social Sciences Edition), 000*(004), 50-52.

[5] Hao Xuehui. (2019). *Study on grammatical errors in high school students' English writing based on the Pigai System* (Unpublished master's thesis). Yan'an University, China.

[6] Liu Xianghong. (2008). A case study of a non-English major's writing: Error analysis. *Foreign Language Research,* (02), 140-142.

[7] Sun Yang & Chen Xueqin. (2020). An analysis of passive voice errors in Chinese middle school students' English Writing based on CLEC corpus. *Journal of Kaifeng Vocational College of Culture & Art, 40*(4), 198-199.

[8] Wang Lei & Xu Xiaoyu. (2015). Error analysis of engineering university students' English Writing based on Correcting Network. *Journal of Xi'an Shiyou University (Social Sciences), 24*(1), 95-100.

[9] Zheng Ke & Li Daobo. (2008). Verb errors in the writings of Test for English Majors (Grade Four). *Journal of Jiangxi Normal University (Philosophy and Social Sciences Edition), 41*(01), 144-147.

[10] Zhang Huihui. (2020). A corpus based analysis of collocation errors in senior high school students' English Writing. *Overseas English*, (6),86-87.