

# An Analysis of Lexical Errors in Senior High School Students' English Writing: A Corpus-Based Study

Maojuan Ding

School of Foreign Languages China West Normal University, Nanchong 637000, China.

---

**Abstract:** When we see the previous studies on English writing errors, we can find some interesting phenomenon. First, most of these research objects are college students majoring in English and non-English majors. There are few studies on writing errors of middle school students. Therefore, this paper focuses on senior high school students and focuses on the teaching of English writing in middle schools, which can fill the gap of previous studies to some extent. Second, in terms of research content, many previous studies only classified writing vocabulary errors without comprehensive causes analysis, or did not give specific suggestions to reduce the frequency of errors. Based on this, this study will break through these shortcomings. It is also a corpus-based study and conduct a deeper analysis of high school students' English writing vocabulary mistakes and the relationship between it and writing quality.

**Keywords:** English Writing; Vocabulary Errors; Senior High School Students; Corpus

---

## 1. Introduction

With the continuous development of globalization, English learning has become more and more increasingly prominent. As is known to us all, writing is a key factor to measure students' comprehensive language competence. The Senior High School English Curriculum Standard (2020 Revision) not only requires students to be able to clearly express their own views and attitudes in writing, but also requires them to have logical thinking, critical thinking and innovative thinking in their writing.

In addition, the rapid development of modern technology also provides more rapid progress for foreign language teaching. Due to the convenience of corpus, the research on foreign language teaching using corpus has gradually extended to the analysis of words, phrases and syntax. All of these provide the possibilities for my research of this paper.

## 2. Literature review

In 1967, Corder pointed out that making mistakes is an inevitable process of foreign language learning, and he revealed the general rules of foreign language learning. Error refers to the defect of learners' oral and written expression, which is the part where learners deviate from the standard of language behavior in conversation and writing (Dulay, Burt & Krashen, 1982). Hu Zhuanglin, Liu Runqing et al. (1988) point out that error usually refers to the misuse or misunderstanding of the target language, which may be grammatical or pragmatic. Carl James (2001) believes that mistakes are the unsuccessful use of language.

For this study, the author mainly refers to GUI Shichun's and Yang Huizhong's criteria for the classification of vocabulary errors, and makes a systematic analysis of the vocabulary errors in senior high school students' compositions into two categories: morphological errors and lexical errors.

## 3. Research Design

Based on the second language input and output theory, error analysis theory and interlanguage theory, this paper analyzes the lexical errors in English writing of 149 senior high school students from six middle schools in Sichuan and

Chongqing province.

This project adopts the five steps of error analysis by Corder(1981), and carries out the research according to the standard of vocabulary error classification in The Chinese Learners English Corpus by GUI Shichun and Yang Huizhong (2003). It uses the corpus tools to analyze typical examples and studies students' compositions with the help of text analysis method. This supplement explains the characteristics and possible causes of lexical errors, and finally puts forward relevant suggestions.

This paper focuses on the following two research questions: (1) What are the most common vocabulary errors in high school students' English writing? (2) What's the relationship between vocabulary errors and writing quality?

## 4. Results and discussion

Before we do this task,we have to know the lexical error-tagging system.In this article,the author uses the explanation of the classification of lexical errors by Gui and Yang.

By sorting out various classification of lexical errors,it can be found that Gui and Yang's classification is most suitable for the preset study in that classification is based on careful analysis of the Chinese learner English corpus.Therefore,the present study employs the classification made by Gui and Yang.

The descriptive statistics of lexical errors are presented in the following table.

Table 1 Descriptive statistics of lexical errors

type	N	Minimum	Maximum	Mean	Std.Deviation
Wd1	23	0	5	0.28	0.81
Wd2	120	0	8	2.24	1.82
Wd3	135	0	11	2.79	2.23
Wd4	107	0	9	2.97	2.18
Wd5	86	0	7	1.97	1.79
Wd6	7	0	2	0.29	0.56
Wd7	39	0	3	0.71	0.86
Total number	149	1	12	8.22	6.05

From this table,we can found every student commits lexical errors.The subjects at least commit 1 error in their compositions.There is a student committing 12 errors in his writing.On average,each student commits 8.22 errors.It means that there are 8.22 errors in each composition.

### 4.1 Data processing and analysis

According to statistics, there are 3098 language errors in this sample, including 1753 vocabulary errors, that is, vocabulary errors account for 56.58% of language errors.Among these errors, spelling, capitalization, ellipsis and redundancy accounted for 86.53% of vocabulary errors.In particular, spelling errors account for a high proportion of total errors,( Haggan,1991)In particular, spelling errors accounted for 44.3% of the total errors.

This article lists several common spelling mistakes:

- (1) wrong alphabetical order: "against" is written as "agianst".
- (2) ellipsis error: "environment" is written as "environment".
- (3) error in adding letters: "widespread" is written as "widespread".

According to the current college entrance examination, the vocabulary in high school grammar mainly includes verbs, nouns, pronouns, articles, adjectives, adverbs and conjunctions. (Jia Guangmao, 2017) Therefore, this paper also makes statistics and analysis on the lexical errors of different parts of speech one by one.

Incorrect use of verbs

According to the characteristics of middle school students' compositions on this topic and the relevant literature, the verb use errors in this paper mainly include tense errors, voice errors, subject predicate inconsistencies, non finite verb errors,

etc. As shown in the table below, subject predicate inconsistencies account for the most.

Voice errors are mainly reflected in that some words used by students in writing do not reflect the passive voice, and the frequency of voice errors is relatively low.

Subject predicate agreement means that the singular and plural number of predicate verbs in a sentence should be consistent with the singular and plural number of subject in the sentence.(Jiang Jinyang,2021)In fact, students know this grammar, but they always make such mistakes due to carelessness.

Errors in the use of nouns are mainly reflected in the singular and plural and whether nouns are countable.For example,China has many smoker, who are making air pollution even worse.

#### Incorrect use of articles

The error of using articles is mainly reflected in the mixed use of a, an and the, especially the use of the. But we all know that the is generally used for someone or something, which has appeared in the previous text.

## 4.2 Correlation Between Vocabulary Errors and Writing Quality

The second question asks about the correlation between lexical errors and the quality of the composition as measured by the final score given by two English teachers.SPSS Pearson Correlation was run to test the correlation between the two variables.Table 4 offers the results.

Table 2 Correlation between lexical errors and writing score

		Writing scores	Number of lexical errors
Writing Scores	Pearson Correlation	1	-.581**
	Sig.(2-tailed)		.000
	N	149	149
Number Of errors	Pearson Correlation	-.581**	1
	Sig.(2-tailed)	.000	149
	N	149	

The result of the correlation test revealed a correlation coefficient of  $r=-.581$ ,which points to a significant negative correlation.The negative value of the correlation implies that,the scores of composition goes up as the number of lexical errors a composition displays,the lower will it score.The significant correlation implies that the number of lexical errors influences the writing quality very much.

## 5. Major Findings of the Study

The purpose of writing is not for a certain exam, which will cause pressure on students and fail to feel the fun of writing.(Zhao Dongqin,2008) Therefore, it is particularly important to diversify the ways of writing training and cultivate students' fun and habits of writing.At the same time, teachers should pay attention to the teaching of basic knowledge and emphasize the differences between English and Chinese expressions in ordinary classes.

This study can provide some theoretical reference value for senior high school English teachers in vocabulary teaching, writing teaching and feedback evaluation strategies. In addition, it can also provide specific practical guidance for the application of corpus in foreign language teaching. The application of this study to senior high school English teaching can improve students' vocabulary errors in writing to a certain extent.

## 6. Limitations and suggestions for further research

It is hoped that unnecessary interference factors can be eliminated as far as possible in future research to ensure the objectivity and authenticity of data acquisition. And the analysis of data can be more in-depth and thorough.

## References

[1] Brown, H.D.(2003).Principles of Language Learning and Teaching. Beijing: Foreign Language Teaching and Research Press.

- [2] Corder, S.P. (1967). The significance of learners' errors. *International Review of Applied Linguistics* (5):161-170.
- [3] Dulay, H. Burt, M. and Krashen, S.D. (1982). *Language Two*. Rowley, MA: Newbury House.
- [4] Granger, S. (1994). New insights into the learner lexicon: A preliminary report from the International Corpus of Learner English. In Flowerdew L & Tong KK (eds). *Proceedings of the Joint Seminar on Computers and Lexicology*. Hong Kong: Hong Kong University of Science and Technology, 128-145.
- [5] Haggan, M. (1991). Spelling errors in native Arabic-speaking English majors: a comparison between re-medical students and fourth year students. *System*(19):45-61.
- [6] James, C. (2001). *Errors in Language Learning and Use: Exploring Errors Analysis*[M]. Beijing: Foreign Language Teaching and Research Press, 2001:1-78. Lennon Paul. *Principles of Language Learning and Teaching*. San Francisco State University(4): 28-50.