

A Study of Errors Analysis in Online Commentary Writing of English Sophomores: Implications for Writing Teaching

Hongyu Hu¹

¹ School of Foreign Languages, Wuhan University of Technology, Wuhan, Hubei, China

ABSTRACT

Commentary writing is used to describe the profound essence of things, which can improve students' linguistic competence and cognitive mode. Nowadays, with the development of science and technology, the form of writing has changed. Except for offline writing, online writing has gained immense popularity. When writing, the phenomenon of errors is supposed to be unavoidable in the process of language acquisition and learning, previous studies have mainly focused on offline writing. However, little research has been implemented on errors made in online writing. This paper attempts to conduct an analysis of errors encountered in two online commentary writing assignments submitted by English sophomores, trying to study different types of English writing errors from the aspects of vocabulary, grammar and discourse, and identifying the sources of the errors students have made in writing. The results reveal that students make more intralingual errors rather than interlingual errors. The findings contribute to providing valuable perception into the types and frequencies of errors committed by English sophomores in online commentary in English. Furthermore, insightful recommendations are proposed in this study for both students and teachers, with the aim of enhancing the effectiveness of teaching and learning strategies in commentary writing.

Keywords: Error analysis, Online commentary in English, English sophomores.

1. INTRODUCTION

Commentary writing constitutes a form of critical analysis that entails providing insights, interpretations, and evaluations of a particular topic, text, or event. It can be considered as one of the most effective ways to know whether students have achieved the teaching objectives. With the progression of the digital era, an increasing number of commentary tasks necessitate completion in an online environment. Compared with writing offline, writing online will not only promote students' operation capability of using the Internet but also examine students' linguistic competence.

In this English course, students were required to upload their comments on an app. Based on Error Analysis, this study collects 72 online writing assignments chosen from two comments tasks to set up a self-built corpus, and tries to study students' errors that have been made in these two tasks. After

analyzing the errors students have made and the reasons for these errors, this paper will put forward some suggestions for teachers in teaching, and it will also help students better notice their errors and learn to avoid or reduce unnecessary errors.

2. LITERATURE REVIEW

2.1 The Classification of Errors

Since the 1950s, scholars have shown their own interpretations of errors. Corder (1967), who was the pioneer in classifying and putting forward the concept of Error Analysis (EA) which refers to a method that can be used to examine EFL learners' speech or written performance so as to understand the process of second language learning and make second language teaching more targeted. Errors in language learning have been categorized by researchers according to various conceptual frameworks. Corder (1971) categorized errors into

interlingual errors and intralingual errors based on their sources. Interlingual errors are attributed to the negative transfer of the learner's first language, where elements or structures from the mother tongue interfere with the acquisition and use of the target language. In contrast, intralingual errors arise from faulty or incomplete learning of the target language itself, reflecting misunderstandings or misapplications of its rules and structures. James (2014) divided errors into the aspects of the occurrence levels: substance, text and discourse.

Gui and Yang (2003) not only extended this category, but also classified errors in detail based on a self-built learners' corpus, summarizing 11 categories of errors, that is, morphology errors (fm), word errors (wd), collocation errors (cc), verb errors (vp), noun errors (np), pronoun errors (pr), adjective errors (aj), adverb errors (ad), prepositional errors (pp), conjunction errors (cj) and sentence errors (sn). More details can be seen in "Table 1".

Table 1. Classification of Speech Errors (Total: 61)

Morphology		Verb Phrase		Noun Phrase		Pronoun	
Encoding	Type	Encoding	Type	Encoding	Type	Encoding	Type
fm1	spelling	vp1	pattern	np1	pattern	pr1	reference
fm2	word building	vp2	set phrase	np2	set phrase	pr2	anticipatory it
fm3	capitalization	vp3	agreement	np3	agreement	pr3	agreement
		vp4	finite/non-finite	np4	case	pr4	case
		vp5	non-finite	np5	countability	pr5	wh-
		vp6	tense	np6	number	pr6	indefinite
		vp7	voice	np7	article		
		vp8	mood	np8	quantifiers		
		vp9	modal/auxiliary	np9	other determiners		
Adjective Phrase		Adverb		Prepositional Phrase		Conjunction	
Encoding	Type	Encoding	Type	Encoding	Type	Encoding	Type
aj1	pattern	ad1	order	pp1	pattern	cj1	pattern
aj2	set phrase	ad2	modification	pp2	set phrase	cj2	set phrase
aj4	-ed/-ing confusion						
aj5	predictive/attributive						
Word		Collocation		Sentence			
Encoding	Type	Encoding	Type	Encoding	Type		
wd1	order	cc1	noun/noun	sn1	run-on sentence		
wd2	part of speech	cc2	noun/verb	sn2	sentence fragment		
wd3	substitution	cc3	verb/noun	sn3	dangling modifier		
wd4	absence	cc4	adj/noun	sn4	illogical comparison		
wd5	redundancy	cc5	verb/adv	sn5	topic prominence		
wd6	repetition	cc6	adv/adj	sn6	coordination		
wd7	ambiguity			sn7	subordination		
				sn8	structural deficiency		
				sn9	punctuation		

2.2 Previous Studies in Error Analysis

Research in Error Analysis has mainly focused on written errors, with numerous studies investigating the frequency, types, and sources of errors in L2 learners. Some scholars find that the

most frequent errors students make are grammatical errors, especially verb errors (Nguyen & Nguyen, 2021; Shirban Sasi & Mark Lai, 2021; Ehsanzadeh & Dehnad, 2024). But some scholars argue that learners make article errors more frequently (Jia and Qiao, 2014).

Research also consistently demonstrates that language transfer significantly contributes to writing errors. Maqbool, Ghani & Khan (2018) discovered that learners from L1 backgrounds without an article system commit more errors in English writing compared to those with an article system.

3. METHODOLOGY

3.1 Research Purposes and Research Questions

Although there is a lot of research on writing errors, few studies have focused on the errors of English majors. The aim of this research is to investigate the types of errors English sophomores make in online commentary, trying to find out the possible sources of those errors. The three research questions are designed as follows:

- What are the types of errors among students in online commentary in English?
- What are the causes for errors in English writing of English sophomores?
- What are the distribution features of errors among students with different English achievements?

3.2 Research Participants and Materials

All of the participants in this research are English sophomores from two different classes (N=40) in the same university and they are taught by the same teacher and these two classes have close scores in the final exam last semester. Each of the participants is given the same two topics to

write down comments and they are required to hand in their assignments at the same time. All the students are asked to type their assignments online and they are allowed to look up some information online within a limited time. After removing some invalid samples, a total of 72 copies were collected.

3.3 Research Methods

Two methods are used in this study, that is, corpus and text analysis. This study follows the five main steps proposed by Ellis (1997): collect the samples from learners' language, identify errors in the self-built corpus, describe the errors, and give explanations for errors. After collecting 72 samples, in order to better understand Chinese college students' errors, this study decides to start based on the taxonomy of Gui and Yang (2003). After all samples are processed, the data will be retrieved and analyzed by the corpus tool AntConc. Also, text analysis will be used in this research to analyze the statistical data. And the collected data are used to analyze the sources and reasons for students' writing errors.

4. DISCUSSION

4.1 Data Analysis

With the help of AntConc, the number of tokens in this self-compiled corpus is 10265. There are 194 errors in this corpus. The data can be seen in "Table 2". It shows that 35.57% of the students' writing errors are vocabulary errors, 41.24% are grammar errors, and 23.71% are discourse errors. More details can be seen in "Table 3".

Table 2. The distribution of errors

Type	Frequency	Percentage
Vocabulary	68	35.05%
Grammar	80	41.24%
Discourse	46	23.71%

Table 3. The distribution of errors based on CLEC

Type	Frequency	Percentage
fm	27	13.92%
wd	30	15.46%
cc	11	5.67%
vp	44	22.68%
np	16	8.25%
aj	2	1.03%
ad	1	0.52%

Type	Frequency	Percentage
pp	7	3.61%
pr	10	5.15%
cj	2	1.03%
sn	44	22.68%

4.1.1 Errors in Vocabulary

Vocabulary errors can be divided into morphological errors (fm), word errors (wd) and collocation errors (cc). After collecting the data,

there are 68 vocabulary errors, which accounts for 35.05% in proportion of total errors. Based on the corpus, more word errors have been made in word and morphology in vocabulary errors.

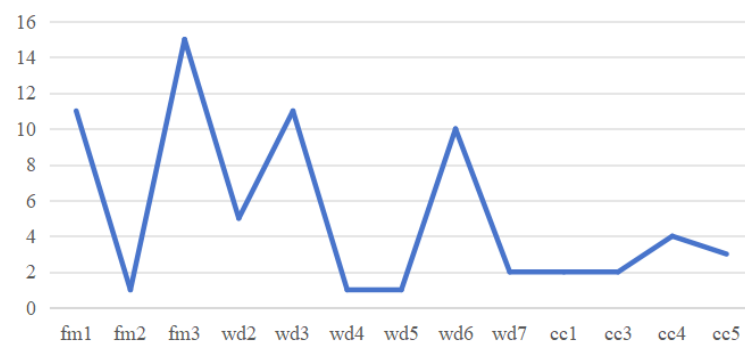


Figure 1 Frequency of errors in vocabulary.

Firstly, it is obvious to see in “Figure 1” that in morphological errors (fm), students always make errors in capitalization (fm3) and word spelling (fm1). There are 15 capitalization errors and 11 word spelling errors, which accounts for 13.40% in proportion of total errors. And some typical errors in morphology are listed as follow:

- I want to recommend the book *the Lychee of Changan* <fm3> (correct: *The Lychee of Changan*).
- After I finish appreciating the light-hearted film, My <fm3> best favourite figure is kim <fm3> (correct: my; Kim).
- The first kind of book owner is not recommended unless <fm1> you are wealthy beyond description (correct: unless).
- I’d like to remmend <fm1> *And Then There Were None* (correct: recommend).

As is shown above, students make errors in capitalization in different ways. In the first example, students make capitalization errors, which reveal that they do not master the format of the book title. As for the second example, student makes an error in the name of a character and he also forgets about

the comma and therefore he does not use the low-case form of “my”. In last two examples, students make wrong spellings of “unless” and “recommend” respectively because they cannot accurately remember the correct forms of those words.

Then according to the data, students also make a lot of errors in word including part of speech (wd2), substitution (wd3), repetition (wd6) and ambiguity (wd7), which accounts for 15.46% of the total number of errors. For example:

- (1)...but scarcely evoke intellectual resonance out of their merely <wd2> respect for books’ physical appearance (correct: mere).
- (2) I was most impressed by Michael, the protagonist’s ex-fiancé <wd3> (correct: ex-fiancé).
- (3) Her determination and persistence not only touched the role in the film, but also deeply moved me and moved me <wd6> (correct: moved me).
- (4) And what Julianne chose to do at the ending of the film has justified her

kindness, ration <wd7> and courage
 (correct: reason).

In these examples, students make errors in word choice. In the first example, “merely” is an adverb which cannot be used to link the pronoun “their” to the noun “respect”. For the second example, although students choose the right part of speech of the word, they ignore the slight difference between the word that they really want to express and the word they actually use. And for some students, they use their words only according to its literal meaning instead of thinking twice, which show that they do not completely master these words. In example (2), Michael is a male lead, however, the word “ex-fiancé” refers to a female which is improper. So “ex-fiancée” is more accurate. And students sometimes want to emphasize too much which causes unnecessary repetition. In examples (3), the repetition of “moved” redundant. Some students also make errors in memorizing the meaning of words. In example (4), “reason” instead of “ration” is more proper to show the girl’s characteristics. And under some circumstances, students will make errors in grammar because of the omission of a word.

Besides, students will also make collocation errors which means that they do not totally understand the collocation rules. It is common for students to make “adjective+noun” collocation errors which makes the expressions nonnative. Some examples are provided.

- (1) From this book, I saw the difficulties of small people <cc4> in the face of power, and also saw the darkness of feudal society (correct: few people).
- (2) He is full of understanding, understanding the awkward and uneasy friends <cc4>. (correct: the awkwardness of friends).

In example (1), “small people” is completely interfered by their first language. “Few people” may be more appropriate if the student wants to express the meaning of “a some amount of people”. And in example (2), “awkward and uneasy friends” is a little unclear in context and semantics. Both “awkward” and “uneasy” describe the uncomfortable or difficult situation. But it is strange to use these two words directly to modify the word “friends”. So here “the awkwardness of friends” may be better.

4.1.2 Errors in Grammar

There are 80 grammar errors in the corpus, consisting of verb errors (vp), noun errors (np), adjective errors (aj), adverb errors (ad), preposition errors (pp) and pronoun errors (pr). Among these grammar errors, verb errors account for the largest proportion, and few errors have been made in adjective and adverb. Detailed data can be seen in “Figure 2”.

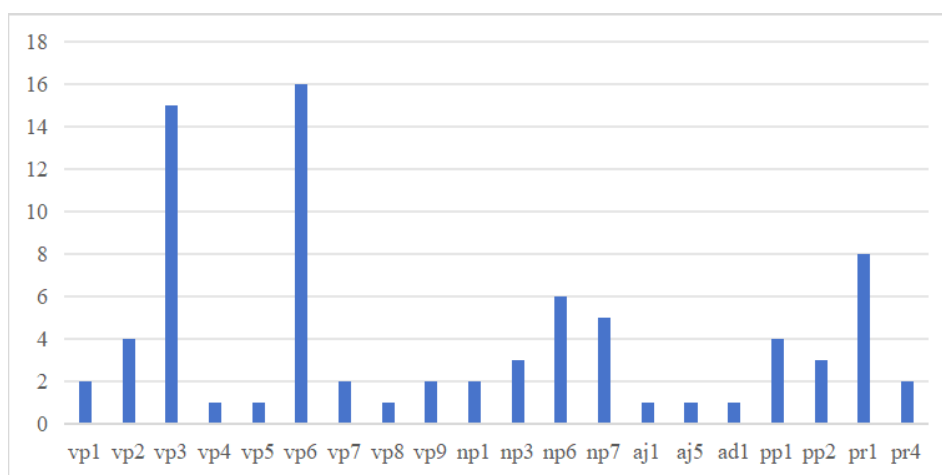


Figure 2 Frequency of errors in grammar.

Verb errors occur the most frequently and the number of verb errors is 44 which accounts for 22.68% in the whole error types. Students always make errors in set phrase (vp2), agreement (vp3), and tense (vp6). And a small amount of errors are

made in pattern (vp1), finite/non-finite (vp4), non-finite (vp5), voice (vp7), vp8 (mood) and modal/auxiliary (vp9). For instance:

- (1) These book owners seem have <vp1> no individual preference or judgement on

books, thus they tend to lose interest in the books they bought before long (correct: seem to have).

- (2) From the point of view of friends, he is really a friend worth making. He heavy love heavy righteousness, originally hate to fly he did two planes <vp2> in three days to help her (correct: took two planes).
- (3) Her actions embodies <vp3> themes of love, friendship, and personal growth, making her more sincere (correct: embody).
- (4) At that time, the most impressive plot of the book is <vp6> that the heroine chooses <vp6> to commit suicide after the hero's sacrifice (correct: was; chose).
- (5) This plot is <vp7> the most moved me, because to do something <vp5> you don't like for a person shows that this person is very important in your heart (correct: moves; doing).

The first example shows that students misuse the verb phrases because of their unclear memory. In this sentence, "seems to do" belongs to a particular phrasal verb pattern where "seem" is the main verb and "to do" is the infinitive phrase of the verb as a complement to "seem". So "seem have" should be modified to "seem to have". And for some students, they are more likely to make errors because of the influence of their first language. The second example also illustrates that the errors students make in verb phrases are also strongly affected by their mother tongue. It seems that "did two planes" means "took two flights" in literal meaning. However, it refers to "making two planes" in reality. And compared with English, Chinese verbs do not have any tense change, which also causes the errors in agreement and tense. The third example and the fourth example both show that students will write without considering the form of the subject. In example (3), the verb should be suitable with "actions". In the example (4), errors are made in tense. The phrase "at that time" is always related to the past tense, so the verbs "is" and "chooses" should be changed. Moreover, the student makes errors in the use of voice and the use of infinite. The relationship between "plot" and "me" should be active rather than passive. And here "doing something" would be more appropriate, because the emphasis here is on an ongoing behavior, not a single event.

Except for the errors in verbs, noun errors are also common which account for 8.25% in total errors. And some examples are listed as follows:

- (1) Fortunately, she was aware of her guilty <np1> and halted her evils promptly (correct: guilt).
- (2) As a minority group <np3>, he is not inferior or concealed (correct: group member).
- (3) After all, she maintained a vague and subtle relationships <np6> with Michael, not daring to confess her love (correct: relationship).
- (4) In the film, several scenes indeed leave a <np7> indelible impression on me (correct: an).

As for the first example, it shows that although students have formed a certain awareness of affixes, they have not fully mastered them. They take it for granted that "-ty" is a noun suffix. However, the word "guilty" is an adjective, and the word "guilt" is the noun form. Students also tend to ignore articles before nouns when writing online. In example (2), the noun "group" refers to a number of people or things. So it is supposed to be changed to "group member". And in example (3), "relationships" should be correct to "relationship" in order to match with the preceding indefinite article "a". Sometimes students choose the wrong article, even though they notice the singular form of the noun. In the last example, errors about "a/an confusion" occur. "Indelible" is a noun starting with vowel phonemes. So "a" should be replaced with "an".

Students make errors in adjectives, adverbs, prepositions and pronouns at the same time. For instance:

- (1) who abandoned his wife and children, abandoned the life that others seemed to be rich and delighted <aj1> (correct: delightful).
- (2) This plot is the most <ad1> moved me, because to do something you don't like for a person shows that this person is very important in your heart (correct: moves me the most).
- (3) The whole story is set in a grocery store, through <pp2> the perspective of the owner, Atsushi Kanzaki, tells the stories of different customers, and permeates a profound life philosophy in each story (correct: from the perspective of the owner).
- (4) She is brave and confident, but also hesitant, he <pr1> rarely shows his <pr4>

love for the people he <pr1> loves (correct: she; her; she).

In the example (1), there is an error in combination with the noun “life”. So here “delighted” should be changed to “delightful”. In the second example, the adverbial phrase “the most” should be used after the modified verb “moves”. And students often ignore the formation of idiomatic prepositional phrases. In example (3), “from the perspective of” is a fixed phrase. There are also incorrect pronoun references caused by carelessness. The subjects of (4) are women, so the pronoun should be changed.

4.1.3 Errors in Discourse

Students make fewer errors in discourse than in vocabulary errors or in grammar. The proportion of discourse errors is 23.71%. The errors in discourse can be divided into sentence errors and conjunction errors. In contrast to conjunction errors, students make more errors in sentences, including run-on sentences (sn1), fragment sentences (sn2), coordination (sn6), subordination (sn7), structural deficiency (sn8) and punctuation (sn9). More details about the discourse errors can be seen in “Figure 3”. And some examples are written as follows:

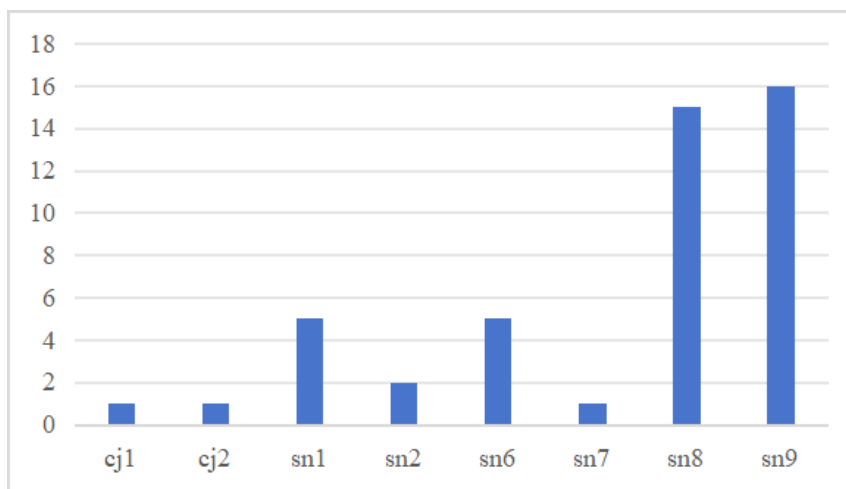


Figure 3 Frequency of errors in discourse.

- (1) ... and the third type of reader is the person who will understand the content thoroughly, <sn1> is the most beneficial person (correct: the third type of reader who will understand the content thoroughly).
- (2) Of course, he is not only a blunt persuade <sn6> a good friend <sn2>, he is still in your hesitation <sn8> (correct: Of course, he is not only a blunt persuader but also a good friend who supports you even in your moments of hesitation).
- (3) Draws the reader to read on <sn8> (correct: It draws the reader to read on).
- (4) The first kind of readers, I think they are more like ";<sn9> "Collector" <sn9> (correct: ... I think they are more like “collectors”).

The student makes an error in the run-on sentence in the first example. The sentence improperly joins two clauses without appropriate punctuation or conjunction. The part “is the most beneficial person” is an independent clause that is

improperly connected to the preceding part. Examples (2) and (3) both make structural deficiency errors, which make the sentences broken. Furthermore, in the second example, the phrase “not only a blunt persuade a good friend” lacks parallel structure and “a good friend” as a standalone element is a fragment that does not form a complete sentence. The sentence “he is still in your hesitation” is confusing, which needs correcting. In the fourth and fifth examples, students use the wrong punctuation.

According to the collected data, fewer errors have been made in conjunction, which only accounts for 1.03% in the total errors. For example:

- (1) She seemed to be arrogant and careless while it was indeed not the same case, in contrast <cj1>, she embraced the characteristics of vulnerability... (correct: She seemed to be arrogant and careless, but in reality, she embraced the characteristics of vulnerability).

In this example, the conjunction “while” is used incorrectly in combination with the phrase “it was indeed not the same case”.

4.1.4 The Possible Sources of Errors

After sorting out all the errors, it is important and necessary to find out why students make these errors. This study simply investigated the sources of each error from the perspective of interlingual errors and intralingual errors. Intralingual errors account for 75.26%, while interlingual errors account for 24.74% which is far more than interlingual errors. That is to say, although Chinese English majors will make errors because of the interference of their first language, they will make more intralingual errors.

4.2 Results and Implications

4.2.1 The Distribution of the Errors

Firstly, for vocabulary errors, it is found out that the frequency of word errors is the highest and the collocation errors happens the least in students’ writing samples. Among the word errors in this corpus, the frequency of substitution errors and repetition errors rank higher while the frequency of absence and redundancy errors occur less. And the highest frequency of vocabulary errors students make is capitalization errors. Secondly, for grammar errors, verb errors take up the highest proportion. Among these verb errors, most errors

occur in agreement and tense. In addition, many students also make a lot of errors in nouns and pronouns. As for discourse errors, sentence errors take up a higher proportion than conjunction errors. And most sentence errors occur in structural deficiency and punctuation.

4.2.2 Distribution Features among Students with Different English Achievements

This study attempts to separate the writing samples into two parts according to the grades of TEM4. And the results are shown in “Table 4”, “Table 5” and “Table 6”. Students who got scores higher than the average are defined as high achievers and the rest students are viewed as low achievers. It is normal to find that high achievers write down more words than low achievers, and they make more errors as well. However, according to the ratio of errors and total words, it is found that the error rate of low achievers is higher than that of high achievers. And the errors high achievers made seem to appear more in the dimension of discourse. In comparison to high achievers, the errors committed by low achievers tend to primarily reside at the grammatical level. For both high and low achievers, the intralingual errors are higher than interlingual errors. High and low achievers made similar interlingual errors, though high achievers had a slightly higher rate, which suggests that advanced learners may commit more intralingual errors.

Table 4. Errors of high and low achievers

	High achievers	Low achievers
Number	22	18
error rate	1.92%	1.99%
interlingual error rate	0.51%	0.51%
intralingual error rate	1.56%	1.28%

Table 5. The distribution of errors of high achievers

Type	Frequency	Prop
sn	30	25.86%
vp	21	18.10%
fm	18	15.52%
wd	18	15.52%
np	14	12.07%
cc	5	4.31%

Type	Frequency	Prop
pr	4	3.45%
aj	2	1.72%
pp	2	1.72%
cj	2	1.72%
ad	0	0.00%
total	116	

Table 6. The distribution of errors of low achievers

Type	Frequency	Prop
vp	23	29.49%
sn	14	17.95%
wd	12	15.38%
fm	9	11.54%
pr	6	7.69%
cc	6	7.69%
pp	5	6.41%
np	2	2.56%
ad	1	1.28%
aj	0	0.00%
cj	0	0.00%
total	78	

4.3 Implications

Through the data collected, it is found that students will make the same errors to some extent when they write their comments online. They make the most errors in grammar. And their writings online still are strongly affected by language transfer. Firstly, in order to solve the interlingual errors, the teachers still need to emphasize the difference between Chinese and Western cultures which will help students reduce the barriers caused by negative transfer. According to this study, most errors are made in grammar. So it would be better if the teachers could compare Chinese grammar rules with English grammar rules to deepen students' impressions. Also, it is significant for teachers to let students understand various expressions in Chinese and English so as to improve their English thinking ability. Secondly, it is clear to see that students make more intralingual errors compared to interlingual errors. That's to say, students still have trouble understanding the rules and patterns of English. So more input is suggested to be provided. Students are supposed to enlarge their reading materials, which can strengthen their understanding of English. And since students with different achievements show some differences in writing

errors, teachers are supposed to adopt stratified teaching differentiated guidance to better adapt to the needs of students. And when writing, students are suggested to check their online writings several times to make sure that some avoidable errors will not occur.

5. CONCLUSION

This study investigates and analyzes the types and frequency of writing errors online from English major students, finding that the proportion of grammatical errors is the highest and discourse errors are relatively fewer, and then it tries to analyze the sources of these errors finding out that although students are interrupted by negative transfer, they make more interlingual errors. In order to reduce and avoid these errors, this paper also attempts to put forward some suggestions. For teachers, they need to emphasize the differences between Chinese and English. And they are advised to adopt different strategies for students with various achievements. For students, they need to expand their reading.

And there is also some room for improvement in this paper. More data needs to be collected to make it more comprehensive and universal. Since

this study only takes comment writing into consideration, scientific works of other sorts of writings need to be deepened to make the research holistic.

REFERENCES

- [1] Corder, S.P. The Significance of Learner's Errors[J]. *International Review of Applied Linguistics in Language Teaching*, 1967 (5): 161-170.
- [2] Corder, S.P. Idiosyncratic dialects and error analysis[J]. *International Review of Applied Linguistics in Language Teaching*, 1971, 9.
- [3] Ehsanzadeh S J, Dehnad A. Analysis of high-frequency errors and linguistic patterns in EFL medical students' English writing: Insights from a learner corpus[J]. *BMC Medical Education*, 2024, 24(1): 1264.
- [4] Ellis, Rod. The study of second language acquisition research[M]. Oxford: Oxford University Press, 1997.
- [5] Gui Shichu, Yang Huizhong. Chinese Learner English Corpus[M]. Shanghai: Shanghai Foreign Language Education Press, 2003. (in Chinese)
- [6] Maqbool, Saira; Ghani, Mamuna; Khan, Asif. Error Analysis of English Language Speakers in the Use of Articles in Pakistan[J]. *Journal of Educational Research*, 2018, Vol.21(1): 177-188.
- [7] James C. Errors in language learning and use: Exploring error analysis[M]. Routledge, 2014.
- [8] Jia Guanjie, Qiao Liangwen. Research on Language Error Analysis in Master's Theses of English Majors[J]. *Foreign Language World*, 2014, (3): 63-69, 96. (in Chinese)
- [9] Nguyen K D, Nguyen U H T. Common errors in writing of EFL sophomores in a context of the mekong delta[J]. *Eurasian Journal of Applied Linguistics*, 2021, 7(2): 46-57.
- [10] Shirban Sasi A, Lai J C M. Error analysis of Taiwan region University students' English essay writing: A longitudinal corpus study[J]. *International Journal of Research in English Education*, 2021, 6(4): 57-74.