

# *Study on Discourse Cohesion in Chinese English Learners' English Argumentative Compositions Based on Coh-Metrix*

Aohang Yuan\*

*School of Foreign Languages, Southwest Jiaotong University, Chengdu, 610031, China*

*\*Corresponding author: oliveryuanaohang@outlook.com*

**Keywords:** English composition; discourse cohesive devices; corpus linguistics; Coh-Metrix; Chinese English Learner

**Abstract:** Writing, as an important component of English teaching, plays an important role in cultivating language output and communication skills and is also a key link in cultivating critical thinking. The cohesive devices in the composition not only reflect the author's thinking ability but also their ability to comprehensively use language. Based on the cohesion theory proposed by Halliday in systemic functional linguistics, this study establishes a small corpus and uses Coh-Metrix, the discourse cohesion analysis tool, to research the use of discourse cohesion devices in Chinese students' (L2) English compositions and native English students' (L1) English compositions. It has been found that there is a significant difference in the use of discourse cohesion between L1 and L2. At the same time, this study explored the reasons for the differences in the distribution of discourse cohesion types through semi-structured interviews and concluded that the influence of mother tongue, textbooks, teachers' attention to grammar cohesion teaching, and students' subjective motivation can explain this difference. This study aims to draw the attention of teachers to the teaching of discourse cohesive devices, as well as the attention of English learners to the application of cohesion.

## 1. Introduction

The College English Teaching Guide (2020 Edition) issued by the Ministry of Education of the People's Republic of China (hereinafter referred to as the Guide) points out that college English courses aim to promote and expand English teaching, and the main purpose is to further improve students' English proficiency in listening, speaking, reading, writing and translation after high school English learning. Specifically, for college students, the main focus should be on improving their comprehensive English proficiency. As an important part of English teaching, composition plays an important role in training language output and cultivating communication skills, and is also a key link in cultivating critical thinking ability. Currently, with the deepening of China's opening up to the world and the advancement of globalization, whether contemporary college students can write in authentic English is related to whether China can communicate with countries around the world.

However, Huang<sup>[1]</sup> argues that there are problems in undergraduate students' English compositions,

such as discontinuous topics, weak logic, and improper selection and use of cohesive devices. At the same time, compared with native English speakers, Chinese students' English composition generally has problems with unidiomatic words and unnatural language expression, especially in the use of cohesive devices in their compositions, which reflects the problem of discourse coherence in Chinese college students' English compositions. Essentially, one of the ways to evaluate the quality of composition depends on the cohesive devices in it. To some degree, the cohesive devices in composition not only reflect the thinking ability but also reflect the author's comprehensive language application ability. In English composition, good cohesive devices can be regarded as an important aspect of assessing students' critical thinking ability. Therefore, to put forward more solutions to the problem of discourse cohesive devices, it is necessary to find the difference in the use of discourse cohesive devices between Chinese students and English students L2, which calls for the comparative study of the types of discourse cohesive devices in Chinese students' (L2) English compositions and native English students' (L1) compositions, analyze the distribution of various types of cohesion devices, and explore the causes for the differences.

To successfully achieve the above research objectives and make the research conclusions not only statistically but also practically significant and universal, two core tasks need to be addressed: firstly, conducting comparative analysis based on a large number of texts, and secondly, utilizing reliable and effective discourse cohesive devices analysis tools. At present, the rapid development of corpus technology makes it a new method used in applied linguistics research. Xu <sup>[2]</sup> holds the view that Corpus linguistics, based on language use, highlights statistical probability, pays attention to methods, and attaches importance to context, which can not only build language theory but also guide language use. A major feature of corpus technology is its ability to effectively process a large number of language texts and quantitatively describe the relevant information in the text. At the same time, with the development of technology, in recent years, some researchers have begun to use Coh-Metrix 3.0 to automatically analyze cohesive devices in texts. This technology can provide 106 analysis indices. Combined with manual analysis, discourse cohesion can be comprehensively examined more precisely.

Therefore, based on the above ideas, this study tends to utilize Coh-Metrix 3.0 and corpus technology to compare the distribution characteristics and differences of discourse cohesive devices in Chinese students' (L2) English compositions and native English students' (L1) English compositions, explore the causes for the differences, and propose solutions to improve the application of discourse cohesive devices in Chinese students' English compositions.

## 2. Literature Review

This chapter is devoted to reviewing relevant literature from four aspects. The first part explains the definition of certain key concepts. The second part presents previous empirical studies concerning cohesive devices in English compositions. The third part reviews current research based on Coh-Metrix. Finally, the last part concludes the limitations of previous studies.

### 2.1 Definition of Cohesion Theory

Cohesion theory is one of the important theories in Systemic functional linguistics. It was proposed by Halliday and Hasan in their book *Cohesion in English* in 1976. Halliday and Hasan <sup>[3]</sup> argued that sentences are the textual mechanism of a discourse, connected by various bonds, and these cohesive bonds are cohesive devices. From a semantic perspective, cohesion is a semantic concept that refers to the meaning of relationships that exist in the text. As an important content of discourse characteristics, discourse cohesion can be divided into two categories based on different cohesive functions: grammatical cohesion and lexical cohesion

Furthermore, grammatical cohesion mainly includes reference, connection, ellipsis, and substitution; Lexical cohesion can be divided into lexical reiteration and lexical collocation. And the way of operationalizing these concepts into quantitative indices will be elaborated in Chapter 3. The following will briefly introduce the basic concepts and content of different grammatical and lexical cohesion.

### 2.1.1 Grammatical Cohesion

Reference includes pronouns and other means to represent the same semantics. Based on different relationships, reference can be subdivided into personal reference, demonstrative reference, and comparative reference. Personal reference functions mainly through Personal pronouns (I, you, he, etc.), subordinate Determiner (my, your, his, etc.), and subordinate pronouns; demonstrative reference is mainly achieved through demonstrative words, such as this, that, the, here, there, etc.; Comparative reference involves comparative meaning and is implemented through words such as less, more, etc. In addition, according to different classifications, reference can also be divided into anaphoric reference and cataphoric reference, which will not elaborate further in this paper due to research purposes.

Ellipsis is a cohesive device that avoids repetition by omitting objects that have already been described in the previous text. The common forms of ellipsis include noun ellipsis, verb ellipsis, and participle ellipsis. For example:

A: Here are two pencils. One is red, and the other is blue.

B: I'd like the red. (the noun of Pencil is omitted)

Substitution and omission have similar functions and are the means of avoiding repetition and connecting contexts. Its core is to replace non-critical words in sentences with substitute words. The discourse cohesion function of substitution is achieved through substitute words (such as one, the same, do, etc.). Common forms of substitution include nominal substitution, verbal substitution, and clausal substitution.

Noun substitutes such as one (s), the same, and the kind are used to achieve Nominal substitution which replaces the components that have already appeared in the previous text. For example:

I have already read this book. Show me another one.

The use of the nominal substitute word "one" instead of "book" makes the composition concise and avoids repetitive expressions.

Verbal substitution is achieved through variants of do for substitution, such as:

A: Does Mary enjoy this trip?

B: Yes, she does.

The word "does" here replaces "enjoy this trip".

So/not is used in sentences to act as a clausal substitution. For example:

A: Do you think he will go to this party?

B: Yes, I think so. /No, I think not.

Connection is one of the most common means of linking contexts. Halliday argues that connection is different from the other three types of grammatical cohesion; Connectives themselves are not the key to cohesion, but rather the logical and semantic connection between the preceding and following text they connect. In *Cohesion in English*, Halliday classified connections into four types: additive, transitional, causal, and temporal connection.

### 2.1.2 Lexical Cohesion

In addition, Halliday's cohesion theory also involves lexical cohesion. Lexical cohesion mainly includes lexical reiteration and lexical collocation. In addition to lexical repetition, lexical repetition

also involves the repetition of synonyms and synonyms, as well as the repetition of synonyms and synonyms. For example:

I think interest is very important. Without interest, we cannot achieve anything. (repetition)

The upsurge in oil prices is unbelievable this year and I think it will soar to the next level. (Recurrence of synonyms, where upsurge and soar are synonyms)

I really like flowers and I often buy a bunch of rose and lily once a week. (Recurrence of synonyms with flowers being synonyms of rose and lily)

Collocation refers to a class of words that appear around the same topic or field, including antonyms and complements. For example:

My dress is old. I want to buy a new one. (Old and new are antonyms, forming an antonymous relationship)

I am going to go upstairs. (The use of upstairs here automatically excludes the use of downstairs, forming a complementary relationship between these two words)

## 2.2 Empirical Studies on cohesion devices of learners' English composition

As one of the important theories of Systemic functional linguistics, scholars at home and abroad have studied discourse cohesive devices from many aspects, such as pragmatics and translation. For discourse cohesion, research mainly focuses on analyzing its functions and types in different discourses. Hessamy<sup>[4]</sup> compared the frequency of the use of cohesive devices in independent and integrated essays; Firdaus *et.al*<sup>[5]</sup> analyzed the textual interpretations which correspond to selected high cohesion texts and explored whether or not they fulfilled the comprehension of reading; Samadian *et.al*<sup>[6]</sup> Studied Iranian intermediate EFL learners' problems in cohesion and coherence of writing performance as well as the extent to which they utilized cohesion and coherence in their writing. Masadeh<sup>[7]</sup> investigated cohesion and coherence in the writings of Saudi undergraduates majoring in English based on corpus and found the common problems of discourse cohesive devices made by Saudi undergraduates. Sutopo<sup>[8]</sup> studied the use of cohesive devices in articles written by graduate students.

Some Chinese researchers also studied cohesion devices of learners' English composition. Some proposed the idea of improving English composition through discourse cohesion by analyzing common errors in discourse cohesion. Hong<sup>[9]</sup> studied the local coherence characteristics of English composition discourse among non-English major students; Wang<sup>[10]</sup> examined the cohesion of students' oral English from a cognitive pragmatic perspective by summarizing the phenomenon of ellipsis in Chinese and English students' oral English; Zhang<sup>[11]</sup> explored the interaction between fluency and discourse coherence.

To sum up, current research on discourse cohesive devices in students' composition mainly focuses on exploring the types and characteristics of students' use of discourse cohesion, as well as analyzing how to better enhance students' use of discourse cohesion in composition.

## 2.3 The Current Study on Cohesion Devices based on Coh-Metrix

For the study of discourse cohesion in texts, some researchers rely on manual annotation and corpus-related software for analysis. However, this method may lead to a high workload and is difficult to handle a large amount of corpus, which may cause the lack of credibility of research results. Considering the above issues, in recent years, some researchers have started using Coh-Metrix for discourse cohesion analysis. Coh-Metrix 3.0 was developed by Danielle McNamara of the University of Memphis in the United States. It can analyze 106 language indicators, including word length, sentence length, word frequency, syntactic complexity, readability indices, etc. The core function of this software is to measure the cohesive features of a text. In recent years, research based on Coh-

Metrix has shown a growing trend. The relevant research involves the following three types:

Firstly, based on the core discourse cohesion analysis function of Coh-Metrix, some researchers have focused on studying cohesion in discourse. For example, Crossley<sup>[12]</sup> used the Coh-Metrix technique to analyze whether some indices of cohesive devices affect teachers' ratings of essays; Li et al.<sup>[13]</sup> analyzed the cohesive devices in the introduction part of academic papers by Chinese English learners (L2) based on Coh-Metrix and found that L2 has significant differences in the cohesive devices in the introduction compared to L1, with problems such as overuse or underuse.

Secondly, relevant studies analyzed the readability of text through Coh-Metrix. For example, McNamara<sup>[14]</sup> explored how to use Coh-Metrix to provide advanced measurement methods for text and discourse language features, to evaluate the features of the text that make it easy or difficult to read; Chen<sup>[15]</sup> conducted a study on the difficulty level of college English textbooks based on Coh-Metrix.

Thirdly, Coh-Metrix is also used to explore the characteristics of composition language and its prediction of composition quality, and related research has played a promoting role in the development of machine-automated scoring. For example, Latifi<sup>[16]</sup> analyzed how to use Coh-Metrix indices to effectively and accurately grade student essays automatically; McNamara conducted a study on 120 compositions by native English speakers and found that high-quality compositions have more complex syntax, diverse vocabulary, and use more obscure vocabulary, while cohesion indices are not directly related to composition quality; Wang<sup>[17]</sup> evaluated the use of Coh-Metrix indices as an approach for defining the characteristics of essays related to their scoring difficulty and finding that the subjective interpretation of the underlying unfounded scale can be enhanced by examining essay characteristics with Coh-Metrix indices; Liang<sup>[18]</sup> found through research that the Coh-Metrix indices of local cohesion and overall cohesion are significantly correlated with EFL essay performance, and high scoring essays use better overall cohesion methods, while low scoring essays rely more on local cohesion methods.

In conclusion, with the iterative updates of Coh-Metrix, more and more researchers are considering it as a reliable tool for discourse cohesion analysis and have conducted many studies relying on Coh-Metrix. Based on this, this study tends to adopt Coh-Metrix to conduct this research.

## 2.4 Limitations of the Previous Studies

Throughout previous studies, although there has been research on grammatical cohesion, there are still shortcomings in related research: 1) Some scholars' research only focused on one type of grammatical cohesion device, without examining various grammatical cohesion; 2) Part of the research is limited to traditional text analysis, only using manual analysis and data processing methods, which makes it difficult to process a large amount of corpus; 3) The relevant research mainly focused on the grammatical coherence of English compositions by middle school students, while there is a lack of research on the grammatical coherence of English compositions by undergraduate students.

In addition, it should be noted that some current studies have focused on lexical cohesion in discourse cohesion, namely the phenomenon of lexical repetition and lexical co-occurrence. For example, Alotaibi<sup>[19]</sup> explored the relationship between lexical cohesion and essay quality by analyzing the role of lexical cohesion; Sidabutar<sup>[20]</sup> analyzes the proportion of different lexical cohesion by analyzing students' written language, and puts forward relevant teaching suggestions based on it; Liu<sup>[21]</sup>, based on lexical cohesion theory, extracted some language features related to written language coherence of Chinese English learners. At the same time, the above research findings have also promoted composition teaching, such as encouraging students to expand their vocabulary to achieve synonym substitution in their compositions. However, there are still not many

related studies, and at the same time, relevant studies have not analyzed the differences in cohesive devices between L1 and L2 from a comparative perspective.

Given this, this study is based on Coh-Metrix and focuses on exploring the distribution of different types of discourse cohesion in L1 and L2 English compositions. Through this study, the researcher hopes to raise the attention of teachers to discourse cohesive devices, encourage more researchers to analyze discourse cohesion, and attract teachers to pay attention to discourse cohesion as a way to make students improve the coherence of discourse in students' composition and promote their comprehensive English application ability.

### 3. Research Design

#### 3.1 Research questions

To explore the use of discourse cohesion in composition by L2, this study focuses on the following two issues:

- 1) Is there a significant difference in the use of cohesive devices among L1 and L2 compositions?
- 2) What are the causes for the significant differences in cohesive devices among L1 and L2 compositions?

#### 3.2 Research Tools

Coh-Metrix 3.0 is used in this study as a tool to examine the cohesive features of composition. It provides a total of 106 measurement indicators. This study selects 21 indices related to discourse cohesion and divides them into the following two major categories based on Halliday's (1976) cohesion theory (As shown in Table 1).

Taking the research purpose and the indices that can be analyzed by Coh-Metrix into consideration, this study only focuses on personal reference and connection, both of which are the most used grammatical cohesive devices in discourse according to Halliday (1976). There are 11 indices in grammatical cohesion in which six indices (WRDPRO, WRDPRO, WRDPRP1p, WRDPRP2, WRDPRP3s, and WRDPRP3p) are used to evaluate the feature of reference in compositions; and five indices (CNCAil, CNCCaus, CNCLogic, CNCADC, and CNCTemp) which can be used to analyze discourse conjunction such as causal, logical, adversative, and contrastive conjunction are selected.

14 indices are selected for the analysis of lexical cohesion. Since lexical reiteration is the main type of lexical cohesion and is directly related to discourse cohesion (Halliday, 1976), this research focuses on lexical reiteration. Indices of CRFNO1, CRFAO1, CRFSO1, CRFNOa, CRFAOa, CRFSOa, CRFCWO1, and CRFCWOa are used to analyze the overlap of noun, argument, stem, and content words, which corresponds to the concept of lexical reiteration. LSASS1 and SMCAUSwn are used to determine the degree of semantic similarity between sentences, and can also be used to measure the degree of lexical repetition; The indices of polysemy and hypernymy can also be used to determine lexical reiteration.

It is worth noting that Coh-Metrix cannot directly analyze ellipsis in grammatical cohesion, alternative means, and lexical collocation in lexical cohesion. Therefore, to ensure the reliability of research data and the consistency of research methods, this study will not conduct research and analysis on the above cohesive means for the time being. Even if certain types of discourse cohesion are excluded from this study, this study can still effectively evaluate the use of cohesion devices in L1 and L2 compositions due to the fact that the types of discourse cohesion that this study focuses on account for the majority of discourse cohesion devices, and this study ensures that the analyzed cohesion devices can be analyzed through Coh Metrix, which ensures consistency in research methods.

Table 1: Indices for Measuring Cohesive Devices in Coh-Metrix 3.0

Types of Cohesive Devices	Subtype of Cohesive Devices	Indices	Abbreviation
Grammatical cohesion	Reference	Pronoun Incidence	WRDPRO
		first person, single form	WRDPRO
		first person, plural form	WRDPRP1p
		second person	WRDPRP2
		third person, single form	WRDPRP3s
		third person, plural form	WRDPRP3p
	Conjunction	All connectives incidence	CNCAll
		Causal connectives incidence	CNCCaus
		Logical connectives incidence	CNCLogic
		Adversative and contrastive connectives incidence	CNCADC
Temporal connectives incidence		CNCTemp	
Lexical Cohesion	Lexical reiteration	Noun overlap, adjacent sentences, binary, mean	CRFNO1
		Argument overlap, adjacent sentences, binary, mean	CRFAO1
		Stem overlap, adjacent sentences, binary, mean	CRFSO1
		Content word overlap, adjacent sentences, proportional, mean	CRFNOa
		Noun overlap, all sentences, binary, mean	CRFAOa
		Argument overlap, all sentences, binary, mean	CRFSOa
		Stem overlap, all sentences, binary, mean	CRFCWO1
		Content word overlap, all sentences, proportional, mean	CRFCWOa
		LSA overlap, adjacent paragraphs, mean	LSASS1
		WordNet verb overlap	SMCAUSwn
		Polysemy for content words, mean	WRDPOLc
		Hypernymy for nouns, mean	WRDHYPn
		Hypernymy for verbs, mean	WRDHYPv
		Hypernymy for nouns and verbs, mean	WRDHYPnv

### 3.3 Research subjects and corpus sources

Table 2: Basic Information of the Two Corpora Used in this Research

CORPUS NAME	CORPUS TYPE	NUMBER OF SELECTED FILES	NUMBER OF TOKENS
CHINESE COLLEGE ENGLISH COMPOSITION CORPUS	Monolingual Corpus	136	26,028
LOCNESS CORPUS	Monolingual Corpus	136	27,456

The corpus of Chinese college English learners used in this study is a small self-built corpus that comes from the compositions of 136 freshmen from Southwest Jiaotong University in 2022. The collection of essays has been approved by the participants. The type of composition in this research is argumentative essay, with a word count requirement of between 150 and 200 words. Table 2 provides descriptive statistical information for the corpus. The reference English native language corpus is a LOCNESS corpus composed of argumentative papers by college students from the UK and the US.

### 3.4 Research Methods

This study adopts a combination of quantitative analysis and qualitative research based on semi-structured interviews.

Firstly, the researchers randomly selected 50 articles respectively from two corpora, transcribed them into .txt format, and input them into Coh-Metrix 3.0 analysis to obtain data. The data analysis stage includes descriptive statistical analysis and inferential statistical analysis. After obtaining the data, the researcher first integrates the data and then conducts descriptive statistical analysis. Since the mean values of different cohesive devices in L1 and L2 compositions are the data basis for answering the first research question, the researchers summarized the mean values of different discourse cohesive devices and found the distribution characteristics of different types of discourse cohesive devices in L1 and L2 compositions; At the same time, researchers used SPSS 26.0 to conduct independent sample t-tests on the data, analyzing whether there are significant differences in the use of discourse cohesive devices between L2 and L1 for further discussion and analysis.

To further explore the reasons for the use of different types of cohesion in Chinese students' English compositions, explore the causes for the significant differences in the types of cohesion in L1 and L2 English composition, and answer the second research question, this study, based on the preliminary analysis of the corpus, randomly selected five students (randomly labeled S1, S2, S3, S4, and S5) from the composition writers of the Chinese college students' English composition corpus for semi-structured interviews. The interview is conducted online, using one-to-one individual communication, and the interview time is approximately 10 to 20 minutes. The interview outline can be found in the appendix, and the interview content will be appropriately modified based on the use of cohesion in the student's essay. The interview content was manually transcribed and then translated into English by the researcher and then organized and studied.

## 4. Results

By conducting independent sample t-tests on the data detected by Coh-Metrix, Table 3 shows the t-values and p-values of the cohesion indices in L1 and L2 English compositions.

As shown in Table 3, out of the 21 indices, a total of 12 showed significant differences, which indicates a difference in the use of cohesion devices among the compositions of L1 and L2.

In terms of grammatical cohesion, L2 is higher than L1 in terms of the Personal Pronoun index. Among them, there are significant differences between L1 and L2 in the use of the singular number of the first Personal pronoun and the second Personal pronoun, while there is no significant difference in the use of the Third-person pronoun, which indicates that L2's composition has an obvious tendency to use too many Personal pronouns; In the indices of conjunctions, there was no significant difference between L1 and L2 in their compositions. Only in the use of causal conjunctions, L1 was significantly higher than L2.

In terms of lexical cohesion, L1 is significantly higher than L2 in the use of noun word overlap, argument overlap, and stem overlap, and L2 is significantly higher than L1 in content word overlap. At the same time, L1 is significantly higher than L2 in the index of content word polysemy, noun



synonyms, and verb synonyms. The above data indicate that L2 students have significant differences in vocabulary cohesion compared with native speakers.

Table 3: T-values and p-values of cohesion indices in L1 and L2 English compositions

Cohesion Indices	Measurement Indices	Means of L2	Means of L1	t-value	p-value
Grammatical Cohesion	WRDPRO	14.630	13.050	0.352	.025*
	WRDPR1s	18.530	10.690	1.539	.029*
	WRDPR1p	6.230	3.900	0.955	0.342
	WRDPR2	9.120	6.060	0.949	.046*
	WRDPR3s	14.507	13.586	0.251	0.803
	WRDPR3p	34.219	29.424	1.393	0.168
	CNCAII	45.877	45.548	0.085	0.933
	CNCCaus	16.286	21.111	2.102	.038*
	CNCLogic	19.41	18.448	0.412	0.681
	CNCADC	0.326	0.376	1.302	0.196
CNCTemp	0.524	0.536	0.307	0.759	
Lexical Cohesion	CRFNOa	0.336	0.394	1.648	.003*
	CRFAOa	0.095	0.151	5.364	.000*
	CRFSOa	0.076	0.113	4.354	.000*
	CRFCWOa	88.983	60.259	3.312	.001*
	LSASS1	94.163	93.477	0.129	0.897
	SMCAUSwn	0.561	0.555	0.262	0.794
	WRDPOLc	4.523	4.676	3.393	.001*
	WRDHYPn	6.303	6.633	3.076	.003*
	WRDHYPv	1.449	1.497	1.344	.082*
	WRDHYPnv	1.528	1.766	5.446	.000*

(The data marked with an asterisk indicates a significant difference in this index)

## 5. Discussion

### 5.1 The Influence of Mother Tongue

The negative transfer is the phenomenon that linguistic rules and grammatical structures of people's mother tongue tend to impede the acquisition of a second language. Lin<sup>[22]</sup> argues that when the structure of the mother tongue is inconsistent with the targeted language, it will hinder the process of acquisition. In Chinese students' compositions, incorrect use of cohesive devices can be detected such as the overuse of "because" and "so". This can be explained by the negative influence of the Chinese. Also, since the discourse written in Chinese tends to use the personal pronoun, it to some extent affects Chinese students, making them tend to overuse cohesive devices of personal reference in composition, which is consistent with the data detected by Coh- Metrix.

Most college students in China still communicate mainly in Chinese, except for limited English class time, so Chinese has a greater influence on their acquisition of English, which makes students accustomed to using the same expressions as in Chinese when writing. According to the statistical data, due to the influence of Chinese, L2 students are significantly more likely to use personal pronouns than L1 students, which causes a tendency to favor the overuse of personal pronouns. This analysis is corroborated by the following data from the interviews of the two interviewees:

"I feel that Chinese has a significant impact on my composition... Sometimes when I write, I get

used to using the way I use in Chinese to achieve cohesion in my compositions, and I feel that the way I write is not wrong.” (S1)

“The influence of Chinese is one reason. Because I use cohesive devices in Chinese, so when I express a similar meaning, I tend to apply the Chinese way into my English compositions.” (S2)

## 5.2 Students' Motivation

We found through semi-structured interviews that students' attitude on exams, composition, and the use of cohesive devices is also the factor that leads to significant differences in the use of cohesive devices between L1 and L2. The results of semi-structured interviews indicate that due to limitations in composition time and the inability to form a sense of cohesion, some students rarely consider using grammatical and lexical cohesion devices when writing. At the same time, some students consciously reduce the use of certain cohesive devices to achieve good grades and avoid errors in composition, to write "correct" composition. For example:

“I haven't paid much attention to this issue. I only use conjunctions, and I think that's enough... I don't think much about it during exams...Also when I use ‘but’ too many times I will probably use ‘however’.” (S3)

“Time is too tight during examination and sometimes I can't even finish composition...and that's why I don't want to consider the use of cohesive devices. I have to spend much time thinking how to fit some cohesive devices into my compositions and also if I use it incorrectly, my score will be influenced, which is totally unacceptable.” (S5)

## 5.3 Teachers' Emphasis on Cohesion in Different Discourses

In some Chinese universities, teachers attach great importance to explaining words, sentence structures, grammar, texts, and exam techniques, but rarely provide specialized explanations on the cohesive devices. This leads to students' unclear understanding of cohesive devices and difficulty in applying them to composition. Students lack systematic learning and understanding of discourse cohesion, which may result in a tendency to transfer relevant discourse cohesive devices from their mother tongue in English composition. Through data analysis, some simple discourse cohesive devices such as to sum up, however, and therefore, et.al were detected in most of the compositions written by L2 but it has been found through semi-structured interviews that students, even if knowing how to use cohesive devices in compositions, still lack understanding toward the essence of cohesion and also reported a lack of emphasis on cohesive devices by teachers. The insufficient attention given by teachers to cohesive devices has led some students to believe that cohesive devices are not important and cannot be used correctly and flexibly in composition to achieve discourse cohesion. For example, interviewee S2 reported that:

“The teacher usually talks about words, grammar, and problem-solving skills... memorize some writing templates and master some advanced conjunctions. As for cohesive devices, my teacher always says that we should use 'because', 'but', and 'how' in writings but except these conjunctive words, I have no idea about other ways to make my article cohesive, and it seems that my teacher doesn't pay much attention to it.” (S2)

At the same time, through semi-structured interviews, the study found that Chinese teachers tend to teach lexical diversity of notional words when teaching lexical cohesion. Students are required to use various expressions of the same meaning when writing composition in English, which is consistent with the results of statistical data. The data of Chinese students on the overlap of notional words is significantly higher than L1. However, in terms of the use of noun reduplication, argument reduplication, and stem reduplication, L1's data is significantly higher than L2's, which to some extent indicates that Chinese English teachers have a type preference in teaching discourse cohesion.

“For example, the teacher said that when you write, you need to use some “advanced” synonyms more frequently. For example, my teacher says you need to take turns using importance and significance. However, I only know about synonym replacement and some other vocabulary connection methods. To be honest, I really don't know other way to achieve cohesion in composition.”(S5)

## 5.4 Influence of Textbooks

Alifia <sup>[23]</sup> concluded through research and analysis that the content of cohesion devices in English textbooks can effectively enhance students' ability to apply grammatical and lexical coherence in composition. In recent years, the content of cohesive devices in composition has also begun to appear in college English textbooks in the Chinese Mainland. However, compared with the presentation of vocabulary, grammar, culture, and other contents, the way cohesive devices are presented in most textbooks is relatively simple with insufficient specific examples and practices. Some textbooks even only present some cohesive words, without teaching the use of cohesive devices and the internal logic of cohesion. Although the content of cohesive devices in the textbook can to some extent promote students to imitate, it fails to fully demonstrate the essence and logic behind the use of discourse cohesion devices, resulting in a lack of deep understanding of cohesive devices and susceptibility to negative transfer from their mother tongue. Therefore, while there is no significant difference in statistical data between L1 and L2, the interview results of the study support the above analysis:

“The textbook only describes some connectives, but I'm not sure why I use these words. I only use the words' because 'and 'therefore', and I don't use any other words.” (S4)

At the same time, some English textbooks may have a certain emphasis on the accuracy of language forms and grammatical structures, while ignoring the meaning and context of the language, which may affect Chinese students' English Composition style and the use of cohesive devices; In addition, the examples and exercises in textbooks may lack diversity and cannot fully reflect the actual use of the English language, making it difficult for students to master the latest language cohesion skills and expression methods. The interview result also supports the above analysis:

“I feel that the focus on cohesion in the textbooks is not enough. At least, I believe that the textbooks I encounter do not have specialized sections to present cohesive devices, and I cannot learn the deep logic of using cohesive devices from simple imitation.” (S3)

## 6. Conclusion

Cohesion is one of the important means of connecting semantics in English compositions, and it is also one of the important criteria for reflecting the thinking ability of composition authors and evaluating students' comprehensive language application ability. The data analysis and discussion in this article indicate that there is a significant difference in the use of discourse cohesive devices in Chinese college students' English compositions. On the one hand, this phenomenon is attributed to the negative transfer influence of students' mother tongue; On the other hand, it is also due to the lack of systematic explanation of cohesive devices in textbooks, insufficient emphasis on cohesive devices by teachers, and students' motivation.

We argue that in order to solve the above problems, the most important thing is the action of teachers and the transformation of textbook production models.

Teachers should not only pay attention to the influence of students' mother tongue on their mastery of English device cohesive devices, but they should also focus on explaining the essence of relevant device cohesions in composition teaching and transform students' negative motivations towards cohesive devices to promote students' mastery and appropriate use of cohesive devices; In addition, when explaining the text, teachers should also guide students to pay attention to the cohesive devices

in the text. Tahmasbi et al <sup>[24]</sup> found that appropriate technology such as applying mobile-assisted flipped learning can also ignite students' motivation to learn cohesive devices. Therefore, teachers can combine some new technology to design targeted instructional materials, activities, and assessments to facilitate students' development of cohesive composition skills and can provide feedback that addresses cohesion-related errors, fostering students' improvement in the use of cohesive devices in English composition.

College English textbook writers should include specialized practices on English cohesive devices in their textbooks as a way to allow students to learn and understand how to apply cohesive devices in English composition and improve their composition skills through imitation exercises. We argue that the following four textbook modes can promote students' composition coherence skills:

1) Designing targeted exercises. Textbooks can be designed with targeted exercises on cohesion devices in English compositions such as making students rewrite a paragraph to make it more cohesive; Alternatively, some fill-in-the-blank questions can also allow students to fill in appropriate discourse based on the context to exercise their cohesive skills.

2) Providing composition guidance. Textbooks can include guidance on English composition, such as how to use cohesive devices, and how to handle transitions between paragraphs. These guidelines and methods can help students become more proficient in composition and improve their ability to use cohesive devices.

3) Focusing on practice. Textbooks should provide sufficient practical opportunities, such as designing some English essay questions to enable students to apply the knowledge they have learned in practical composition and improve their composition skills and the application of cohesive devices.

4) Emphasizing language accuracy. Textbooks can emphasize the importance of language accuracy and provide relevant exercises and inspection mechanisms to ensure that students' composition language is accurate, appropriate, and coherent.

This research can contribute to theoretical and practical development. For theoretical development, cohesion plays a pivotal role in achieving communicative competence, as it ensures the clarity and coherence of the written text. By examining the types of cohesive devices employed by L1 and L2, this research contributes to the theoretical framework of language acquisition and proficiency development.

For practical development, understanding the differences in cohesion between L1 and L2 can promote students to attach importance to grammatical and lexical cohesion and their efforts should be made to use relevant cohesive devices in daily composition to enhance their familiarity with the way to make their compositions coherent, thereby boosting their academic performance and communication skills; Also, education policy-makers can utilize the findings of this analysis to inform curriculum development and pedagogical reforms. By recognizing the importance of discourse cohesive devices in English composition, policymakers can allocate resources to support teacher training and the integration of cohesive composition instruction across various educational levels. This, in turn, promotes the overall quality of English education and enhances students' English proficiency.

## References

- [1] Huang, Jian. 2016. *On the Application of Textual Cohesion in Teaching and Learning English Writing*. *Journal of Jiangxi Normal University (Philosophy and Social Sciences Edition)*, 49(05):135-139.
- [2] Xu Jiajin. 2021. *Corpora and Discourse Studies*. Beijing: Foreign Language Teaching and Research Press.
- [3] Halliday, M. A. K. & Hasan. 1976. *Cohesion in English*. London: Longman, 1976.
- [4] Hessamy, G., & Hamed, S. 2013. *A comparison of the use of cohesive devices in EFL Learners' performance on independent vs. integrated writing tasks*. *Study in English Language Teaching*, 1(1), p 121.
- [5] Firdaus, A. F., Hamzah, Y., & Nurhalimah, F. 2022. *Cohesion Analysis on Texts for Reading Comprehension*. *BELTIC2018-1st Bandung English Language Teaching International Conference*, 78-84.

- [6] Samadian S, Mohseny. 2019. *Analysis of cohesion and coherence in writing performance of Iranian intermediate EFL learners [J]. Issues in Language Teaching*, 8(2): 213-242.
- [7] Masadeh T.S. 2019. *Cohesion and coherence in the writings of Saudi undergraduates majoring in English[J]. Journal of Social Sciences and Humanities*, 5(3): 200-208.
- [8] Sutopo D, Faridi A. 2020. *The Use of Cohesive Devices in English Education Journal Articles Written by Graduate Students of UNNES [J]. English Education Journal*, 10(2): 208-213.
- [9] Hong, Ming, Xu, Yanxia. 2016. *A Study on Local Coherence Characteristics of Non-English Majors' English Essays. Technology Enhanced Foreign Language Education*, (02):59-64.
- [10] Wang, Meng, Lin, Zhengjun. 2016. *A Research on Ellipsis in English Majors' Spoken English from a Cognitive and Pragmatic Perspective. Foreign Languages and Their Teaching*, (06):76-84+149-150.
- [11] Zhang, Chao, Liang, Wenhua. 2022. *Multi-dimensional interactive patterns of the L2 writing textual fluency and discourse cohesion [J]. Foreign Language Teaching and Research*, 54(04):557-568+639-640.
- [12] Crossley, S. A., Weston, J. L., McLain Sullivan, S. T., & McNamara, D. S. 2011. *The Development of Writing Proficiency as a Function of Grade Level: A Linguistic Analysis. Written Communication*, 28(3), 282–311.
- [13] Li Jiaqi, Ma Jianju, Zhumu Langma. 2022. *A Coh-Matrix Study on Cohesive Devices in Chinese English Learners' Research Article Introductions. Modern Linguistics*, 10(12): 2989-3000.
- [14] McNamara, D. S., Louwse, M. M., McCarthy, P. M. & Graesser, A. C. 2010. *Coh-Matrix: Capturing linguistic features of cohesion [J]. Discourse Processes*, (4): 292-330.
- [15] Chen, Annie, Guo, Aiping. 2019. *A Study on the Difficulty of College English Textbooks Based on Coh Matrix: Taking the Third Edition of New Horizon College English as an Example. Educational Theory and Practice*, 39 (06): 46-48.
- [16] Latifi S, Gierl M. 2021. *Automated scoring of junior and senior high essays using Coh-Matrix features: Implications for large-scale language testing [J]. Language Testing*, 38(1): 62-85.
- [17] Wang J, Engelhard Jr G, Combs T. *Exploring difficult-to-score essays with a hyperbolic cosine accuracy model and Coh-Matrix indices [J]. The Journal of Experimental Education*, 2023, 91(1): 125-144.
- [18] Liang Maocheng 2006. *A Study on the Coherence of Learner Written Discourse. Modern Foreign Languages*, (3): 284-292.
- [19] Alotaibi, H. 2015. *The role of lexical cohesion in writing quality. International Journal of Applied Linguistics and English Literature*, 4(1), 261-269.
- [20] Sidabutar, U. 2021. *An analysis of lexical cohesion on the students' composition. JETAL: Journal of English Teaching & Applied Linguistics*, 2(2), 62-67.
- [21] Liu, Guobing. 2012. *A Study on the Evaluation of Coherence in English Learners' Written Discourse from the Perspective of Lexical Cohesion Theory. Foreign Language Education in China*, 9(02):33-43+88.
- [22] Lin, Zhengjun, Liu, Yongbing. 2012. *English grammar teaching: A perspective of Construction Grammar. Foreign Language Learning Theory and Practice*, (06):76-84+149-150.
- [23] Alifia Junita Cendraa Sari, Januarius Mujiyanto, Dwi Rukmini. 2022. *Assessing the use of cohesive devices in reading texts of English textbooks [J]. English Education Journal* 12(2): 264-273.
- [24] Tahmasbi S, Rabani Ebrahimi Pour K. 2023. *The Effect of Mobile-Assisted Flipped Learning on Iranian EFL Learners' Cohesive Devices Improvement in Writing [J]. International Journal of Research in English Education*, 8(2): 82-95.