

Article

Features of Connector Use in Chinese College English Students' Writing

Nairui Chen

Northern Arizona University, USA

Wen Zhao*

Jinan University, China

Received: 15 September 2021/Accepted: 31 January 2022/Published: 30 March 2022

Abstract

This chapter introduces a study which investigates the features of connector use in Chinese non-English major college students' writing based on a home-developed learner corpus (3,908,816 tokens). The research findings reveal that Chinese college students use more one-word connectors than multiword ones to express the meaning of enumeration and addition, and that connectors are usually placed in simple declarative sentence order with inverted sentence order or complex sentence patterns rarely used. With reference to the English Grammar Profile (EGP), the criterial features of grammatical use based on the CEFR, it can be found that students' use of connectors spreads across levels, with most connectors clustered on lower levels. It is expected that the current empirical study can inform scale descriptors and criterial features of the Cohesion Competence of the China's Standards of English Language Ability (CSE).

Keywords

Connectors, English grammar profile, China's standards of English language ability, learner corpus

1 Introduction

Writing has long been considered a manifestation of critical thinking and language ability. Constructing cohesive texts, which are cohesive and coherent in meaning, is important for writers. According to Halliday and Hasan (1976), cohesion is part of the text-forming component in the linguistic system realized through grammar and vocabulary while coherence is associated with the overall continuity. Connectors, also known as a type of cohesive device, conjuncts, and cohesive ties (Halliday & Hasan, 1976; Milton & Tsang, 1993; Quirk & Crystal, 1985), play a significant role in achieving cohesion and coherence in both L1 English writing (Altenberg & Tapper, 1998) and in EFL academic writing (Li, 2011; Yang & Sun, 2012; Zhao, 2003). Mastery of the lexical-grammatical features of connectors can help both native speakers and EFL learners improve their English writing quality (Biber et al., 1999).

*Corresponding author. Email: wenzhao@jnu.edu.cn

2 Literature Review

The term “connectors” can be defined as words or phrases whose function is to show some logical relationship between two or among more basic sentences (Celce-Murcia et al., 1983). In terms of categorization, Halliday & Hasan (1976) classify connectors into four categories: *additive*, *causal*, *adversative* and *temporal*, with the categories further elaborated into seven subcategories. Biber and his colleagues (1999) further summarize the meaning-based classification into *enumeration* and *addition*, *result* or *inference*, *contrast* or *concession*, *summation*, *apposition* or *transition*, and classify connectors according to their parts of speech into three types: *conjunctions* (e.g., and, but, or, because, either ... or), *prepositional phrases* (e.g., in that), and *adverbial phrases* (e.g., never, hardly ... when), which provides a guideline for connector studies.

Studies have been conducted to investigate connectors used in both L1 and L2 learners’ English writing, with their foci mainly on two aspects. One involves studies on patterns, similarities, or differences of connectors used by native speakers and L2 learners. The research studies disclose that EFL learners use connectors more frequently than other grammatical categories because of their language background. For example, compared with L1 native speakers, Chinese EFL learners overuse connectors expressing *listing* and *enumeration*, and they underuse connectors for ellipsis (Liu & Braine, 2005; Yang & Sun, 2012; Zhang, 2000). Milton & Tsang (1993) point out that EFL learners, including Chinese EFL learners, have apparent difficulty in making use of connectors to compose a coherent and cohesive English written text. As for learners with other L1 background, Kang (2005) made a comparative study between Korean EFL learners and American native speakers, and found that Korean students overused connectors expressing *inference*, rather than connectors expressing *Listing* and *enumeration*. Blagoeva’s (2004) study of Bulgarian EFL learners has similar research findings as Kang’s (2005).

The other aspect centres on the issue of whether using cohesive devices properly can lead to better writing quality. Some researchers claim that connectors have no effect on writing quality (e.g., Granger & Tyson, 1996; Jin, 2001; Liu & Braine, 2005; Norment, 1994; Witte & Faigley, 1981; Song & Xia, 2002; Yang & Sun, 2012), while others argue that the appropriate use of connectors is significantly correlated with the quality of writing (e.g., Crossley & McNamara, 2016; Khalil, 1989; Medve & Takač, 2013; Zhang, 2000) (see Table 1).

Table 1

Studies on the Relationship Between Connectors and Writing Quality

Research study	Sample	Relationship between connectors & writing quality
Witte & Faigley, 1981	10 L1 college learners	Yes
Norment, 1994	30 EFL college learners (L1: Chinese)	Yes
Granger & Tyson, 1996	over 200 EFL learners (L1: Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Tswana, Turkish)	Yes
Jin, 2001	3 intermediate + 3 advanced EFL college learners (L1: Chinese)	Yes
Song & Xia, 2002	364 EFL college learners (L1: Chinese)	Yes
Liu & Braine, 2005	50 EFL college learners (L1: Chinese)	Yes
Yang & Sun, 2012	30 intermediate + 30 advanced EFL college learners (L1: Chinese)	Yes
Crossley & McNamara, 2016	184 EFL college learners	Yes
Khalil, 1989	20 EFL college learners (L1: Arabic)	No
Karasi, 1994	135 EFL college learners (L1: Chinese)	No
Zhang, 2000	111 EFL college learners (L1: Chinese)	No
Todd et al., 2007	8 EFL college learners (L1: Thai)	No
Medve & Takač, 2013	30 EFL college learners (L1: Croatian)	No

Table 1 indicates that there is no consensus agreement in terms of the relationship between connectors and writing quality. The variation may result from the inconsistent assessment and the size of sampling. The sample size varies from 8 participants (e.g., Jin, 2001; Todd et al., 2007) to more than 300 (e.g., Granger & Tyson, 1996; Song & Xia, 2002), and the language background of learners varies from English as their first language to English as a second or foreign language. Multiple research methods are adopted in the studies from quantitative, qualitative, mixed-methods to corpus-based or corpus-driven methods, from different theoretical perspectives.

In view of connector distribution across language scales, the Common European Framework of Reference for Languages (CEFR) and its companion volume (Council of Europe, 2001, 2018) give qualitative description in the sections of ‘Grammatical Accuracy’ (2018, p. 132) and ‘Coherence and Cohesion’ (2018, p. 141) across CEFR’s 3 stages and 6 levels (i.e., Basic user (A1-A2), Independent user (B1-B2), Proficient user (C1-C2)). The descriptors cover the following aspects: degree of control, prominent mistakes, linking elements, and cohesive devices at the level of the sentence/utterance and complete text (see Table 2).

Table 2

Connector Use Across CEFR’s Levels

	Grammatical Accuracy	Coherence and Cohesion
C2	Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g., in forward planning, in monitoring others’ reactions).	Can create coherent and cohesive text making full and appropriate use of a variety of organizational patterns and a wide range of cohesive devices.
C1	Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot.	Can produce clear, smoothly flowing, well-structured language, showing controlled use of organizational patterns, connectors and cohesive devices. Can produce well-organized, coherent text, using a variety of cohesive devices and organizational patterns.
B2	Good grammatical control; occasional “slips” or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect. Shows a relatively high degree of grammatical control. Does not make mistakes which lead to misunderstanding. Has a good command of simple language structures and some complex grammatical forms, although they tend to use complex structures rigidly with some inaccuracy.	Can use a variety of linking expression efficiently to mark clearly the relationships between ideas. Can use a limited number of cohesive device to link their utterances into clear, coherent discourse, though there may be some “jumpiness” in a long contribution. Can produce text that is generally well-organized and coherent, using a range of linking expressions and cohesive devices. Can structure longer texts in clear, logical paragraphs.
B1	Communicates with reasonable accuracy in familiar contexts; generally good control, though with noticeable mother-tongue influence. Errors occur, but it is clear what they are trying to express. Uses reasonably accurately a repertoire of frequently used “routines” and patterns associated with more predictable situations.	Can introduce a counter-argument in a simple discursive text (e.g., with “however”). Can link a series of shorter, discrete simple elements into a connected, linear sequence of points. Can form longer sentences and link them together using a limited number of cohesive devices, e.g., in a story. Can make simple, logical paragraph breaks in a longer text.
A2	Uses some simple structures correctly, but still systematically makes basic mistakes; nevertheless, it is usually clear what they are trying to say.	Can use the most frequently occurring connectors to link simple sentences in order to tell a story or describe something as a simple list of points. Can link groups of words/signs with simple connectors (e.g., “and”, “but” and “because”).
A1	Shows only limited control of a few simple grammatical structures and sentence patterns in a learnt repertoire.	Can link words/signs or groups of words/signs with very basic linear connectors (e.g. “and” or “then”).

The CEFR, as a concertina-like reference tool, has been criticized for the opaqueness of some descriptors to be applied (North, 2007). In order to give a clearer description of the distinctive lexical and grammar features that learner demonstrate across language scales, the English Profile (EP) program was launched after 2012. A main focus of the EP is the identification of distinguishing criterial features, one that distinguishes higher levels from lower levels (Hawkins & Filipović, 2012). According to Salamoura and Saville (2010), criterial features are linguistic properties from all aspects of language (phonology, morphology, etc.) which can serve as a basis for the estimation of a learner's proficiency level. The features are characteristic and indicative of language proficiency at each of the six levels in the CEFR.

On the basis of the Cambridge Learner Corpus (CLC), the EP aimed to identify criterial features to make a level distinct from other ones. The criterial features of the EP consist of two linguistic properties: lexical as well as grammatical, which are specified in the English Vocabulary Profile (EVP) (vocabulary.englishprofile.org) and the English Grammar Profile (EGP) (grammar.englishprofile.org) respectively (Hawkins & Filipović, 2012). In the EGP, the criterial grammatical features of cohesive competence are mostly conjunctions and adverbs that are used in multiple sentence patterns with distinctive features. Descriptors of cohesion in the CEFR and examples of criterial features of connectors in the EGP are shown in Table 3.

Table 3

Criterial Features of Connectors in the EGP Across CEFR Levels

CEFR	EGP	
Stages	Levels	
	Criterial features of Connectors	
Proficient user	C2	neither/nor + inverted auxiliary/'be' + subject + a previous related negative clause; and yet + sentences; hardly + had + clauses + when; in that + clauses/sentence
	C1	(at the beginning of a sentence) Whatever, etc. + clauses/sentence; not only + phrases/clauses + but also; not only + inverted auxiliary 'do' + but also; yet + phrases/clauses; either + more complex strings of clauses/sentences + or
Independent user	B2	both + phrases/clauses + and; neither + words/phrases/clauses + nor; never + auxiliary/modal verb + clauses
	B1	therefore, etc. (as adverbs) + clauses/sentence; both + noun phrases + and; either + words/phrases/clauses + or; plus + noun, often in relation to numbers; and, but, or, so, then + more complex strings of clauses/sentences
Basic user	A2	whatever, etc. (as adverbs) + clauses/sentence
	A1	and, but, or + single noun/adjective/phrases; because + clauses

EGP's criterial features of the connector use across CEFR's levels (see Appendix) make the qualitative description in terms of control, linking elements and cohesive devices more specific with level of degrees, and are more helpful for English language teaching, learning and assessment.

Like the CEFR, the China's Standards of English Language Ability (CSE) (MoE, 2018), is the first national English proficiency framework, which also has 3 stages and 9 levels (i.e., Elementary (1-3), Intermediate (4-6), and Advanced (7-9)). In the grammatical competence description, the cohesion competence contains qualitative scale descriptors of learners' ability to construct coherent and cohesive texts (see Table 4).

Table 4

Descriptors of Connectors in the CSE

Stages	Levels	Descriptors
Advanced	9	Can flexibly use a variety of cohesive devices to create coherent and well-structured content.
	8	Can skillfully and appropriately use various cohesive devices to express him/herself logically and coherently.
	7	Can use cohesive devices to indicate relationships between paragraphs, including progression, cause and effect, and comparison and contrast. Can use lexical cohesive devices such as linking words and pronouns to reinforce the coherence of a text.
Intermediate	6	Can use lexical cohesion, omission, and other means to achieve the coherence of a complex text. Can use cohesive devices to indicate relationships between paragraphs (e.g., comparison and contrast, cause and effect, transitions). Can use appropriate cohesive devices to transition topics or viewpoints in a text.
	5	Can use various cohesive devices to logically organize utterances and texts. Can use cohesive devices to demonstrate relationships among sentences (e.g., comparison and contrast, cause and effect, progression, transition).
	4	Can effectively use references and linking words to realize text coherence. Can use words to express contrast, addition, and/or other logical relationships (e.g., however, although, nevertheless, similarly).
Elementary	3	Can use different linking words to connect ideas (e.g., when, if, because, although). Can use linking words and phrases to indicate relationships including addition, contrast, and sequential order (e.g., for example, then, first, second)
	2	Can use simple linking words to connect information (e.g., and, so).
	1	Can recognize textual features of simple short stories (e.g., "Once upon a time" often used to start a story).

Unlike the EGP, which list the connector profile across CEFR's six levels on the basis of learner corpus, the CSE only adopts scale descriptors to specify the cohesion competence of learners while lack criterial or distinctive features of connector use across its nine levels.

3 Research Questions

The current study hence aimed to conduct an empirical study to investigate the following two research questions:

1. How do Chinese college English students use connectors in their writing?
2. What are the distinguishing features that students demonstrate in their connector use?

4 Data Set

The data of the study were English writing composed by Year 1 to Year 4 Chinese college students from a national key university in the Northeast of China and *pigai wang* (www.pigai.org), an online automatic

scoring system. As shown in Table 5, nearly 17,580 English written texts were collected to build a learner corpus, consisting of four sub-corpora with 4,587 texts from Y1s, 4,998 from Y2s, 4,994 from Y3s, and 3,000 for Y4s. Each of the texts contain 200 words on average, covering eight themes (i.e., culture, education, environment, health, people, science and technology, ethics, and life) and four genres (i.e., narrative, expository, descriptive, and persuasive).

Table 5

Learner Corpus Information

Items	Year 1	Year 2	Year 3	Year 4
Texts	4,587	4,998	4,994	3,000
Types	18,473	22,693	20,789	19,611
Tokens	944,565	1,054,559	1,107,985	801,707
Mean word length (in characters)	4.61	4.65	4.65	4.70

The raw data were cleaned through eliminating incomplete texts, which have symbols, tables, figures or graphs. The corpus was then annotated (Liang, 2009) with Tree Tagger 3.0 for part of speech (POS) tagging. Then one out of every 400 texts was randomly selected to test for precision and recall. The accuracy rate was 96.74% and the recall rate was 93.38%. Research instrument, such as WordSmith 5.0 (Scott, 2008), was utilized for data analysis.

5 Results

The raw frequency and percentage of connector use is summarized in Table 6. Four connectors (i.e., *and*, *but*, *or*, *because*) were of higher frequency, with *and* the most frequent (69.20%), *hardly ... when* the least (0.002%). The use of one-word connectors exceeded multiword ones. The connector use was consistent from Year 1 to Year 4 students, with Year 4 students demonstrating more varied use.

Table 6

Frequency of Connector Use

Connectors	Year 1	Year 2	Year 3	Year 4	Total	Percentage
and	24,476	28,094	30,468	22,827	105,865	69.20%
but	4,410	4,593	4,956	3,171	17,130	11.20%
or	2,397	2,785	3,576	2,201	10,959	7.16%
because	1,380	1,420	1,773	1,203	5,776	3.78%
not only...but also	1,025	721	781	560	3,087	2.02%
therefore	906	769	837	428	2,940	1.92%
although/though	646	723	811	523	2,703	1.77%
never	587	557	686	338	2,168	1.42%
so	365	404	405	299	1,473	0.92%
as long as	158	102	171	88	519	0.34%
as if/as though	31	29	32	19	111	0.07%
as soon as	29	29	19	17	94	0.06%
in that	24	15	27	19	85	0.05%
either...or	6	20	8	17	51	0.03%
neither...nor	4	2	6	6	18	0.01%
hardly...when	--	1	2	--	3	0.002%

Note: percentage = raw frequency of one connector/total raw frequency of all connectors*100%

Table 7 displayed the patterns and positions of one-word connectors. The patterns of one-word connectors were usually positioned in the middle of statement sentences to connect words, phrases, clauses, and sentences, or placed at the beginning or in the inverted order of sentences. Year 3 students showed a strong tendency to use connectors at the sentence beginning while Year 2 students liked to use connectors in the middle of sentences.

Table 7

Positions and Patterns of One-word Connectors

Positions	Patterns	Year 1	Year 2	Year 3	Year 4	Total (%)
head	one-word connectors + words (n., 6,109		6,497	8,462	5,326	26,394
	adj., v., etc.)/phrases					(17.25%)
middle	one-word connectors + clauses/	2,369	2,165	2,365	2,108	9,007
	sentences					(5.89%)
	one-word connectors + words (n., 21,321		27,721	28,387	16,980	94,409
	adj., v., etc.) or phrases					(61.71%)
	one-word connectors + clauses/	4,822	6,930	4,096	6,324	23,172
	sentences					(15.15%)

As for multiword connectors, Table 8 revealed that the connector *not only ... but also* in statement sentence order was used more frequently than other multiword connectors. The inverted sentence order of *not only ... but also* was underused. The pattern ‘*Not only* + modal verb/*be*/verb + subject, *but also* + modal verb/*be*/verb’ was used more frequently by Year 1 students than ‘*Not only* + subject + verb, *but also* + verb’ by Year 3 students. The connector *as long as* is relatively of higher frequency. The rest of connectors (i.e., *as soon as*, *as if/as though*, *hardly ... when*, *either ... or*, *neither ... nor*) were all underused by students, with the sum of total frequency lower than 100.

Table 8

Positions and Patterns of Multiword Connectors

Connectors	Patterns	Y1	Y2	Y3	Y4	Total
not only...but also	subject + (modal v.) + not only + v., but also + (modal v.) + v.	910	654	697	499	2,760
	Not only + subject + v., but also + v.	27	31	36	29	123
	Not only + modal v./be/v. + subject, but also + modal v./be/v.	88	36	48	32	204
as long as	principal clause + as long as + subordinate clause	42	29	37	18	126
	as long as + phrases	1	1	1	--	3
as if/as though	as if/as though + indicative mood clause	24	25	22	13	84
	as if/as though + subjunctive mood clause	7	4	10	6	27
as soon as	as soon as + temporal adverbial clause	29	29	19	17	94
hardly...when	subject + (modal v.) + hardly + v., when + clause	--	1	1	--	2
	Hardly + be/modal v./auxiliary + v. + clause + when + clause	--	--	1	--	1
either...or	either + word/phrase/clause + or	6	15	8	14	43
	Either + be/modal v./aux. + v. + or	--	--	--	--	--
neither...nor	neither + word/phrase/clause + nor	4	2	6	6	18
	Neither + be/modal v./aux. + v. + nor	--	--	--	1	1

In view of meaning, Table 9 disclosed that connectors were mainly used to express *enumeration* and *addition*, *result* or *inference*, *contrast* or *concession*, and *summation*. The meaning of *enumeration* and *addition* was of the highest frequency while that of *summation* the lowest. Students from Y1 to Y4 showed no greater discrepancy.

Table 9

Meaning Conveyed by Connectors

Meaning	Connectors	Year 1	Year 2	Year 3	Year 4	Total (%)
enumeration and addition	and	24,476	28,094	30,468	22,827	119,980 (78.43%)
	or	2,397	2,785	3,576	2,201	
	not only...but also	1,025	721	781	560	
contrast	either...or	6	20	8	17	19,301 (12.61%)
	neither...nor	4	2	6	6	
	but	4,410	4,593	4,956	3,171	
	never	587	557	686	338	
	hardly...when	--	1	2	--	
result or inference	because	1,380	1,420	1,773	1,203	7,334 (4.79%)
	so	365	404	405	299	3,316 (2.17%)
concession	in that	24	15	27	19	
	although/though	646	723	811	523	
	as long as	158	102	171	88	
	as soon as	29	29	19	17	
summation	as if/as though	31	29	32	19	111 (0.07%)
	therefore	906	769	837	428	2,940 (1.92%)

With reference to the criterial features in the EGP, it can be seen from Table 10 that the connectors spread across all six CEFR levels, with the majority at the A1 (i.e., *and*, *or*, *but*, *because*) and B1 (i.e., *so*, *therefore*) levels and in statement sentence order. The inverted sentence order was mainly used by students of higher language proficiency from B2 to C2. The most frequently used connectors (i.e., *and*, *or*, *neither ... nor*, *not only ... but also*) are those expressing enumeration and addition. The meaning distribution across levels was imbalanced, with one-word connectors (i.e., *but*, *never*, *although/though*, *because*, *so*) expressing the meaning of *contrast*, *concession*, *result* or *inference* used more often than their multiword counterparts. The patterns of connector use were not complicated.

Table 10

Distribution of Connectors With Reference to the EGP

Connectors	Patterns	Meanings	Total	EGP
and, or	and, or + n. / adj. / phrases / clauses/sentences	enumeration & addition	116,824	A1
but	but + n. / adj. / phrases / clauses/ sentences	contrast	1,7130	A1
because	because + clauses/sentences	result or inference	5,776	A1
although/though	although/though + clauses/ sentences	concession	10,548	B1
so	so + more complex strings of clauses/sentences	result or inference	4,247	B1

therefore	therefore + clauses /sentences	result or inference	4,034	B1
either...or	either + word/phrase/clause + or	enumeration & addition	43	B1
hardly...when	Either + be/modal v. /aux. + v. + or subject + (modal v.) + hardly + v., when + clause	contrast	--	C1
	Hardly + be/modal v. /aux. + v. + clause + when + clause		2	B1
never	never + aux. + clauses /sentences	contrast	1	C2
as soon as	as soon as + temporal adverbial clause	concession	2,168	B2
as if/as though	as if/as though + indicative/ subjunctive mood clause	concession	94	B2
as long as	as long as + phrases	concession	3	B2
	principal clause + as long as + subordinate clause		126	C1
neither...nor	neither + word / phrase /clause + nor	enumeration & addition	18	B2
	Neither + be/modal v. /aux. + v. + nor		1	C2
not only...but also	subject + (modal v.) + not only + v., but also + (modal v.) + v.	enumeration & addition	2,760	C1
	Not only + subject + v., but also + v.		123	
	Not only + modal v./be/verb + subject, but also + modal v./be/v.		204	
in that	principal clause + in that + subordinate clause	result or inference	81	C2

6 Discussion

In terms of connector use, one-word connectors (e.g., *and*, *but* and *or*) are used more frequently than multiword ones, and are more likely to be positioned in declarative statement sentence orders. The frequent use of one-word connectors by college students from Y1s to Y4s may result from L1 transfer and students' insufficient knowledge of English grammar (Zhao, 2003). Multiword connectors are more difficult for Chinese college students to acquire. The study has disclosed that students' mastery of connectors does not improve with years of learning. The declarative statement sentence order is used more often than the inverted ones. The connectors with the meanings of *enumeration* and *addition* are of significantly high frequency, which may lie in the differences between Chinese and Westerner mindset and their frequent use in academic and argumentative writing (Chen & Jiang, 2015).

As for the scale distribution of connector use, Chinese college students are inclined to use connectors at lower EGP levels (A1 (e.g., *and*, *or*, *but*, *because*) and B1 (e.g., *so*, *therefore*), with the majority of which at A1 level, in one-word form and in declarative statement sentence order (Li, 2011). The results also indicate that the connector use lacks of variety, with some either overused or underused (Xu & Zhang, 2006). The students from Y1 to Y4, though displayed some variation in connector use, did not show greater discrepancies.

Based on the empirical corpus-driven study and with reference to the EGP, the cohesion competence that students demonstrate in their writing in relation to the CSE stages and levels can be summarized in Table 11. The frequency of connector use over 40 are counted as salient distinctive features.

Table 11
Summary of Students' Cohesive Competence in the CSE

Stages & levels	Connectors	Patterns	Meanings
Advanced Levels 7-9	never	never + auxiliary + clauses / sentences	contrast
	not only ... but also	subject + (modal verb) + not only + verb, but also + (modal verb) + verb Not only + subject + verb, but also + verb Not only + modal verb/be/verb + subject, but also + modal verb/be/verb	enumeration & addition
	as long as	principal clause + as long as + subordinate clause	concession
	as soon as in that	as soon as + temporal adverbial clause principal clause + in that + subordinate clause	concession result or inference
Intermediate Levels 4-6	although /though	although/though + clauses / sentences	concession
	so	so + more complex strings of clauses / sentences	result or inference
	therefore either ... or	therefore + clauses/sentences either + word/phrase/clause + or	result or inference enumeration & addition
Elementary Levels 1-3	and , or	and, or + nouns / adjectives / phrases / clauses / sentences	enumeration & addition
	but	but + nouns / adjectives / phrases / clauses / sentences	contrast
	because	because + clauses/sentences	result or inference

As shown in Table 11, the connector use by college students are fairly limited in terms of their variety in forms, meanings and patterns as described in the cohesion competence across CSE's 9 levels (see Table 3). With reference to the EGP, the distinctive features of students' connector use patterns in their English writing can be made salient, and further mastery of connector use in forms, meanings and patterns can be emphasized in college English teaching.

7 Conclusion

The current study has disclosed that although the connector use in Chinese college students' writing spread across all levels, students in general demonstrate a strong tendency of overusing some one-word English connectors while underusing those multiword ones, with connectors clustering at lower CEFR levels (i.e., A1 & B1) or CSE's elementary and intermediate levels, and connectors at higher levels rarely used. Among all the connectors, the one-word connector *and* is of the highest frequency, followed by *but* and *or* respectively. For the multiword connectors, the most frequently used is *not only...but also* in the pattern of 'subject + (modal verb) + *not only* + verb, *but also* + (modal verb) + verb', with *because of* coming the next in the pattern of '*because of* + phrases/nouns'.

In terms of sentence order, the connectors are mostly used in declarative statement sentence order and in the middle of sentences rather than in inverted sentence order. In view of meaning expressed, connectors expressing *enumeration* and *addition* (i.e., *and*, *or*, *not only ... but also*) are the most frequently used, followed by connectors expressing *contrast* (i.e., *but*, *never*), *result or inference* (i.e., *because*, *so*) and *concession* (i.e., *although/though*) and *summation* (i.e., *therefore*). Across four grades

of college students, the connector use is quite consistent and there is no distinctive discrepancy from Year 1s to Year 4s. On the basis of the empirical study, we can safely come to the conclusion that the high frequency of certain connector use also indicates the lack of variety of connectors in terms of forms, orders, meanings and patterns.

The implication of the current corpus-based study lies in that it presents an overall pattern and meaning distribution of college students' cohesion competence and their connector use in their English writing with reference to the CEFR and the EGP, which can provide significant insights on future CSE-based cohesion competence study on the aspects of forms, patterns and meanings. It is expected that the study can shed light on college English curriculum and syllabus design to take more effective measures to improve students' connector use as well as their cohesion competence in their English writing. It is also expected that the study can further enrich scale descriptors and criterial features of the CSE.

Acknowledgements

This paper was based on the project "Corpus-based Criterial Feature Parameter Study of College English language Competence" funded by National Planning Office of Philosophy and Social Science Program (13AYY005).

Appendix

Criterial Features of Connectors in the EGP

Connectors	Part of speech	Patterns	Examples	Level
and, but, or	conj.	and, but, or+single noun/ adjective.	I like to eat soup, eggs, meat and salad or banana and cheese. (A1 BREAKTHROUGH; 2006; Polish; Pass)	A1
		and, but, or+phrases	She likes to go to the cinema or to the theatre. (A1 BREAKTHROUGH; 2006; French; Pass)	A1
		and, but, or+clauses.	There are lots of shops in the town and the town is only five minutes away. (A1 BREAKTHROUGH; 2008; Urdu; Pass)	A1
		and, but, or+sentences	I like it because there are lots of shops, it is not too busy and there is lots of parking. And I like my home because it's not too big and it's a very nice flat. (A1 BREAKTHROUGH; 2009; Tamil; Pass)	A1
so, then, etc.	conj.	and, but, or, so, then+more complex strings of clauses/ sentences.	Then we went to his limousine and we went to his studio and he was singing 'Thriller', then I heard a strange voice, I woke up, and there was my mom waking me up. (Colombia; B1 THRESHOLD; 2009; Spanish - Latin American; Pass)	B1
plus	conj.	plus+noun, often in relation to numbers	In my house we are six people, my mother, my father, my sister and my two brothers, plus my two cats, which are members of the family too. (B1 THRESHOLD; 2003; ; Pass)	B1

		plus+clauses/sentences	It will be fun because everyone is going to be there. Plus, the park is the best place for a picnic, we can run and just have fun! (Philippines; B1 THRESHOLD; 2009; Tagalog; Pass)	B1
either...or	conj.	either+words/phrases/clauses+or	The men wear suits, either white or black, and the bride and women wear special long dresses, usually the same colour. (Yugoslavia; B1 THRESHOLD; 2006; Serbian; Pass)	B1
		either+more complex strings of clauses/sentences+or	And finally, about the present, you have two options; either you get something typical from your country, which they may like because you know how much they like collecting things from all over the world, or I get something from the wedding list store and sent it to them with your name on it and we can sort it out when you arrive here. (Spain; C1 EFFECTIVE OPERATIONAL PROFICIENCY; 1998; Catalan; Pass)	C1
both...and	conj.	both+noun phrases+and	I felt sorry for you because both friends and family are very important and having to pick one is very hard. (Turkey; B1 THRESHOLD; 2008; Turkish; Pass)	B1
		both+phrases/clauses+and	A lot of people think that television can both entertain and educate. (Italy; B2 VANTAGE; 2004; Italian; Pass)	B2
neither/nor	conjunctions	neither/nor+inverted auxiliary/'be'+subject+a previous related negative clause	Nevertheless, I have no recollection of ever leaving my diary. In fact, I can clearly remember that I kept it with me all the time I was working. Neither do I remember losing it during the time I was on the platform waiting for my train to pull in. (Spain; C2 MASTERY; 1993; Catalan; Pass)	C2
neither...nor	conj.	neither+ words/phrases/clauses+nor	We can neither eat nor drink during the lessons. (Poland; B2 VANTAGE; 2000; Polish; Pass)	B2
yet	conj.	yet+phrases/clauses	This college is famous for its art teachers and very skilled students, and yet it does not have its own gallery. (C1 EFFECTIVE OPERATIONAL PROFICIENCY; 2006; Polish; Pass)	C1
and yet	conj.	and yet+sentences	I felt very embarrassed and I was sure I would get the sack. And yet the manager forgave me and I continued to work there. (Poland; B2 VANTAGE; 1993; Polish; Pass)	C2
not only...but also	conj.	not only+phrases/clauses+but also	If we go back in time we will see that the difference between the old and the modern films is not only the number produced but also the content. (Greece; C1 EFFECTIVE OPERATIONAL PROFICIENCY; 2007; Greek; Pass)	C1

		not only+inverted auxiliary 'do'+but also	For me, not onlydo extreme sports provide satisfaction of achievement, but they also help you to express yourself. (Greece; C1 EFFECTIVE OPERATIONAL PROFICIENCY; 2008; Greek; Pass)	C1
because	conj.	because+clauses/sentence	I want a big garden because I have children. (Arabic - Other)	A1
if, when, so, while (limited range of one-word connectors), etc.	conj.	simple conjunctions+clauses/sentence	All of us were dancing while my father sang my favourite song. (Arabic - Meghreb) My father bought it for me when he saw my grades at school. (Greek)	A2
as, after, before since, until, although, whether, so (that), though (increasing range of simple conjunctions), etc.	conj.	simple conjunctions+clauses/sentence	I usually read during the afternoon after I have finished my homework. (Italy; B1 THRESHOLD; 2005; Italian; Pass) But I didn't remember my ring until I got back home. (Korea, South; B1 THRESHOLD; 2000; Korean; Pass) Although the holiday was fantastic, I wished that you were with us. (Libya; B1 THRESHOLD; 2003; Afrikaans; Pass)	B1
once, whereas, unless, except (that) provided (that) (wide range of simple conjunctions), etc.	conj.	simple conjunctions+clauses/sentence	Also, despite some customers are always in the shops, the electricity is not really 'wasted'. (Japan; B2 VANTAGE; 2001; Japanese; Pass) Unless the situation improves, we will be forced to try a new supplier. (Romania; B2 VANTAGE; 2005; Romanian; Fail)	B2
as long as, as soon as, in order that, despite the fact that, due to the fact that, as if, as though (complex conjunctions), etc.	conj.	complex conjunctions+clauses/sentence	So take a rest in your hotel, in order that you can be prepared for your packed schedule. (Greece; B2 VANTAGE; 1993; Greek; Pass) To sum up, the festival was successful, so I hope most people enjoyed the festival despite the fact that there were some problems with the poor sound system, and so on. (Korea, South; B2 VANTAGE; 1997; Korean; Pass) Due to the fact that the mass of cars pollute the cities and the environment, we are forced to give up our city habits and learn to be more sensible by using public transportation. (Germany; B2 VANTAGE; 1993; German; Pass)	B2
whatever, wherever, however, etc.	conj.	(at the beginning of a sentence) Whatever, etc.+clauses/sentence	Wherever you go, you find shops, hotels and pubs which are full of tourists. (Greece; C1 EFFECTIVE OPERATIONAL PROFICIENCY; 2003; Greek; Pass)	C1
	adv.	whatever, etc. (as adverbs)+clauses/sentence	Yesterday I bought some clothes: trousers and skirts. However, the clothes were cheap. (Spain; A2 WAYSTAGE; 2008; Spanish - European; Pass)	A2

therefore, furthermore, otherwise, etc.	adv.	therefore, etc. (as adverbs)+clauses/ sentence	You don't like cities with a lot of people. Therefore I suggest you go to a small school in the countryside, near your new home. (Switzerland; B1 THRESHOLD; 2008; French; Pass)	B1
never	adv.	never+auxiliary/modal verb+clauses	Never had I thought that my parents would do this to me. (Netherlands; B2 VANTAGE; 1993; Dutch; Fail)	B2
hardly...when	adv.	hardly+had+clauses+when	Hardly had she finished her complaints when Bob stood up, paid and left. (Switzerland; C2 MASTERY; 1993; French; Pass)	C2
in that	conj.	in that+clauses/sentence	It's often said that nowadays people must be proud of medical advances, in that life is getting considerably longer. (Spain; C2 MASTERY; 1993; Catalan; Fail)	C2

References

- Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. *Learner English on Computer*. Routledge.
- Baker, P. (2004). Corpus linguistics and south Asian languages: Corpus creation and tool development. *Literary & Linguistic Computing*, 19(4), 509-524. <https://doi.org/10.1093/lc/19.4.509>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Blagoeva, R. (2004). Demonstrative reference as a cohesive device in advanced learner writing: A corpus-based study. In Aijmer, K., & Altenberg, B. (Eds.), *Advances in Corpus Linguistics* (pp. 297-307). Brill Rodopi. https://doi.org/10.1163/9789004333710_018
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1983). *The grammar book: An ESL/EFL teacher's course*. Newbury House.
- Chen, Z., & Jiang, G. (2015). Analysis on addition connectors in English writing based on the CLEC and Brown corpus. *Journal of Southwest University of Science and Technology (Philosophy and Social Science Edition)*, 6, 62-66.
- Council of Europe. (2001). *Common European Framework of Reference for languages: Learning, teaching, and assessment*. Cambridge University Press.
- Council of Europe. (2018). *Common European Framework of Reference for languages: Learning: Learning, teaching and assessment: Companion volume with new descriptors*. Council of Europe.
- Crossley, S. A., & McNamara, D. S. (2016). Say more and be more coherent: How text elaboration and cohesion can increase writing quality. *Grantee Submission*, 7(3), 351-370.
- Field, Y. (1994). Cohesive conjunctions in the English writing of Cantonese speaking students from Hong Kong. *Australian Review of Applied Linguistics*, 17(1), 125-139. <https://doi.org/10.1075/aral.17.1.07fie>
- Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27. <https://doi.org/10.1111/j.1467-971X.1996.tb00089.x>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- Hawkins, J., & Filipović, L. (2012). *Criterion features in L2 English: Specifying the reference levels of the Common European Framework*. Cambridge University Press.

- Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied Language Learning*, 12(2), 111-132.
- Jin, W. (2001). *A quantitative study of cohesion in Chinese graduate students' writing: Variations across genres and proficiency levels*. In the Symposium on Second Language Writing (West Lafayette, IN, September 15-16, 2000) (pp.1-34). (ERIC document reproduction service no. ED452726).
- Kang, J. Y. (2005). Written narratives as an index of L2 competence in Korean EFL learners. *Journal of Second Language Writing*, 14, 259–279. <https://doi.org/10.1016/j.jslw.2005.10.002>
- Khalil, A. (1989). A study of cohesion and coherence in Arab EFL college students' writing. *System*, 17(3), 359-371.
- Leech, G. (2006). *Glossary of English grammar*. Edinburgh University Press.
- Li, P. (2011). A Corpus-based comparative analysis on connectors in college English writing. *College English Teaching and Research* (5), 25-31.
- Liang, M. (2009). Retrieving on POS corpora and writing regular expression. *Foreign Language Education in China*, 2, 65-73.
- Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, 33, 623–636. <https://doi.org/10.1016/j.system.2005.02.002>
- Matthews, P. H. (2000). *Oxford concise dictionary of linguistics*. Shanghai Foreign Language Education Press.
- Medve, V. B., & Takač, V. P. (2013). The influence of cohesion and coherence on text quality: A cross-linguistic study of foreign language learners' written production. In Piechurska-Kuciel, E., & Szymańska-Czaplak, E. (Eds.) *Language in cognition and affect* (pp. 111-131). Springer. https://doi.org/10.1007/978-3-642-35305-5_7
- Milton, J. C., & Tsang, E. S. C. (1993). A corpus-based study of logical connectors in EFL students' writing: Directions for future research. In Pemberton, R., & Tsang, E.S.C. (Eds.) *Studies in Lexis. Proceedings of a seminar on lexis* (pp. 215-246). The Hong Kong University of Science and Technology Language Centre. DOI: <http://dx.doi.org/10.7575/aiaac.ijalel.v.5n.1p.113>
- Ministry of Education (MoE) of the People's Republic of China, & National Language Commission (NLC). (2018). *China's standards of English language ability*. <http://cse.neea.edu.cn/html1/folder/1505/244-1.htm>
- Norment, N. (1994). Contrastive analyses of cohesive devices in Chinese and Chinese ESL in narrative and expository written texts. *Chinese Language Teaching Association Journal*, 29(1), 49-81.
- North, B. (2007). The CEFR illustrative descriptor scales. *The Modern Language Journal*, 91(4), 656-659.
- Quirk, R., & Crystal, D. (1985). *A Comprehensive grammar of the English language*. Longman.
- Salamoura, A., & Saville, N. (2010). Exemplifying the CEFR: Criterial features of written learner English from the English Profile Programme. *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*, 1, 101-132.
- Scott, M. (2008). WordSmith Tools version 5, Lexical Analysis Software.
- Song, M., & Xia, W. (2002). Combination of textual cohesive ties and textual teaching for the teaching of English writing—a statistical analysis of the good and poor compositions written by non-English major freshmen. *Foreign Language World*, 6, 40–44.
- Todd, R. W., Khongput, S., & Darasawang, P. (2007). Coherence, cohesion and comments on students' academic essays. *Assessing Writing*, 12(1), 10-25. <https://doi.org/10.1016/j.asw.2007.02.002>
- Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication*, 32(2), 189-204.
- Xu, W., & Zhang, B. (2006). Corpus-based contrastive studies on the causal conjunctions in English/

Chinese classics. *Foreign Language Teaching and Research*, 38(4), 292-296.

Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education*, 23(1), 31-48. <https://doi.org/10.1016/j.linged.2011.09.004>

Zhang, M. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal*, 31, 61–95. <https://doi.org/10.1177/003368820003100104>

Zhao, W. (2003). A quantitative comparative analysis on logic connectors in Chinese EFL learners' writing. *Foreign Language Education*, 24(2), 72-77.

Nairui Chen is a PhD student in Applied Linguistics a GTA in Interdisciplinary Writing Program (IWP) at Northern Arizona University. Her research interests are task variation, L2 writing, and writing quality of EFL learners. ORCID: 0000-0002-6228-4895

Wen Zhao is Dean of the School of Foreign Studies at Jinan University, Guangzhou. Her main publication and research interests are in corpus linguistics, English curriculum and instruction, and EFL writing. She has been working and researching in English language teaching and learning, and has been involved in national English curriculum development for senior secondary vocational education and College English education. ORCID: 0000-0003-4965-0146