A Contrastive Study on Lexical Bundles in English Debating by Chinese and American Debaters Based on SPSS Software

Limin Han[1*], Aiyun Chen[1], Rongjie Zhai[2]
[1]School of Foreign Studies, Anhui Polytechnic University, Wuhu, Anhui, China
[2]School of Artificial Intelligence, Anhui Polytechnic University, Wuhu, Anhui, China
*Corresponding Author.

Abstract:

Based on two self-built corpuses of English debating speeches, a contrastive study on lexical bundles used by Chinese and American debaters was conducted, by applying software Concapp and SPSS. Firstly, through software Concapp, lexical bundles (connectors and exemplifiers) in the two corpuses were searched on the basis of encoding them by CLAWS－－an automatic encoding software. Then, by adopting software SPSS, the mathematical and statistical analysis of using connectors and exemplifiers was carried out from two perspectives: density and richness degree. It was discovered that: the density and richness degree of connectors used by Chinese debaters was lower than that of connectors used by American debaters; the density and richness degree of exemplifiers used by Chinese debaters was higher than that of exemplifiers used by American counterparts. It was hoped that the research could provide some pedagogical implication and guidance on debate teaching practice.

Keywords: Lexical bundle, Corpus, Concapp & SPSS, Statistical analysis.

I. INTRODUCTION

With the arrival of economic globalization and further deepening of China's reform and opening-up, nation-wide English learning is in full swing, with the number and passion of learners expanding with each passing day. Against this national background, English debate, by attaching much importance to training students' oral expression and affective language usage, is increasingly popular among Chinese English learners and educators. Therefore, the research on English debate gradually blooms in the world. Kwak [1] investigated the different types of politeness strategies that were used between the different participants based on the theory of politeness strategy in relation to U.S. Presidential debates. Young [2] examined the learning effects of discussion &debate-based learning and analyzed the students' further suggestions for effective discussion and debate. The research provided effective strategies for the debate-based learning in English literature classes. Susana [3] compared in-class university structured debate implementation methods among undergraduate nursing students, and identified the effect of such debate methods in the students' learning. Hawkins [4] discovered that the debate-dominant course structure could increase accountability of learning and the debate-style elective course was perceived to be beneficial to students.

Praharaj [5] suggested that debate-based learning, with adequate opportunities for interaction, followed the principles of adult learning and was well appreciated than the conventional seminars. Furthermore, debates help to enhance the critical thinking ability of the participants.

In China, some scholars also explored the effect of BP Debate training on oral English of English majors in China. The approaches of cultivating students' critical thinking ability through English debate were introduced and studied [6-8]. WANG [9] conducted a contrastive study on vague language in English debates by Chinese and American college students.

From what is mentioned above, it can be discovered that the research on English debate mainly focuses on two aspects. The first focus lies in English debate itself, including debating strategies, debating manners and how to conduct a debate in practical teaching. The other research focus is the function of English debate in enhancing students' oral English and critical thinking abilities. To sum up, the above research is qualitative research and theoretical research.

However, against this research background, the quantitative study on English debate by applying computer technology and corpus seems to be rarely seen. In the era of big data, computer and corpus technology has been widely applied in linguistic research and greatly broaden research scope and deepen research depth. Therefore, in order to enrich, broaden and deepen research on English debate, this study, by building corpuses and adopting computer technology, conducts a contrastive study on using lexical bundles in English debating speeches by Chinese and American debaters. Firstly, two corpuses of debating speeches are built by recording and transcribing speeches of Chinese and American debaters. Then, through software Concapp, lexical bundles in the two corpuses (connectors and exemplifiers) are searched on the basis of encoding them by CLAWS. Finally, by adopting the software SPSS 22.0, the mathematical and statistical analysis of using connectors and exemplifiers is carried out.

II. LEXICAL BUNDLES

2.1 Definition of Lexical Bundles

In 1957, the British linguist Firth, J.R. mentioned firstly that the meaning of a certain word should be decided from its collocation with other words [10]. Thereafter, the concept "lexical bundle" began to draw wide attention of many linguists. The lexical bundle is a special type of multi-word combination, which appears in the oral communication with high frequency [11-13]. It is the patterned usage of fixed semantic meanings, grammatical rules and pragmatic contexts. Lexical bundles has many different terms, such as "lexical phrases" , "formulae" , "routines", and "prefabricated patterns" , "recurrent word combinations" , "multi-word combinations", "chunks" and so on [14-17]. However, among so many names, there is a great consensus that they have the following common features: (1) Lexical bundles are the basic units of language use; (2) Lexical bundles appear with high frequency, so that they can be conveniently stored and applied; (3) Different from traditional vocabulary, the form of lexical bundles can be diversified, including multi-word combination, some fixed phrases and even sentence patterns [18].

With fixed structures, convenient storage and easiness of being memorized, lexical bundles become the most basic language units and play a key role in learners' writing and speaking. The competence of using lexical bundles can greatly contribute to one's language proficiency. The linguist Cortes held the opinion that "the frequent use of lexical bundles seems to signal competent language use within a register…" [19-23]. Therefore, a mastery of lexical bundles is commonly regarded as an essential step leading to "native-like" production of language. In view of their wide use in speeches, it is of great importance to examine how English learners use them in their debate so that some implication can be gained for practical foreign language learning and teaching.

2.2 Two Types of Lexical Bundle

The classification of lexical bundles can be diversified, with no definite conclusion drawn until now. However, earlier study made by Lewis, Nattinger & Decarricoand Sun Xiaoqing revealed that connectors and exemplifiers were most widely applied in debating speech, because they contributed a lot to expressing personal preference, stance and attitude [24]. Hence, the present study aims to explore the application of the connectors and exemplifiers in English debating by Chinese and American debaters.

III. RESEARCH QUESTIONS

This study focuses on features of using connectors and exemplifiers in English debates by Chinese and American debaters from 2 aspects: density and richness degree. Therefore, it attempts to answer the following research questions:

(1) Is there any difference on density of using connectors and exemplifiers in debating by Chinese and American debaters? If there is, what is the difference?

(2) Is there any difference on richness degree of using connectors and exemplifiers in debating by Chinese and American debaters? If there is, what is the difference?

(3) What pedagogical implication can this research bring to English debate teaching?

IV. RESEARCH METHODS

4.1 Corpus Building

With the development of computer technology, the corpus, with characteristics of convenience, speediness and accuracy, has been widely used in applied linguistic study, because this textual body can be read directly by computers [25-26]. The present study is based on two corpus built by the author: CEDCD (Corpus of English Debate of Chinese Debaters) and CEDAD (Corpus of English Debate of American Debaters). CEDCD is built via recording and transcribing 32 best debaters' speeches in the final

competition of 2021 "FLTRP Cup" English Debating Competition. Likewise, CEDAD is set up via collecting 32 best debaters' speeches from 2020 United States Universities Debate Championships. It is proved that the two corpuses are scientific and valid for the study. On the one hand, both competitions are conducted in the form of British Parliamentary Debate, with 7 minutes' speech time for each debater from two parties. Thus, the size of the two corpuses is almost the same, which can guarantee the research validity. On the other hand, debating topics of the two competitions cover a rich field, including environmental protection, medical care, cultural exchange, global development or whatever, so that there is compatibility in terms of topic and language style.

4.2 Lexical Bundles Encoding

The lexical bundles in above-mentioned 64 speeches are encoded through automatic encoding software CLAWS, designed by Lancaster University in UK. The software identifies part of speech of words automatically through a detailed and complex encoding set, classifying words into 10 groups, i.e. 10 kinds of lexical bundles. Therefore, this software provides much guidance and convenience for conducting research on lexical bundles and syntax. The main codes associated with this study are listed as following:

<AAT1>=definite article; <CCB>= adversative conjunction; <CS>= subordinating conjunction; <CC>= coordinating conjunction; <CSA>= conjunction, as; <CSN>= conjunction, than; <CSW>= conjunction, whether; <AAT2>=indefinite article; <IFF>= proposition, for; <II>= proposition; <IO>=of; <RP>=adjective; <VBZ>=is.

4.3 Lexical Bundles Searching

In corpus-based research, the searching of corpus text information is often realized by searching software. The text searching software Concapp can read directly texts of various forms with no need to transform so that the text searching and identifying is accurate, rapid and convenient. Therefore, in this study, software Concapp is adopted to identify and analyze lexical bundles (connectors and exemplifiers) in 64 speeches in two corpuses. The searched connectors and exemplifiers used by Chinese and American debaters are presented in TABLE I and TABLE II, respectively.

4.4 Statistical Analysis

Based on searched lexical bundles by software Concapp, statistical analysis of using connectors and exemplifiers in terms of density, richness degree and Independent Sample T test, is carried out by adopting statistical software SPSS22.0. The analysis result will be reported in the following chapter.

V. ANALYSIS AND DISCUSSION OF RESEARCH RESULTS

Based on research questions, the contrastive study on features of using connectors and exemplifiers in debate by Chinese and American debaters is conducted. The research results are discussed and analyzed as

follows.

5.1 Lexical Bundles Searching Result

Firstly, connectors used in English debate by Chinese and American debaters are searched and shown in TABLE I.

From TABLE I, it can be concluded that connectors used by Chinese debaters are less various than those used by American debaters.

Meanwhile, exemplifiers used in English debate by Chinese and American debaters are searched and shown in TABLE II.

From the above table, the conclusion can be arrived at that exemplifiers used by Chinese debaters are more various and diversified than those used by American debaters. The diversity of using exemplifiers is a special feature of speaking and writing of Chinese debaters.

TABLE I. Connectors used by Chinese and American debaters

DEBATERS	LEXICAL BUNDLE TYPE(CONNECTORS)
CHINESE DEBATERS	first, second, finally, last but not least, in addition, moreover, furthermore, in a word, to sum up, above all, to conclude, because, so, but, if, when, or, however, since, in contrast, on the country, thanks to, due to, because of., on the other hand
AMERICAN DEBATERS	first of all, to start with, for a start , to begin with, to start with, moreover, in addition, nevertheless, next, similarly, equally, hence, consequently, particularly, in particular, furthermore, likewise, alternatively, admittedly, incidentally, to cap it, to get back to the point, to change the subject, with reference to, on the country, however, in addition, in spite of, instead, owing to, due to whereas, in brief, briefly, on the whole, to summarize, in short, in closing, in all, above all, in conclusion

5.2 SPSS Analysis Results

The software of SPSS is good to analyze language phenomenon [27, 28]. Through SPSS analysis, it can be discovered that there is noticeable difference in density and richness degree of connectors and exemplifiers used by Chinese and American debaters.

TABLE II. Exemplifiers used by Chinese and American debaters

DEBATERS	LEXICAL BUNDLE TYPE(EXEMPLIFIERS)
CHINESE DEBATERS	for example, such as, for instance, namely, in other words, that is, to illiterate, in fact, as an illustration, to name a few, a case in point, to put it in another way, take... as an example, let's say
AMERICAN DEBATERS	for example, for instance, namely, I mean, such as, to give you an example, and son on, like, to illiterate

(1) Density of connectors

Density refers to the ratio between the number of lexical bundles and that of words in whole text. A larger density means that there are more lexical bundles within a text with a certain number of words. Through SPSS22.0 statistical analysis, the density of using connectors in debating by Chinese and American debaters can be expressed in Fig 1.
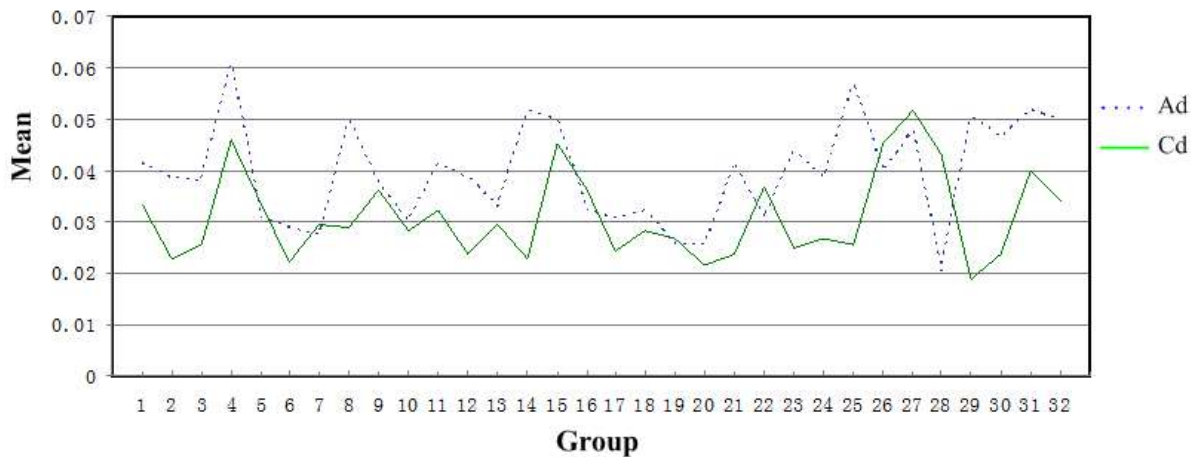

Fig 1: Density of connectors
(Ad represents density of using connectors by American debaters, whereas Cd represents density of using connectors by Chinese debaters)

From Fig 1, it can be found that the density of using connectors by American debaters is from 0.032 to 0.050, while the density of using connectors by Chinese debaters is from 0.021-0.034. The result indicates that American debaters use 3 to 5 connectors every 100 words, while Chinese debaters use 2 to 3 connectors every 100 words. Moreover, TABLE III reveals that through Independent Sample t Test, the density of using connectors by Chinese and American debaters is distinctively different ([t (62) = 3.534, p = 0.002 < 0.05]).   The combination of Fig 1 and TABLE III reveals that the density of using connectors by American debaters is obviously higher than that by Chinese debaters.

TABLE III. Independent sample t test of density of connectors

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | 95% Confidence Interval of the Difference | |
		F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Ld	Equal Variances Assumed	1.077	.304	3.534	62	.002	.008435	.003415	.003689	.013368
	Equal Variances not Assumed			3.534	59.201	.002	.008435	.003415	.003689	.013368

It is revealed that Chinese debaters apply fewer connectors than American ones in debating, which leads to loose structure, chaotic sentences and weak arguments. The difference can be explained by different linguistic features of Chinese and English. Generally speaking, Chinese is a language of parataxis, coordination of clauses in succession without connectors, while English is a language of hypotaxis, syntactical subordination of clauses by the use of connectors. Naturally, Chinese college students are influenced by the parataxis feature of Chinese, failing to use as many connectors in debating as their American counterparts.

(2) Density of exemplifiers

Through SPSS 22.0 statistical analysis, the density of using exemplifiers in debate by Chinese and American debaters can be expressed in Fig 2.

From Fig 2, it can be found that the density of using exemplifiers by Chinese debaters is from 0.012 to 0.021, while the density of using exemplifiers by American debaters is from 0.005 to 0.018. The result indicates that Chinese debaters use 1 to 2 exemplifiers every 100 words, while American debaters use 1 exemplifier every 200 words. Moreover, through Independent Sample t Test, the density of using exemplifiers by Chinese and American debaters is distinctively different ([t (62) = -3.901, p = 0.001 < 0.05]. The combination of Fig 2 and TABLE IV reveals the density of using exemplifiers by Chinese debaters is obviously higher than that by American debaters.

This difference can find its root in people's thinking pattern, which has a close relationship with learners' choice of words in speech. People's thinking pattern in the world is hugely different due to rich variety of historical background and geographical condition. For instance, Chinese debaters, influenced by long-established concrete thinking pattern and teachers' overstating function of examples, tend to overuse exemplifiers in debates to enhance power and vividness of argument. Therefore, it is easy to explain why Chinese debaters use more exemplifiers than their American counterparts.
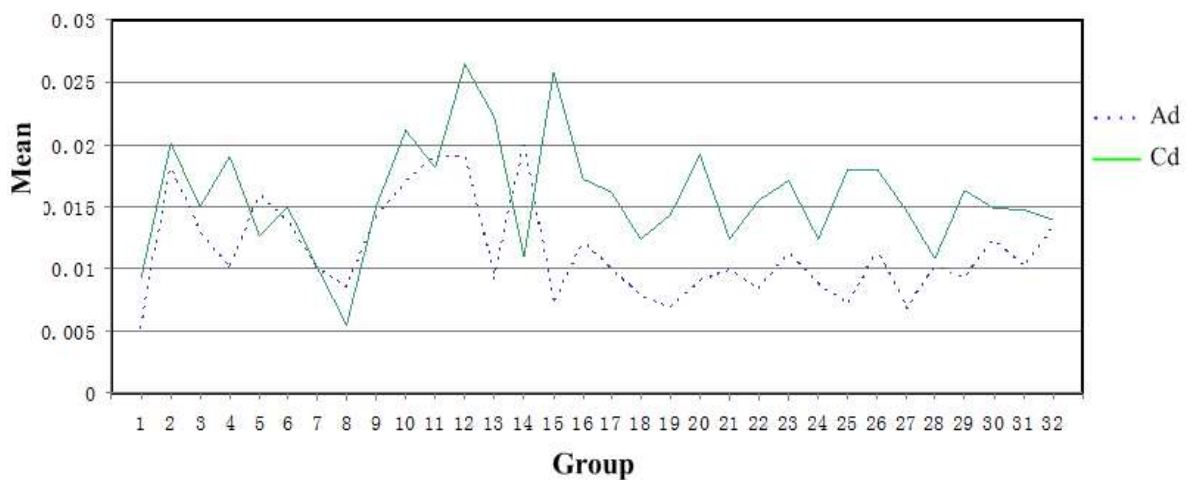

Fig 2: Density of exemplifiers

TABLE IV. Independent sample t test of density of exemplifiers

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
									Lower	Upper	
Ed	Equal Variances Assumed	0.5	.825	-3.901	62	.001	-.005333	.001241	-.006546	-.002310	
	Equal Variances not Assumed			-3.901	60.841	.001	-.005333	.001241	-.006546	-.002409	

(3) Richness degree of connectors

Through SPSS 22.0 statistical analysis, the richness degree of using connectors in debating by Chinese and American debaters can be expressed in Fig 3.

From Fig 3, it can be found that the richness degree of using connectors by American debaters is from 10 to 20, while the richness degree of using connectors by Chinese debaters is from 7 to 14. Moreover, TABLE V reveals through Independent Sample T test that the richness degree of using connectors by Chinese and American debaters is obviously different ([t (62) = 4.669, p = 0.002 < 0.05]). A conclusion can be arrived that the richness degree of using connectors by Chinese debaters is obviously lower than that of American ones.

This result indicates that, compared with their American counterparts, Chinese debaters use fewer types of connectors, which means that there is still much room for Chinese debaters to learn from American native speakers.
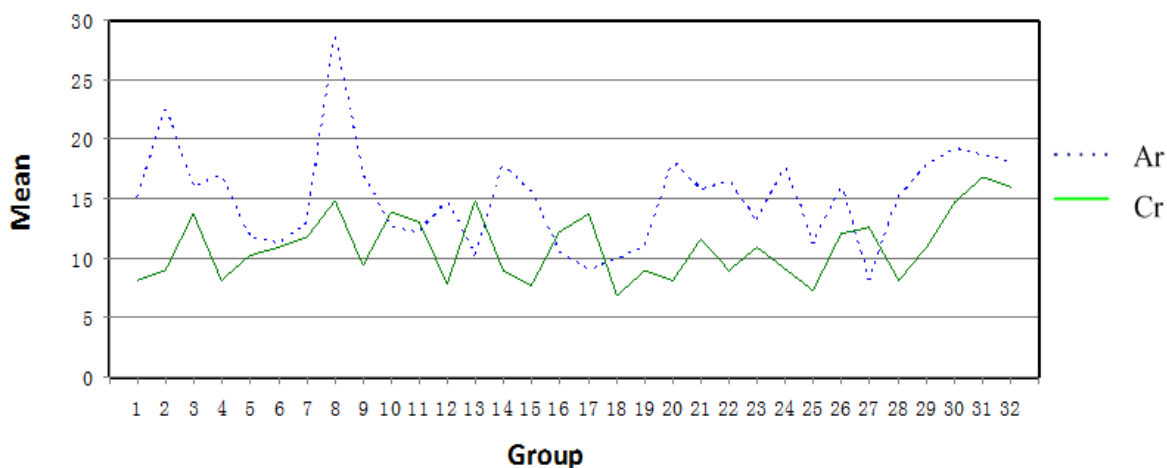

Fig 3: Richness degree of connectors

(4) Richness degree of exemplifiers

Through SPSS 22.0 analysis, the richness degree of using exemplifiers in debates by Chinese and American debaters can be expressed in Fig 4.

From Fig 4, it can be found that the richness degree of using exemplifiers by American debaters is from 3 to 5, while the richness degree of using exemplifiers by Chinese debaters is from 4 to 6. Moreover, TABLE VI reveals through Independent Sample T test that the richness degree of using exemplifiers by Chinese and American ones is obviously different ([t (62) = -2.855, p = 0.004 < 0.05]). It can be found that the richness degree of using exemplifiers by Chinese debaters is obviously higher than that by American debaters. This result indicates that, Chinese debaters are equally competent in using exemplifiers with their

American counterparts.

Two reasons can account for Chinese debaters' equal competence in using exemplifiers with American ones. The first reason is that Chinese, influenced by concrete thinking pattern, attach much importance to accumulation and application of exemplifiers. The other reason is that variety of exemplifiers in English is much poorer than that of connectors, so it is comparatively easier for Chinese to grasp exemplifiers.

TABLE V. Independent sample t test of richness degree of connectors

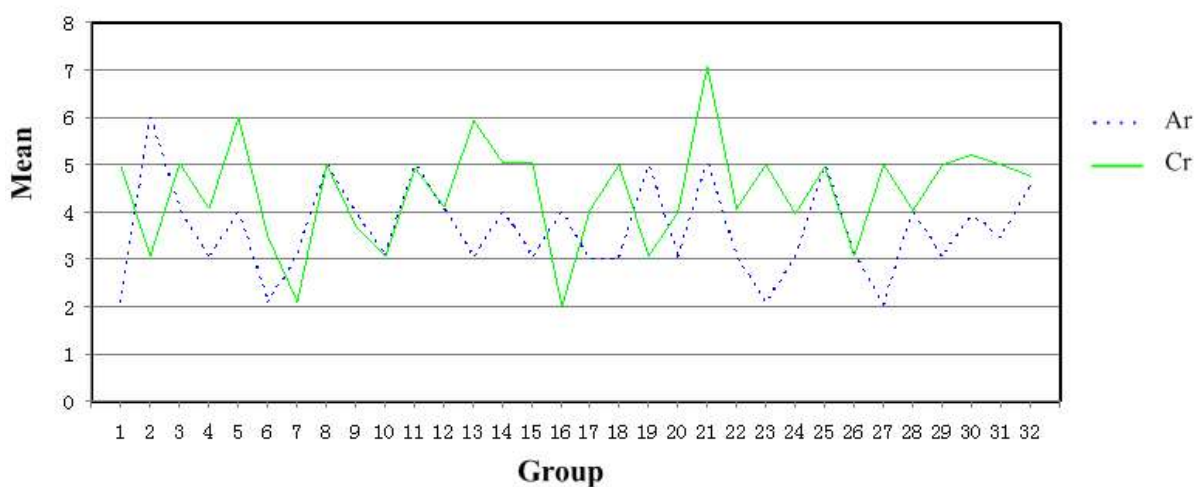|		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
									Lower	Upper	
Lv	Equal Variances Assumed	3.8515	.506	4.669	62	.002	4.225	.908	2.425	6.046	
	Equal Variances not Assumed			4.669	50.053	.002	4.225	.908	2.425	6.046	

Fig 4: Richness degree of exemplifiers

To conclude, through statistical analysis of SPSS, some conclusions can be arrived: density and richness degree of using connectors by Chinese debaters is lower than that by Americans; density and richness degree of using exemplifiers by Chinese is higher than that by Americans.

TABLE VI. Independent sample t test of richness degree of exemplifiers

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | 95% Confidence Interval of the Difference | |
		F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference		Lower	Upper
Lv	Equal Variances Assumed	.034	.854	-2.855	62	.004	-.700	.260		-1.362	-2.38
	Equal Variances not Assumed			-2.855	61.616	.004	-.700	.260		-1.362	-2.38

Based on above conclusions, some pedagogical implication on debate teaching can be achieved:

(1) English teachers in China should teach learners to use more and richer connectors accurately and appropriately to enhance debating speeches' coherence, transition and unity.

(2) English teachers in China should encourage learners to reduce the density and improve richness degree of using exemplifiers so that their speeches could be more truthful, native-like and authentic.

(3) Chinese debaters should read more authentic English reading materials and pay more attention to connectors and exemplifiers application by native speakers.

(4) Chinese debaters should reduce their examples taking in debate. Instead, they should enhance their ability of logical argument and analytical illustration so that their debate can be more convincing and powerful.

VI. CONCLUSION

Based on two self-built corpuses, this paper adopts computer technologies to conduct a contrastive study on lexical bundles in debating speeches by Chinese and American debaters. On the basis of encoding lexical bundles by CLAWS, connectors and exemplifiers in the two corpuses are searched through software Concapp. Then, by adopting software SPSS, the mathematical and statistical analysis of using connectors and exemplifiers is carried out from two perspectives: density and richness degree. Based on the above statistical analysis, some discoveries are achieved: density and richness degree of connectors used by Chinese debaters is lower than that of connectors used by American debaters; by contrast, density and richness degree of exemplifiers used by Chinese debaters is higher than that of exemplifiers used by American debaters.

To conclude, the research results reveal obvious difference on using connectors and exemplifiers in English debate by Chinese and American debaters. It is hoped that this research can provide some valuable pedagogical implication on debate teaching and insights for enhancing Chinese debating quality, which in turn, can greatly enhance overall language application skills of English learners in China.

ACKNOWLEDGEMENTS

REFERENCES

[1] Kwak Myunsun and Sangcheol Lee (2020) English Politeness Strategies in the American Television Presidential Debates: Focusing on the Presidential Election in 2016. The Journal of Studies in Language 35:485-502

[2] Young Chang Sun and Kim Hyejin (2019) Quantitative Analysis of the Reflection Style and Learning Outcomes and Qualitative Analysis of Learning Effects and Further Suggestions for Discussion &Debate-based Learning in English Literature Classes of Higher Education. The Journal of Learner-Centered Curriculum and Instruction 19:601-628

[3] Carinanos-Ayala Susana, Arrue M and Zarandona, J (2021) The use of structured debate as a teaching strategy among undergraduate nursing students: A systematic review. NURSE EDUCATION TODAY 98:104766

[4] Hawkins W Anthony, Fulford Michael and Phan Stephanie V. (2019) Using debates as the primary pedagogy to teach critical care in a PharmD curriculum elective course. Currents in pharmacy teaching & learning 11:943-948.

[5] Praharaj, SK and Bhandary, RP (2019) Igniting minds: Debates enhance engagement in small group teaching for postgraduates 34:188-189.

[6] Kang Xu and Beishui Liao(2021) Characterizing Argumentation Frameworks with an Extension. Studies in Logic 14:41-67

[7] Zhe Yu, Kang Xu and Beishui Liao (2018) Structured Argumentation: Restricted Rebut vs. Unrestricted Rebut. Studies in Logic 3:3-18

[8] Shang Biwu (2017) Current Issues and Critical Debates in World Literature and Comparative Literature Studies: An Interview with Professor Thomas Oliver Beebee. Foreign Literature Studies 39:1-8

[9] WANG Yan (2018) Comparative Studies Between NEST and NNEST. Heilongjiang Researches on Higher Education 36:23-30

[10] Rachel Allan (2016) Lexical Bundles in Graded Readers: To What Extent does Language Restriction Affect Lexical Patterning? System 59:61-72

[11] Lee Hye-Kyung(2020) Lexical bundles in linguistics textbooks. Linguistic Research 37: 121-145

[12] Sung-Min Lee and Daehyeon Nam (2019) Russian Lexical Bundles in Spoken and Written Registers. 31: 51-84

[13] Jihyun Ahn (2020) A Corpus-Based Classification of Lexical Bundles in Business Email. Journal of Language Sciences 27:147-166

[14] Daehyeon Nam(2020) Functional Distribution of Lexical Bundle in Native and Non-Native Students' Argumentative Writing. Journal of Asia TEFL 14:703-716

[15] Alasmary Abdullah A. (2018) Participant-oriented Lexical Bundles in Graduate-level Math Textbooks: A Corpus-based Exploration. Corpus Linguistics Research 4:99-114

[16] Hong Ji-yoon (2020) Connecting Lexical Bundles and Moves by Korean EAP Students at Different Academic Levels 36:3-31

[17] Kim Eun-Soo and Lee Eun-Joo (2020) A Corpus-Based Study of Lexical Bundles and Moves by English L1 and L2 Writers in Medical Journal Abstracts 20:768-800

[18] Liang Xiaolei and Jhang Se-Eun (2022) Functional Analysis of Lexical Bundles and Rhetorical Moves of L1 and L2 Research Article Abstracts in Applied Linguistics 29:289-314

[19] Kim Eun-Soo and Lee Eun-Joo (2021) A Corpus-Based Analysis of Lexical Bundles between English L1 and L2 Writers in Medical Journal Abstracts: Focusing on Structures and Functions 18:142-160

[20] Garrido MR (2022) The evolution of language ideological debates about English and French in a multilingual humanitarian organisation. LANGUAGE POLICY 21:47-73

[21] Zhang Yi(2020) On the Dilemma and Breakthrough of College English Linguistics Teaching from the Perspective of Ecological Linguistics. BASIC & CLINICAL PHARMACOLOGY & TOXICOLOGY 127:221-221

[22] Yu Hana and Oh Sun-Young (2020) Stance Markers in Expert versus Novice English Academic Writing in Applied Linguistics. Korean Journal of Applied Linguistics 36:83-108

[23] Hinh Duc Nguyen(2016) Metaphors as a Window into Identity: A study of Teachers of English to Young Learners in Vietnam. System 60:66-78

[24] Nattinger, J and J Decarrico (1992) Lexical Phrases and Language Teaching. Oxford: Oxford University Press ISBN9780194371643

[25] Jia Liu, Lina Han (2015) A Corpus-based Environmental Academic Word List Building and Its Validity Test. English for Specific Purposes 39:1-11

[26] Khadim Dramé, Gayo Diallo (2014) Reuse of Termino-ontological Resources and Text Corpora for Building a Multilingual Domain Ontology: An application to Alzheimer's Disease. Journal of Biomedical Informatics 48:171-182

[27] Hossain, M M (2022) Applied univariate, bivariate, and multivariate statistics: Understanding statistics for social and natural scientists, with applications in SPSS and R, 2nd edition. JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES A-STATISTICS IN SOCIETY 185:727-728

[28] Van Wingerde Bart and Van Ginkel Joost (2021) SPSS Syntax for Combining Results of Principal Component Analysis of Multiply Imputed Data Sets using Generalized Procrustes Analysis. Applied psychological measurement 45:231-232