

Lexical Bundles in Argumentative and Narrative Writings by Chinese EFL Learners

Yanfeng Yang¹

¹ School of Foreign Studies, Lingnan Normal University, Guangdong Province, China

Correspondence: Yanfeng Yang, School of Foreign Studies, Lingnan Normal University, No.29 Cunjin Road, Chikan District, Zhanjiang City, Guangdong Province, China, 524048. Tel: 1-343-035-7003. E-mail: anna01124@hotmail.com

Received: January 5, 2017 Accepted: January 26, 2017 Online Published: February 9, 2017

doi:10.5539/ijel.v7n3p58

URL: <http://dx.doi.org/10.5539/ijel.v7n3p58>

Abstract

Previous studies have shown that lexical bundles are important building blocks of discourse and a significant component of fluent linguistic production. However, little research was found to investigate lexical bundles in narrative writings, a basic text type on which the other text types (discourses) build upon. The present study tries to fill the gap and investigates lexical bundles in argumentative and narrative writings by Chinese EFL learners. The lexical bundles were retrieved by *kwNgram* and then manually refined and classified into structural and functional categories respectively based on Biber et al.'s (1999) and Biber et al.'s (2003) frameworks. The findings show that (1) students used much more four-word bundles in argumentative writings than those in narrative writings; (2) no big difference was found in the structural patterns of the four-word lexical bundles used by the students across the two text types; (3) students relied much more on stance bundles than the other functional types of bundles in their argumentative writings, while they turned to referential expressions other than stance bundles or discourse organizers in their narrative writings. The functional purposes of various discourses explain the students' selection of different functional patterns across the text type.

Keywords: lexical bundles, corpus study, Chinese EFL learners, narrative writing, argumentative writing

1. Introduction

Lexical bundles are recurrent sequences of words, which have been studied under many rubrics, including "lexical phrases", "formulas", "routines", "fixed expressions", "pre-fabricated patterns", "n-grams", and "clusters" (Biber, 2006; Biber & Barbieri, 2007). In addition to high frequency, another significant feature of lexical bundles is the important role they play in discourse construction. Biber, Conrad, & Cortes (2004) noted that "they [lexical bundles] are important building blocks of discourse, associated with basic communicative purposes" (p. 400). Hyland (2012) also pointed out that lexical bundles are a significant component of fluent linguistic production. High frequency of lexical bundles is not by chance and needs explanation (Biber et al., 2004). Therefore, to reveal the discourse functions of the lexical bundles becomes the major task in research of this domain. Previous studies revealed similarities and differences between lexical bundles across different registers, among which studies on academic writing are the most fruitful (Hyland, 2012). However, little research was found to investigate lexical bundles in narrative writings, a basic text type on which the other text types (discourses) build upon. The present study tries to fill the gap and investigates lexical bundles in argumentative and narrative writings by Chinese EFL learners. The primary purpose of the study is to find out the similarities and differences between the lexical bundles used in the two text types so as to reveal the discourse functions within each text type and shed some light on pedagogy.

2. Literature Review

2.1 Definition of Lexical Bundles

According to Biber, Johansson, Leech, Conrad, & Finegan (1999), lexical bundles are "recurrent expressions regardless of their idiomaticity, and regardless of their structural status" (p. 990). In operation, lexical bundles are identified using a frequency-driven approach. The frequency cut-off is somewhat arbitrary, ranging from 10 (Biber et al., 1999; Biber, 2006) to 20 (Cortes, 2004; Hyland, 2008a, 2008b) to 40 times per million (Biber & Barbieri, 2007). In general, the higher the frequency cut-off is, the more representative the lexical bundles are

and thus have greater significance for investigation. Another defining feature for lexical bundles is that they should be used in multi-texts, normally five or more different texts, in order to guard against idiosyncratic uses by individual speakers or writers (ibid).

2.2 Characteristics of Lexical Bundles

According to Biber & Barbieri (2007), three major characteristics distinguish lexical bundles from other kinds of formulaic expressions. First, lexical bundles are by definition extremely common. Second, most lexical bundles are not idiomatic in meaning and not perceptually salient. Third, lexical bundles usually do not represent a complete structural unit and most lexical bundles bridge two structural units, usually two clauses in speech and two phrases in writing. Although lexical bundles are neither idiomatic nor structurally complete, they are important building blocks in discourse and provide interpretive frames for the developing discourse (ibid).

According to Biber et al. (2004), lexical bundles have three primary discourse functions: (1) stance expressions (2) discourse organizers, and (3) referential expressions. Stance bundles “express attitudes or assessments of certainty that frame some other proposition” (ibid, p. 384). They can be categorized into five sub-categories: epistemic (e.g., *I don't know what, I think it was, the fact that the*), desire (e.g., *if you want to, what do you want*), obligation (e.g., *I want you to, you have to*), intention/prediction (e.g., *I'm not going to, it's going to be*), and ability (e.g., *to be able to, can be used to*). Discourse organizers “reflect relationships between prior and coming discourse” (ibid). They have two categories: topic introduction (e.g., *what do you think, if you look at, I would like to*) and topic elaboration/clarification (e.g., *has to do with, on the other hand*). Referential bundles “make direct reference to physical or abstract entities, or to the textual context itself, either to identify an entity or to single out some particular attribute of the entity as especially important” (ibid). Four major sub-categories are distinguished: referential identification/focus (e.g., *this is one of the, of the things that*), imprecision indicators (e.g., *or something like that, and things like that*), specification of attributes (e.g., *there's a lot of, the size of the, in terms of the*), and time/place/text reference (e.g., *in the United States, at the time of, at the end of*). As claimed by Biber et al., lexical bundles “can be regarded as structural “frames”, followed by a “slot”. The frame functions as a kind of discourse anchor for the “new” information in the slot, telling the listener/reader how to interpret that information with respect to stance, discourse organization, or referential status.” (ibid, p. 399)

2.3 Studies on Lexical Bundles

Biber and his colleagues made great contributions to the studies on lexical bundles. In an earlier study, Biber & Conrad (1999) found that even though only 15% of the lexical bundles present in conversation are recognized as complete units, their analysis is relevant to how language functions. Biber et al. (1999) identified the most frequent lexical bundles in academic prose and conversation based on Longman Spoken and Written English Corpus. They developed the structural categories of those bundles, by which they made comparison of the bundles across registers. Results indicate that “most of bundles in conversation are building blocks for verbal and clausal units, while most lexical bundles in academic prose are building blocks for extended noun phrases or prepositional phrases” (ibid, p. 992). In a subsequent study, Biber et al. (2003) developed a preliminary taxonomy to classify the functional patterns of lexical bundles based on the most frequent bundles found in Biber et al. (1999). Four core categories were identified: stance bundles, discourse organizers, referential bundles, and interactional bundles. Building upon the structural and functional categories developed earlier, Biber et al. (2004) investigated the use of lexical bundles in university classroom teaching and textbooks. Later, they (Biber et al., 2007) extended their study of the use of lexical bundles in a wide range of spoken and written university registers, including both instructional registers and student advising/management registers (e.g., office hours, class management talk, written syllabi, etc.). The findings show that lexical bundles are even more prevalent in non-academic university registers than they are in the core instructional registers. Contrary to previous research finding that bundles were much more common in speech than in writing (Biber et al., 2004), lexical bundles in their research are very common in written course management (e.g., course syllabi). From the above studies, we can see that the specific categories of lexical bundles were adapted according to the bundles generated in certain registers. Interactional bundles are seldom used in university written register.

Based on Biber et al.'s framework of categorizing lexical bundles into structural and functional categories, Cortes (2004) investigated the use of lexical bundles by professional and student writers and found that the bundles used by the students did not correspond to those employed by the professional authors and some bundles frequently occur in published articles were never used by the learners at all. Similarly, Hyland (2008a, 2008b) made a comparison of the use of lexical bundles in published articles and that in high graded master's theses and doctoral dissertations by L2 writers in Hong Kong. He found that the postgraduates used more bundles than the published authors, suggesting the difference between postgraduate genres and the published one (Hyland, 2012).

Later, Amirian, Ketabi, & Eshaghi (2013) conducted a study on the use of lexical bundles in MA theses of applied linguistics by native (English) and non-native (Iranian) post-graduate writers. Significant differences were found between native and Iranian students in the frequency of lexical bundles used and their structural and functional patterns. Iranian postgraduates used more lexical bundles than native counterparts and even more than Chinese students in Hyland's (2008a) study. In terms of structural patterns, Iranian students prefer to use clausal bundles over the native English students. For functional patterns, native English postgraduates show more variety than the Iranian EFL students. More recently, Pan, Reppen, & Biber (2016) narrowed down their research scope to investigate the use of lexical bundles in the published research articles in one single academic discipline, i.e., Telecommunications, by L1 English professional versus L2 English professionals of Chinese. Major structural differences were found between the two groups of expert writers. Bundles consisting of noun phrase and prepositional phrase fragments were preferred by L1 professionals while those consisting of verbs and clause fragments, especially passive verb structures, were used mostly by the L2 professionals. No big difference was found in the functional types of the bundles between the two groups. But L2 professionals used more stance-oriented bundles than their L1 counterparts and some bundles were misused. Different from the previous studies which mainly focused on the frequency of the bundles used, Huang (2015) investigated the accuracy as well as the frequency of the bundles used in the essay writing by Chinese EFL learners. It was found that senior students used the bundles more frequently and with wider variety than junior students but no significant difference was revealed in their accuracy between the two groups.

From the above literature, we can see that the use of lexical bundles varies across different registers, for example, different genres and proficiency levels (Note 1). As stated by Biber et al. (2007), "the overall importance of multi-word units [lexical bundles] in discourse can be fully understood only by undertaking empirical research studies from different perspectives" (p. 372). However, little research was found to investigate the bundles used in the narrative writing, nor the comparative studies between it and the other types of writing, e.g., argumentative writing. The present study tries to fill the gap and aims to find out the bundle patterns used by the Chinese EFL learners in the two types of writings. I will adopt a data-driven approach to select lexical bundles and analyze the bundles based on the structural and functional categories developed by Biber et al. (1999) and Biber et al. (2003). To make the study more focused, only four-word bundles will be chosen for analysis since they are more frequent than five-word bundles and incorporate most of the three-word bundles (Biber et al., 1999). Moreover, four-word bundles present more range of functions and structures than three-word clusters (Hyland, 2008b).

3. Methodology

3.1 Research Questions

The present study addresses the follow research questions.

- (1) What is the overall use of four-word lexical bundles in argumentative and narrative writings by Chinese EFL learners?
- (2) Do the structural patterns of four-word lexical bundles used by Chinese EFL learners in their argumentative writings differ from those in their narrative writings?
- (3) Do the functional patterns of four-word lexical bundles used by Chinese EFL learners in their argumentative writings differ from those in their narrative writings?

3.2 Data Selection

The data for the present study was selected from WECCL (Written English Corpus of Chinese Learners), a sub-corpus of SWECCCL (Spoken and Written English Corpus of Chinese Learners) compiled by Wen et al. (2005). The compilation of SWECCCL is a state-sponsored project of social sciences and it is the first and also the biggest corpus so far in Mainland China consisting of spoken and written corpora by university undergraduate English majors ranging from Grade 1 to Grade 4. The English proficiency for these students can be considered intermediate-to-advanced according to the overall English proficiency of the Chinese EFL learners in mainland China. In order to maintain comparability with the international learner corpora, SWECCCL research team strived to conform to the criteria of corpus design stipulated by Granger (1998). Two sub-corpora can be found in WECCL. One is named "raw data" and the other "tagged data", the former of which consists of another four sub-corpora, i.e., "argumentation", "essays by conditions", "narration" and "years 1-4 essays". The "argumentation" corpus contains timed argumentative writings while "narration" timed narrative writings. The files of the other two sub-corpora, i.e., "essays by conditions" and "years 1-4 essays", were not categorized according to their types of writing. Therefore, "argumentation" and "narration" were selected for the purpose of the present study. All the plain texts in the sub-corpus of "narration" were chosen, which are comprised of 529

pieces of timed narrative writings with 153,859 words in total and involve English majors ranging from Grade 1 to Grade 4. In order to make the two sets of data comparable, the timed argumentative writings were randomly selected from each grade in the sub-corpora of “argumentation”. Since the argumentative writings by the Grade 4 students are fewer than the other three grades, all the 60 plain texts were selected. Moreover, 130 texts from each of the other three grades were chosen in order to balance the number of texts among the three grades and make the narration and the argumentation data sets comparable to each other in terms of their respective total amounts of words. Finally, the argumentative corpus for the present study consists of 450 plain texts with 151,782 words, close to the total amount of words for narrative writings. Table 1 illustrates a general picture of the data for the present study. The mean length of the texts in argumentation (337 words per text) is longer than that in narration (291 words per text). All the argumentative plain texts were merged into one text file and the same was done for the narrative plain texts. The two files served as the two corpora for the present study and were put into further analysis.

Table 1. Description of the data selected for the present study

		Argumentation	Narration
Number of texts	Grade 1	130	91
	Grade 2	130	257
	Grade 3	130	91
	Grade 4	60 (Note 2)	90
Total number of texts		450	529
Total number of words		151,782	153,859
Mean length of texts		337	291

3.3 Retrieval and Identification of Lexical Bundles

KfNgram 1.2.03 published by William H. Fletcher was used to generate lexical bundles in the two corpora respectively. It is a free software for researchers to generate n-grams, also known as lexical bundles, from the text or HML file (see <http://www.kwicfinder.com/kfNgram/kfNgramHelp.html>). What the researcher needs to do is to add the source file and type the number of words for the n-gram (e.g., type “4” for four-word bundles), then click “get wordgrams” under the tap of “tools” and the lexical bundles will be generated within seconds. As mentioned earlier, only four-word bundles were taken into account in the present study. Appendices A and B are the screen shots of the software and the bundles generated for the argumentative writings. The frequencies generated by the software were the raw frequencies and needed to be transferred into norm-referenced frequencies for better comparison with the other studies in this line. The raw frequencies were copied into excel file and calculated for the normed rates of frequency. For example, the bundle of “with the development of” occurred 62 times (raw frequency) in the argumentative writings. When it was transferred into the normed rate of frequency, it turned to 408 times per million words. Another defining feature of lexical bundles is their multi-text (normally at least 4 texts) occurrence. Since the texts in the present study are short and the repetition of the bundles would affect the variety of the article, normally students would avoid repeating the bundles very often in one text. Therefore, most bundles in the present study are widely distributed. The least common bundles in the data of argumentative writings occur in more than four texts, while the more common bundles are distributed more widely.

Moreover, a conservative cut-off point of 40 times per million words was adopted, that is, only those bundles used more than 40 times per million words were selected for further analysis. Some high-frequency bundles were excluded from analysis because they are topic related, e.g., “the most unforgettable event”, “the use of technology”, “gap between parents and”, “between parents and children”, etc. In addition, bundles such as “when I was #” and “was # years old” were also eliminated because “#” can have various tokens. In other words, I followed a conservative sense of token frequency in this study. After elimination, 194 four-word bundles were left for narrative writings and 489 for argumentative writings.

4. Results and Discussion

4.1 Overall Use of Lexical Bundles

Table 2 illustrates the thirty most frequent four-word bundles in each corpus.

Table 2. Thirty most frequent four-word bundles in each corpus

Argumentation			Narration	
1	with the development of	408	at that time i	247
2	more and more people	376	when i was in	188
3	between parents and children	369	i will never forget	136
4	the best way to	343	in my life i	136
5	as a lifelong process	323	for a long time	123
6	education as a lifelong	296	the first time i	110
7	impact of the internet	296	from then on i	104
8	on the other hand	290	for the first time	97
9	of the internet on	283	when i was a	91
10	some people think that	283	to take part in	91
11	more and more popular	270	i didn't know how	84
12	the impact of the	257	the top of the	84
13	convenient fast and inexpensive	231	at the same time	78
14	at the same time	224	to go to school	78
15	become more and more	217	for a while and	78
16	in the first place	211	didn't know how to	71
17	communicate with each other	191	will never forget it	71
18	gap between parents and	178	is one of the	71
19	way to deal with	171	was the first time	71
20	generation gap between parents	165	i went to the	71
21	personal friendly and valuable	165	one of my friends	71
22	the development of the	165	all of a sudden	71
23	to communicate with their	165	i don't want to	65
24	more personal friendly and	152	get to the top	65
25	a generation gap between	145	is an unforgettable event	65
26	has got a cancer	145	if you want to	65
27	problems in the first	145	in order not to	65
28	the problems in the	145	how to deal with	65
29	becoming more and more	138	that time i was	65
30	best way to deal	138	in my life is	65

Notes. Numbers within "Argumentation" and "Narration" are normed rates of frequency (# times per million words) for each bundle.

Obviously, students used much more bundles in argumentations than in narrations. This result indicates that students consider the argumentative texts much more highly structured than the narration, which may encourage them to use the patterns of bundles to help them express their opinions. However, their heavy reliance on the bundles in argumentative writings also indicates that the learners cannot express their ideas in argumentations as freely as that in narrations. Looking into the bundles used in the argumentative writings, we can find that the bundles contained "more and more" were heavily used by the learners (e.g., "more and more people", "more and more popular", "become more and more"). The result is consistent with Huang's (2015) study on Chinese EFL essay writings. It may be due to the L1 transfer. In Chinese, people often use *yue lai yue* (meaning "more and more" in English) in their daily life. As noted by Paquot (2013), "the more frequent a lexical bundle is in the learners' mother tongue, the more likely learners are to use its congruent form in the foreign language" (p. 410).

4.2 Structural Patterns

Since the numbers of the four-word bundles generated from the two corpora were imbalanced, only 194 most frequent bundles were selected from argumentations for better comparison with those bundles from narrations. Finally, 388 four-word bundles, i.e., 194 for each corpus, were classified into structural and functional categories.

Biber et al.'s (1999) structural categories for classifying lexical bundles were modified into the following seven general categories for the present study: (1) verb phrase expressions, including "pronoun/noun + be (+...)", "anticipatory *it* + verb phrase/adjective phrase", "passive verb + prepositional phrase fragments", and "copula *be* + noun phrase/adjective phrase"; (2) dependent clause expressions, including "(verb phrase +) *that*-clause fragment", "(verb/adjective +) *to*-clause fragment", and "adverbial cause fragment"; (3) noun phrase expressions, including "noun phrase with *of*-phrase fragment" and "noun phrase with other post-modifier fragment"; (4) prepositional phrase expressions, including "prepositional phrase with embedded *of*-phrase" and "other prepositional phrase (fragment)"; (5) quantifier expressions; (6) adjectival expressions (e.g., *more and more*

important); (7) unclassifiable fragments. Table 3 shows the structural distribution of four-word lexical bundles across text types.

Table 3. Structural distribution of lexical bundles across text types

Structural category	Argumentation		Narration	
	Frequency	Percentage (%)	Frequency	Percentage (%)
(1) verb phrase expressions	95	49.0	87	44.8
(2) dependent clause expressions	31	16.0	30	15.5
(3) noun phrase expressions	38	19.6	32	16.5
(4) prepositional phrase expressions	21	10.8	37	19.1
(5) quantifier expressions	1	.5	1	.5
(6) adjectival expressions	3	1.5	2	1.0
(7) unclassifiable fragments	5	2.6	5	2.6
Total	194	100.0	194	100.0

From Table 3, we can see that the use of lexical bundles does not show any big difference across the text types. Students mainly rely on verb phrase expressions, dependent clause expressions, noun phrases expressions, and prepositional phrase expressions in their writings. This result also conforms to that found in the previous studies (e.g., Biber et al., 2004; Huang, 2015; Pan et al., 2016). The reason is that those expressions are the basic building blocks in constructing a sentence as well as a text.

4.3 Functional Patterns

Based on Biber et al.'s (2003) functional classification of lexical bundles, the following adaptations were made. The category of "interactional bundles" was not taken into account because they were not found in the present corpora. Moreover, some content bundles such as "communicate with their parents", "of the internet on", etc., were excluded from further analysis because they do not serve a certain function. After that, 64 bundles in the argumentation corpus and 88 bundles in the narration corpus were classified into three functional categories: stance bundles, discourse organizers, and referential expressions. Table 4 illustrates the frequency and the percentage of each category within the text type.

Table 4. Functional distribution of lexical bundles across text types

Functional category	Argumentation		Narration	
	Frequency	Percentage (%)	Frequency	Percentage (%)
Stance bundles	42	65.6	23	26.1
Discourse organizers	5	7.8	4	4.5
Referential expressions	17	26.6	61	69.3
Total	64	100.0	88	100.0

From Table 4, we can see that the uses of discourse organizers are more or less the same in argumentative and narrative writings. However, a sharp contrast can be found in the uses of stance bundles and referential expressions across the text types. In argumentation, 42 stance bundles were used, accounting for 65.6% of the functional bundles in this corpus. In narration, only 23 stance bundles were found, the percentage of which (i.e., 26.1%) is almost 40% lower than that in argumentation. Contrarily, the dominant type of bundles used in the narration is referential expressions, 61 in total and accounting for nearly 70% of the functional bundles in this text type, which is much higher than its counterpart in argumentation (17 in total and 26.6%). The findings indicate that students relied much more on stance bundles than the other functional types of bundles in their argumentative writings. However, in the narrative writings, they turned to referential expressions other than stance bundles or discourse organizers. Table 5 illustrates the 30 most frequent functional four-word bundles in each text type.

Table 5. Thirty most frequent functional four-word bundles in each corpus

Argumentation			Narration	
1	with the development of	408	at that time i	247
2	more and more people	376	when i was in	188
3	on the other hand	290	i will never forget	136
4	some people think that	283	for a long time	123
5	more and more popular	270	the first time i	110
6	at the same time	224	from then on i	104
7	in the first place	211	for the first time	97
8	best way to deal	138	when i was a	91
9	in my opinion the	132	the top of the	84
10	it is difficult to	125	i didn't know how	84
11	think the best way	112	at the same time	78
12	difficult to communicate with	112	for a while and	78
13	for a long time	112	is one of the	71
14	in the western world	112	was the first time	71
15	speaking is more important	105	one of my friends	71
16	to tell the truth	105	all of a sudden	71
17	one of the most	105	didn't know how to	71
18	is difficult to communicate	99	will never forget it	71
19	is the most important	99	that time i was	65
20	that it is difficult	99	i don't want to	65

Notes. Numbers within "Argumentation" and "Narration" are normed rates of frequency (# times per million words) for each bundle.

The students' choices of the functional bundles are actually determined by the primary communicative functions of the text types. It is difficult to find a pure text type consisting of only one function such as purely narrative or persuasive since in every piece of writing, the writer will usually employ more than one function to achieve its purpose. For example, the use of narrative functions in argumentation lays a good background for persuasion purpose. However, the primary functions of the text types are distinct from each other, i.e., the primary function of argumentation is persuasive and that of narration is narrative. According to Jackson & Stockwell (2011),

Discourses and texts with a persuasive function aim to convince a hearer or reader that something is true, or that an opinion is the correct one, or that a course of action is the right one. Persuasive texts will provide arguments and evidence for a particular point of view. The text will generally be carefully structured, with a series of points that lead to a logical conclusion. (p. 85)

For narrative function, they explained that

Discourses and texts with a narrative function are used to tell a story. Typically, they show progression through time; they are in the past tense; and there is explicit reference to the passing of time (next week, the following year, after that). Such time expressions are often used to structure the unfolding story. (ibid, p. 83)

As noted earlier, stance bundles express attitudes or assessments of certainty that frame some other proposition (Biber et al., 2004), which serve the persuasive function of a text. The reason why students rely on stance bundles is that they used them to achieve the persuasive function of the text, the primary aim of the argumentation. Contrarily, the primary function of narration is to tell a story. Referential expressions, which "make direct reference to physical or abstract entities, or to the textual context itself, either to identify the entity or to single out some particular attribute of the entity as especially important" (ibid, p. 384), are good choices for the writers to achieve the narrative function. Therefore, the present study adds evidence to the claim that lexical bundles "are important building blocks of discourse, associated with basic communicative purposes" (Biber et al., 2004, p. 400).

In order to have a deeper understanding of the students' uses of stance bundles and referential expressions, I divided the two functional categories into their sub-categories based on Biber et al. (2003). For the stance bundles, the category of "epistemic stance" in Biber et al.'s (ibid) taxonomy was adapted into three sub-categories: personal (first person), personal except first person, impersonal. Table 6 illustrates the distribution of stance bundles across text types.

Table 6. Distribution of stance bundles across text types

Types of stance bundles		Argumentation	Narration	Example
Epistemic stance	Personal (first person)	6 (14.3%)	4 (17.4%)	i think it is, i don't know how
	Personal except first person	13 (31.0%)	0	some people believe that, some people think that
Attitudinal/modality stance	Desire (personal)	1 (2.4%)	6 (26.1%)	i didn't want to, if you want to
	Obligation/directive (personal/impersonal)	11 (26.2%)	3 (13.0%)	it is difficult to, try their best to
Intention/prediction		10 (23.8%)	8 (34.8%)	i decided to go, try my best to
Ability (personal/impersonal)		1 (2.4%)	2 (8.7%)	it can be used
Total		42 (100.0%)	23 (100.0%)	

Notes. For each number column, the number outside the brackets is the frequency of the occurrence of the bundle type and that inside the brackets is the percentage of the bundle type within the text type.

From Table 6, we can see that in argumentative writings, bundles of epistemic stance account for nearly half (45.3%) of the stance bundles in this text type, followed by bundles of attitudinal/modality stance (28.6%) and those of intention/prediction (23.8%), and least with bundles of ability (2.4%). However, in narrative writings, students rely more on bundles of attitudinal/modality stance (38.1%), followed by bundles of intention/prediction (34.8%) and epistemic stance (17.4%), and least with bundles of ability (8.7%). By comparing the uses of bundles in the specific categories across the text types, we can find that students prefer to use non-first-person bundles such as “some people believe that”, “some people think that”, “it is difficult” etc., in argumentative writings in order to make their points sound objective and persuasive. In contrast, they favored first-person bundles such as “i think it is”, “i don't know how”, “I decided to go” etc., in their narrative writings to express their personal ideas. For the referential expressions, I also divided them into several subcategories based on Biber et al. (2003) (see Table 7).

Table 7. Distribution of referential expressions across text types

Types of referential expressions		Argumentation	Narration	Example
Identification/focus		1 (5.9%)	6 (9.8%)	is one of the, one of my best
Specification of attributes		11 (64.7%)	1 (1.6%)	become more and more, more and more people
Time/place/text references	Place reference	2 (11.8%)	17 (27.9%)	top of the ladder, in front of me
	Time reference	3 (17.6%)	30 (49.2%)	when i was in, the first time i
	Text deixis	0	2 (3.3%)	in my life, in our daily life
	Multi-functional reference	0	5 (8.2%)	at end of, in the middle of
Total		17 (100%)	61 (100%)	

Notes. For each number column, the number outside the brackets is the frequency of the occurrence of the bundle type and that inside the brackets is the percentage of the bundle type within the text type.

From Table 7, we can see that students predominantly used time/place/text references among the various types of referential expressions to help them set the time and the place and achieve the coherence in telling the story in narrative writings. Contrarily, students employed bundles under the category of “specification of attributes” quantify specifications and frame attributes in argumentative writings.

5. Major Findings and Future Study

To answer the three research questions in Section 3.1, the following major findings can be gained. Firstly, students used much more four-word bundles in argumentative writings than those in narrative writings. The bundles contained “more and more” were heavily used in the argumentation, which is consistent with the findings in the previous studies on Chinese EFL essay writings (e.g., Huang, 2015). L1 transfer is the possible reason, that is, learners prefer to use the equivalent forms conforming to their mother tongue. Secondly, no big difference was found in the structural patterns of the four-word lexical bundles used by the students between the two text types. This result also conforms to that found in the previous studies (e.g., Huang, 2015; Biber et al., 2004) because the phrases such as verb and noun phrases are the basic and important building elements for a text. Thirdly, the functional patterns of four-word lexical bundles used by the students in their argumentative writings differ greatly from those in their narrative writings. Students relied much more on stance bundles than the other functional types of bundles in their argumentative writings, while they turned to referential expressions other

than stance bundles or discourse organizers in their narrative writings. These discrepancies can be explained by the different purposes the two text types used to achieve. For argumentation, the major purpose is to express one's viewpoints over certain events and the stance bundles can help achieve the purpose; however, narration is mainly used to describe an event, a person, a place or thing, for which the writer needs to refer to the time, place or thing from time to time. Students relied on the different functional patterns of bundles in different text types to help them achieve the functional or communicative purposes of each text type. The result adds evidence to the claim that lexical bundles "are important building blocks of discourse, associated with basic communicative purposes" (Biber et al., 2004, p. 400).

The present study can shed some lights on ESL or EFL teaching. On one hand, for those bundles heavily used by the students, teachers should try to help students distinguish the different uses between the L2 learners and the native speakers so that the students can avoid overuse of the bundles. On the other, it was found that the narrative writings and the argumentative writings differ greatly in the functional patterns of the four-word bundles. Narration is a basic text type but little research was conducted to investigate the language patterns of it, let alone the learners' performance in this text type. The result of the present study may arouse the researchers' attention to this text type, especially by the novice learners at the beginning stage of their language learning.

The present study is limited in its relatively small corpora. Future studies can extend the research by using larger corpora and incorporate other text types such as description and exposition into the study. Moreover, studies on the lexical bundles used by expert and novice or native and non-native writers need further exploration. Accuracy of the bundles used by the learners is another topic needed to be further explored.

Acknowledgments

The article was supported by the Youth Research Project of Humanities and Social Sciences Sponsored by the Ministry of Education in China (Project No.: 15YJC740118) and the research project sponsored by Lingnan Normal University (Project No.: YW1503)

References

- Amirian, Z., Ketabi, S., & Eshaghi, H. (2013). The use of lexical bundles in native and non-native post-graduate writing: The case of applied linguistics MA theses. *Journal of English Language Teaching and Learning*, 11, 1-29.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam, the Netherlands: John Benjamins. <http://dx.doi.org/10.1075/sc1.23>
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263-286. <http://dx.doi.org/10.1016/j.esp.2006.08.003>
- Biber, D., Conrad, S., & Cortes, V. (2003). Towards a taxonomy of lexical bundles in speech and writing. In A. Wilson, P. Rayson, & T. McEnery (Eds.), *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech* (pp. 71-92). Frankfurt, Germany: Peter Lang.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25, 371-405. <http://dx.doi.org/10.1093/applin/25.3.371>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow, UK: Pearson.
- Breeze, R. (2013). Lexical bundles across four legal genres. *International Journal of Corpus Linguistics*, 18, 229-253. <http://dx.doi.org/10.1075/ijcl.18.2.03bre>
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397-423. <http://dx.doi.org/10.1016/j.esp.2003.12.001>
- Granger, S. (1998). *Prefabricated patterns in advanced EFL writing: Collocations and formulae*. In *Phraseology: Theory, Analysis and Application* (pp. 145-160). Oxford: Clarendon Press.
- Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, 13-23. <http://dx.doi.org/10.1016/j.system.2015.06.011>
- Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18, 41-62. <http://dx.doi.org/10.1111/j.1473-4192.2008.00178.x>
- Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4-21. <http://dx.doi.org/10.1016/j.esp.2007.06.001>

- Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, 32, 150-169. <http://dx.doi.org/10.1017/S0267190512000037>
- Jackson, H., & Stockwell, P. (2011). *An introduction to the nature of functions of language* (2nd ed.). London and New York: Continuum International Publishing Group.
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60-71. <http://dx.doi.org/10.1016/j.jeap.2015.11.003>
- Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics*, 18(3), 391-417. <https://doi.org/10.1075/ijcl.18.3.06paq>
- Wen, Q. F., Ding, Y. R., & Wang, W. Y. (2005). Features of oral style in English compositions of advanced Chinese EFL learners: An exploratory study by contrastive learner corpus analysis. *Foreign Language Teaching and Research*, 35, 268-274.

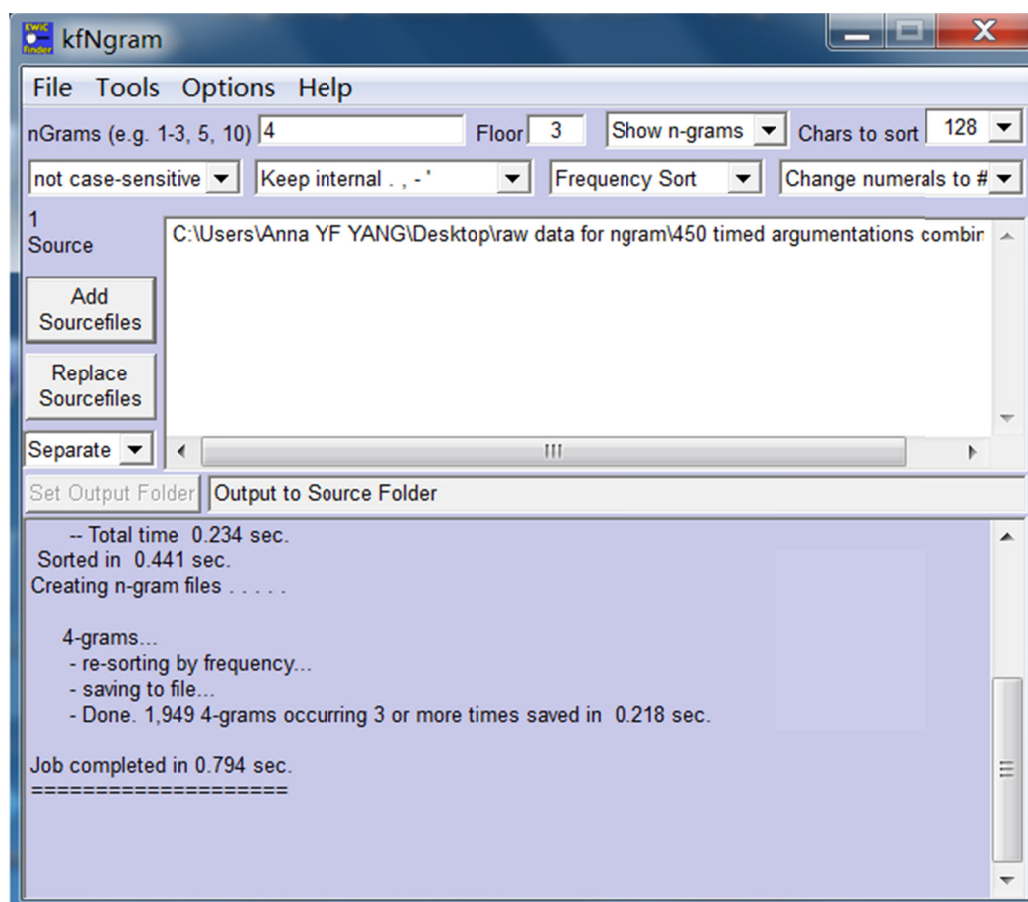
Notes

Note 1. Some studies (e.g., Hyland, 2008a; Breeze, 2013) also found that the use of lexical bundles varies across disciplines and across genres within a certain discipline. Since they are less relevant to the present research, those studies are not reviewed here.

Note 2. The timed argumentative writings by Grade 4 students (60 texts in total) are fewer than those by the student of other grades.

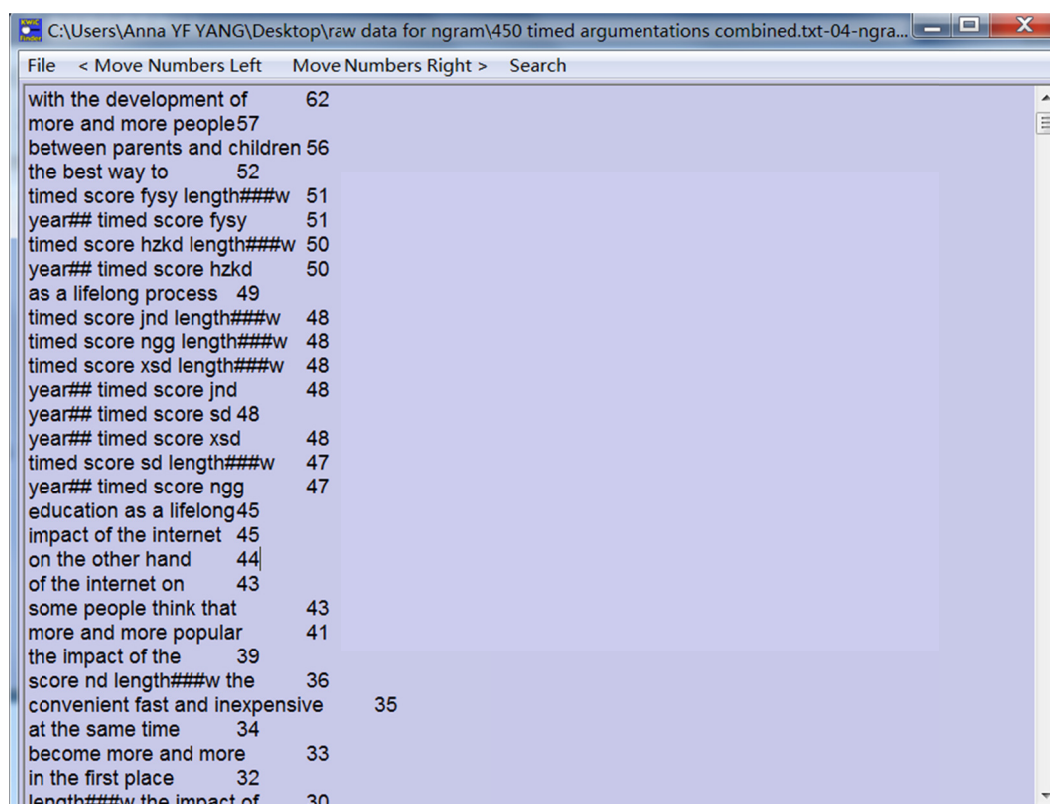
Appendix A

Screen Shot of kfNgram Software



Appendix B

Screen Shot of the Bundles Generated by kfNgram for the Argumentative Writings



with the development of	62
more and more people	57
between parents and children	56
the best way to	52
timed score fysz length###w	51
year## timed score fysz	51
timed score hzkd length###w	50
year## timed score hzkd	50
as a lifelong process	49
timed score jnd length###w	48
timed score ngg length###w	48
timed score xsd length###w	48
year## timed score jnd	48
year## timed score sd	48
year## timed score xsd	48
timed score sd length###w	47
year## timed score ngg	47
education as a lifelong	45
impact of the internet	45
on the other hand	44
of the internet on	43
some people think that	43
more and more popular	41
the impact of the	39
score nd length###w the	36
convenient fast and inexpensive	35
at the same time	34
become more and more	33
in the first place	32
length###w the impact of	30

Appendix C

Sample of Argumentative Writing

<WCOMP><ARG><GRADE1><YEAR03><TIMED><SCORE?><JND><LENGTH330W>

E-mail

With the development of the technology, computers have become part of our lives. Even when communicating with others, no matter far away or near, people prefer to use e-mail. E-mail brings the convenience to people. And also people are tempted by its being fast and inexpensive. For example, you can receive an electric letter immediately after the person in the other side sends it to you, and this procedure costs you little.

So why there are still conventional mails existing while e-mail is considered so perfect? The reason is that some people think that paper letters are more personal, friendly and valuable. I myself don't think it make sense. First, being personal. I think e-mail provides for more security than conventional mails. Any one can open your letter without your knowledge. And you also have risks of losing letters during the transpotion<sp-transportation>. Second, being friendly. Being friendly doesn't mean that one sees other's handwriting. It always takes some time to reach the destination. Everyone will feel more comfort in reading the letter immediately than waiting for the letter with anxious. Third, being valuable. Does being valuable mean that you can put the letters in the deepest box and look it up? Even treating the letters as the gold? I think the letters will valuable as long as the person who send them is important in your life. I think the reasons why some people prefer conventional mails are ridiculous.

We are in the computer world now. We should equip ourselves to adjust to new technology as quickly as possible. So give away those conventional mails. Don't hesitate to change your habits. We must take full advantage of the new technology. People who can absorb new ideas can keep pace with the modern society better. Don't make your friends and your family members wait for your letters any more. Send an electric mail, then you are likely to receive their reply in the next hour. Is it more friendly and valuable?

Appendix D

Sample of Narrative Writing

WCOMP<NAR><GRADE1><YEAR03><TIMED><SCORE?><ND><LENGTH234W>

One Sunday morning last summer, Lan Hua, who was my best friend, and I went to the park near our school. We took along our English books, for we planned to read English in a pleasant place.

About half an hour passed. Then suddenly, a loud noise came into our ears. We had to stop reading, to see what had happened. Then we saw two men fighting. A man was tall and thin, the other was short and fat. Their faces were filled with anger, and they looked very dirty. What was worse, the clothes of the short and fat man was broken. The tall and thin man wasn't better than him. He wore only one shoe. They quarreled, cried and shouted, making more and more noise.

The fighting seemed endless. Lan Hua said to me in an angry voice, "they are ugly! On so beautiful a morning, they makes so much noise." I agreed and hated them very much.

We thought that it wasn't a fine place for us to read English, so we decided to go away. At that very time, we heard a man shouted "Stop!" Then the two men stopped fighting and smiled to each other. We thought it was quite strange and wondered why. Then the two men came to us and said, "We're sorry to trouble you, but we were practicing, for we are actors."

We understood finally.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).