

Lexical Bundles in Selected Children's Fiction: A Corpus-based Analysis

Antermeet Kaur A/P Senthok Singh and Ang Leng Hong

Received: 1 Sep 2019. Accepted: 29 May 2019/Published online: 31 may 2019
© CPLT 2019

ABSTRACT

This study aims to identify four-word lexical bundles in the selected children's fiction. Previous studies on lexical bundles have investigated the existence of lexical bundles in a wide range of genres. However, little has been done on children's fiction with regard to the use of lexical bundles in this genre. Using Biber, Conrad and Cortes's (2004) framework, this study therefore analyses the structural and functional properties of lexical bundles in a corpus of children's fiction. A 1.7 million-word corpus was built comprising 30 well-read children's books. The data was generated and analysed using a corpus analysis tool, WordSmith Tools Version 6.0. The results revealed the presence of lexical bundles in the selected children's fiction. The structural analysis results show that prepositional and verb phrases dominate the children's fiction. With regard to the functional classification of lexical bundles, referential lexical bundles occur the most, followed by action-related expressions and stance bundles. The results are indicative of the presence of lexical bundles in children's fiction which has not received much research attention in phraseology studies. This study has several pedagogical implications which stress on the importance of employing lexical bundles in fiction, textbooks and classroom activities in order to benefit children in their language learning and acquisition. Lists of frequent lexical bundles can be incorporated into English language lessons as a way to expose learners to the phraseological patterns of language.

Keywords: Children's fiction. Lexical bundles. Functional classification. Structural classification

✉ Ang Leng Hong
School of Humanities
Universiti Sains Malaysia, Penang MALAYSIA
E-mail: lenghong@usm.my

1 INTRODUCTION

Lexical bundles are sets of continuous word sequences, in other words, extended collocations which appear more often than expected in written or spoken language (Biber, Johansson, Leech, Conrad & Finegan, 1999; Hyland, 2008). They are structures of grammar which serve important roles in language use, i.e. ensuring fluent and natural use of language (Wray 2002) in various contexts, particularly in written discourse. This makes lexical bundles a great concern in language and linguistic research. Previous studies on lexical bundles, to a large extent, have focused on academic genres. These studies (e.g., Adel & Erman, 2012; Ang & Tan, 2018; Biber & Barbieri, 2007; Biber, Conrad & Cortes, 2004; Biber et al., 1999; Byrd & Cortes, 2002; Conrad & Biber 2004; Coxhead, 2010; Hyland, 2008; Pan, Reppen & Biber, 2016; Perez-Llantada, 2014; Wei & Lei, 2011) have investigated lexical bundles in various academic genres such as university textbooks, research articles, and doctoral dissertations. Although numerous studies on lexical bundles have been conducted on academic genres, Cortes (2004) found that there are still unanswered questions about the use of lexical bundles across different registers and genres. This viewpoint serves as a motivation for the study as very few studies have looked at lexical bundles in the genre of children's fiction. Little is known on the types of lexical bundles present in children's fiction. There is a need to identify lexical bundles and the roles they play in children's fiction as these books are read by generations of children worldwide. It is worth analysing the functional and structural patterns of lexical bundles in children's fiction as the analysis would provide insights into the phraseological tendency of language in this particular genre.

Children's fiction is considered as a useful resource for first language and second language learners to develop and expand their vocabulary (Cheetham, 2015). Nelson (2016) discovered that vocabulary growth is related to reading. He added that when children are exposed to lexis through reading they build their vocabulary. Being aware of this fact, it is crucial to know the kind of input that children take in when they read children's fiction (Nelson, 2016). As such, it is necessary to know the types of word chunks such as lexical bundles that are commonly used in children's fiction to see if they are similar to or different from those found in other genres such as academic writing.

As continuous word sequences, four-word lexical bundles are found to be the focus of many researchers as four-word length has been seen more useful and manageable for learners and researchers (Biber et al., 2004; Cortes, 2004). Following the literature, the study therefore intends to look at four-word lexical bundles in selected children's fiction available worldwide.

2 LITERATURE REVIEW

Biber et al.'s (1999) ground-breaking study on lexical bundles examined the presence of lexical bundles in both spoken and written registers. They reckoned that these word combinations are "too systematic to be disregarded as accidental" (Biber et al., 1999, p. 290). According to Biber et al. (1999), a word combination is regarded as lexical bundle if it appears frequently in a text, for instance over 10 times in every one

million words, with the word lengths of between three to five words. Hyland (2008) added that lexical bundles are “extended collocations which appear more frequently than expected by chance, helping to shape meanings in specific contexts and contributing to our sense of coherence in a text” (p. 4).

Numerous studies have been conducted extensively on lexical bundles in terms of their structural and functional properties. Studies by Biber and his associates (Biber et al., 1999; Biber & Barbieri, 2007; Biber et al., 2004; Conrad & Biber, 2004) revealed that diverse patterns are found in conversation and academic prose. For instance, in Biber et al.’s (2004) study, the researchers discovered that 90 percent of lexical bundles in conversation are verb phrases while 70 percent of bundles in academic prose are noun phrases. This indicates that the structural patterns of lexical bundles are distinguishable across difference registers. It is therefore worth examining lexical bundles in other text types such as children’s fiction as a way to understand the nature of the phraseological sequences in this text type. Besides, a number of researchers attempted comparative studies, for instance Cortes (2002). Cortes compared lexical bundles present in freshman writing and academic prose. She found that the students’ compositions have similar structures of lexical bundles compared with those of academic prose. However, the functions of the lexical bundles in both genres are found to differ significantly. Similarly, Cortes (2004) examined the differences between native students’ writings and writings in academic journals. She found that students use lexical bundles that are different from those used by expert writers in academic journals. To understand the use of lexical bundles by native and non-native speakers, Kashiha and Chan (2015) investigated the use of lexical bundles in classroom discussions. Their analysis revealed that native speakers use more lexical bundles than non-native speakers while conducting discussions in the classroom. The finding was similar to past studies conducted by Adel and Ermen (2012), Chen and Baker (2010) and Karabacak and Qin (2013) among native and non-native speakers. In an attempt to uncover the disciplinary influences, Hyland (2008) examined the use of lexical bundles across four disciplines. He found that students from different fields have different preferences over the use of lexical bundles. He noted that students taking Electrical Engineering use more lexical bundles than those studying Biology.

As mentioned earlier, little is known about the types of lexical bundles which are commonly used in children’s fiction. This study therefore addresses the gap by identifying lexical bundles that are commonly found in children’s fiction. Specifically, this study intends to:

- i) identify the most frequent four-word lexical bundles in the selected children’s fiction.
- ii) classify the structural patterns of the four-word lexical bundles in the selected children’s fiction.
- iii) classify the functions of the four-word lexical bundles in the selected children’s fiction.

3 METHODOLOGY

3.1 Corpus Development

For the purpose of this study, corpus-based methods were employed in compiling the data for analysis purposes. The corpus development involved four steps. First, a total of 30 children's fiction books in the pdf format were downloaded from online sources and stored in a named folder. Each pdf file consisted of one book. Second, each pdf file was converted into plain text file format (.txt) as the corpus analysis tool used in this study only recognises data in plain text format. The plain texts were then stored in a separate named folder. Third, extra information or metadata are discarded from the text file since the researchers only intended to investigate the language contents of the texts. Finally, the texts in the folder were named accordingly, serving as an index for each book in the corpus. The corpus of the study comprises 30 children's fiction books and it contains 1.7 million word tokens.

3.2 Criteria for Data Generation

The criteria suggested by Biber et al. (2004) were referred to in determining if a word combination is qualified as a lexical bundle. The criteria are as follows:

- i) Cut-off frequency- A word combination has to occur at least 10 times to be considered as lexical bundle. In this study the minimum cut-off frequency was set at 20 times to identify highly frequent lexical bundles in children's fiction.
- ii) Incidence of combinations- In order to qualify as lexical bundle, a word combination must occur in at least 5 texts written by different authors. This is to avoid author bias.
- iii) Size of word combinations - Only 4-word bundles were taken into account as they are more useful and manageable as compared to 3-word bundles and 5-word bundles.

3.3 Analysis of Data

This study adapted Biber et al.'s (2004) framework to analyse the structural and functional patterns of lexical bundles found in 30 well-read children's fiction books written by 11 authors from Britain and the United States of America. The selection of Biber et al.'s (2004) framework was due to its comprehensiveness in classifying the structures and functions of lexical bundles in written texts. The corpus analysis tool, *WordSmith Tools Version 6.0* was used to analyse the data and generate the relevant frequency information. By applying the three criteria mentioned above, *WordSmith Tools* was used to generate a list of word combinations as the possible lexical bundles. As the word combinations were automatically generated, manual checking was carried out to identify the meaningful word combinations. The meaningless word combinations such as the ones that cross the syntactic boundaries were discarded. *WordSmith Tools* also generates the concordance lines for the purposes of identifying, analysing and classifying the structural and functional patterns of lexical bundles.

4 RESULTS AND DISCUSSION

74 types of lexical bundles were identified, analysed and classified according to their structural and functional properties (refer to Appendix A for the list of lexical

bundles). Table 1 below exhibits the top 10 lexical bundles that appear in the selected children’s fiction. It can be seen that most of the lexical bundles occur in high frequencies. The lexical bundle, *the rest of the* occurs more than 100 times, while other lexical bundles occur 89 to 156 times. These lexical bundles appear in a minimum of 20 books which were written by 8 or more authors. This shows that the top 10 bundles are free from author bias and they are prevalent in children’s fiction and thus should be given more attention in language classrooms.

Table 1. List of top ten lexical bundles

Number	Frequency	Lexical Bundles
1.	190	the rest of the
2.	156	in the middle of
3.	151	the end of the
4.	149	the top of the
5.	135	the middle of the
6.	130	the edge of the
7.	110	at the end of
8.	100	in front of the
9.	93	out of the window
10.	89	the back of the

4.1 Structural Properties of Lexical Bundles

In this section the lexical bundles were analysed and classified according to their structural properties. According to Biber et al. (1999) even though lexical bundles are not deemed as complete structural units, it is possible to group them in line with their grammatical correlates. In this study, the lexical bundles were classified into 4 main structural categories. They are Noun Phrase (NP) (e.g. *the edge of the*), Verb Phrase (VP) (e.g. *came out of the*), Prepositional Phrase (PP) (e.g. *in the middle of*) and Others. (refer to Appendix B for list of the structural categories and sub-categories of lexical bundles). Figure 1 below illustrates the distribution of the different structural types of lexical bundles found in this study.

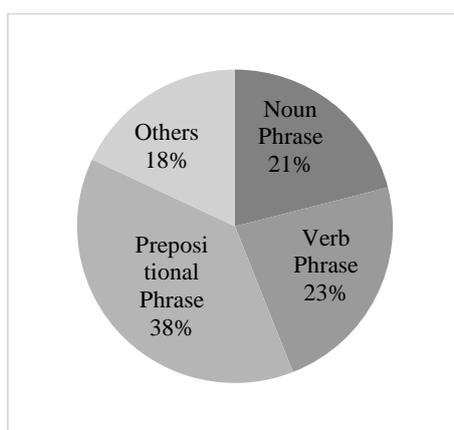


Fig 1. Distribution of structural types

As shown in Figure 1, the most prevalent bundles are in the form of prepositional phrases. These prepositional phrases were further classified into two sub-categories, namely prepositional phrases with *of* (PP + NP fragment containing

of) and other prepositional phrases. An example of a prepositional phrase with *of* is *at the end of*. This type of lexical bundle usually begins with a preposition and ends with *of* following the noun. Verb phrase expressions account for the second highest number of lexical bundles in this study. An example of a verb phrase is *go back to the*. The verb phrases were further divided into six sub-categories, namely verb phrase expressions, anticipatory *it* + verb phrase, yes-no question fragments, *that*-clause fragments, *to*- clause fragments and modal/semi modal expression. An example for the sub-category, *to*- clause fragment is *to be able to*. This is a simple *to*-clause bundle that begins with *to* and indicates possibility/ability (Biber et al., 1999). Another sub-category is *that*- clause fragment. An example of this structure is *that it was a*. This is a simple *that*- clause with *it* as its subject and the copula *is* as the verb. The third highest number of lexical bundles in the structural classification is noun phrase expressions. These noun phrase expressions were further classified into two sub-categories, namely noun phrase with *of* phrase fragment and other noun phrase expressions. An example of a noun phrase is *the back of the*. Similar to academic prose, most of the lexical bundles in this category consist of an “incomplete noun phrase containing an *of* phrase, usually identifying a physical location: the head noun specifies some position (the *back, middle, top, bottom, other side, etc*) with respect to the complement *of*” (Biber et al., 1999, p. 1012). The category, “Others” comprises three sub-categories, namely personal pronoun + lexical verb phrase, *wh*-question fragments and adverbial clause fragment. An example of personal pronoun + lexical verb phrase structure is *I don't want to*. This type of lexical bundle is usually a clause fragment which consists of a subject pronoun followed by a verb phrase. The verb phrase usually follows a complement clause. For the *wh*-question fragment structure, a commonly found lexical bundle associated with this structure is *What are you doing*. This type of lexical bundle usually begins “with a *wh*-question word (e.g., *what, where, how*) and a common verb such as *do* and *say* occurs with it” (Biber et al., 1999, p. 1008).

With regard to the structural characteristics of lexical bundles, the findings show that the most occurring lexical bundles in children's fiction are prepositional phrases followed by verb phrases and noun phrases. Nevertheless, different findings were reported in past studies on lexical bundles. For instance, Biber et al. (2004) discovered that about 90 percent of lexical bundles in conversation are verb phrases and 70 percent of lexical bundles in academic prose are noun phrases. Also, they found that classroom teaching uses a large number of noun phrases and prepositional phrases. This indicates that different structural lexical bundles are used in different genres and registers.

In analysing lexical bundles in Malaysian University English Test (MUET) reading texts, Ong and Yuen (2017) found that lexical bundles in MUET reading texts are mostly in the form of prepositional phrases. This previous research finding is similar to the findings of the study. Despite the fact that the selected children's fiction is of different genre, prepositional phrases are abundant in the selected children's fiction. This shows that prepositional phrases as an indicator of circumstances are indeed needed for readers to understand the context of their reading.

4.2 Functional Properties of Lexical Bundles

In ensuring the cohesiveness and readability of the text, lexical bundles are also continuous word combinations that aid readers' comprehension of the texts. Lexical bundles are always used in texts to introduce topics, compare and contrast ideas and draw conclusions. It is therefore necessary to identify the functions lexical bundles serve in the children's fiction in order to provide appropriate functional description of lexical bundles in the relevant genre. Figure 2 below displays the functional distribution of the lexical bundle types identified in the study. As can be seen, referential bundles are the most frequent expressions. The category, "others" follows with 16%. Stance expressions, discourse organisers and special conversational expressions are less common in children's fiction. Referential bundles are expressions used to identify important entities, or give particular attributes; Stance bundles are useful in conveying epistemic meaning and writer's attitude towards a particular proposition; Discourse organisers mainly function as topic introduction and elaboration (Biber et al. 2004).

It should be noted that the category "Others" is a new category formed to include lexical bundles that do not belong to the existing categories proposed by Biber et al. (2004). Lexical bundles in this category are expressions describing simple actions which take place in various events in the children's fiction. A detailed list of the functional categories and their sub-categories of lexical bundles is available in Appendix C.

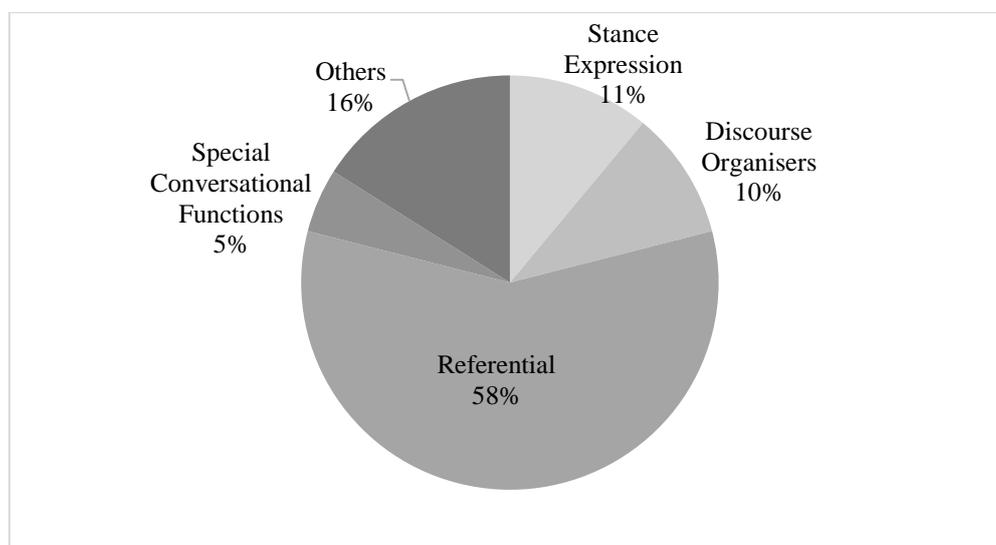


Fig 2. Functional classifications of lexical bundles

Referential bundles appear abundantly in the corpus as they serve a multitude of functions, including specifying attributes and emphasising purposes. Most of the writers in the selected children's fiction use lexical bundles to refer to characters, situations, places, time and events. The use of referential bundles is also dominant in past studies. For instance, Conrad and Biber (2004) discovered that referential bundles are commonly used in academic prose. Kashiha and Chan's (2013) study revealed that lexical bundles serving as referential expressions are frequently used in hard sciences.

As mentioned earlier, lexical bundles used to describe actions or events that take place in the children's fiction are classified as "Others" category in this study (e.g. *what are you doing*). Notably, lexical bundles portraying actions and events are found commonly in the selected children's fiction. This indicates that in children's fiction, it is necessary to understand bundles serving such discourse functions in order to understand the texts.

It is worth noting that the special conversational bundles are rarely used in the selected children's fiction. These bundles are usually used in conversations. In the selected children's fiction, there are not much conversational contents that employ the use of conversational lexical bundles. Furthermore, children's fiction is intended for a younger age group and may not require these bundles.

In line with Biber et al. (2004), stance expressions are also found in the selected children's fiction. For example, the bundles *are you going to* and *he was going to* are used on characters in the story to express their plans. This finding is similar to Biber et al.'s (2004) finding in which they found that stance expressions are dominant in classroom teaching and conversation. Kashiha and Chan (2013) also found that stance expressions are frequently used in spoken and classroom teaching in both hard and soft sciences. Stance bundles are used by writers in different genres to express beliefs and opinions (Biber et al. 2004; Simpson-Vlach & Ellis 2010).

Discourse organisers are also present in the children's fiction, though in small percentage. Lexical bundles classified into this category are used to elaborate or clarify ideas in the texts. Some examples are *on the other hand*, *as soon as he* and *as if they were*. Biber et al.'s (2004) study revealed that discourse organisers are prevalent in classroom teaching but not in conversations, textbooks and academic prose.

5 CONCLUSION

It is evident that the use of lexical bundles is not limited to the academic sphere as different types of lexical bundles are found in the genre of children's fiction, too. This study adapted Biber et al.'s (2004) structural and functional classifications in analysing and classifying lexical bundles found in the selected children's fiction. The findings indicate that writers of children's fiction favour the use of prepositional phrases and verb phrases as compared to noun phrases and other grammatical structures. Past studies on academic writing also revealed similar results where the prepositional phrases and verb phrases occur abundantly in academic texts. With regard to the functional classifications of lexical bundles, the referential bundles are commonly used in the selected children's fiction. These bundles are used to make references to characters, places, time and events. The special conversational bundles show a marginal presence as they are rarely employed by the relevant children's fiction writers. To sum up, some of the findings of this study are consistent with findings of the previous studies on lexical bundles, though they are of different genres.

5.1 Pedagogical Implications

The outcome of the study can raise awareness on the importance of exposing the knowledge of word combinations such as lexical bundle to learners. It also stresses on how an array of children's books can play a role in children's language acquisition directly and indirectly. In relation to language acquisition, Byrd and Coxhead (2010) proposed that teachers incorporate the teaching of word clusters such as lexical bundles and the functions they serve in texts such as children's books in language classroom. Nation (2001) and Gouverneur (2008) suggested that teachers should be given the authority to select and decide the teaching of lexical bundles according to the learners' capability and levels. Also, in classroom teaching, teachers are encouraged to conduct interesting and beneficial activities such as 'Learn a bundle a day' in raising the learners' awareness on the importance of word clusters such as lexical bundles. Frequent repetition of lexical bundles is vital to ensure learners, especially the children understand and remember the word combinations. This will assist the learners to retrieve and use the word combinations correctly when necessary.

ACKNOWLEDGEMENTS

This work was supported by Universiti Sains Malaysia Short Term Grant (304/PHUMANITI/6315044).

REFERENCES

- Adel, A., & Erman (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31, 81-92.
- Ang, L. H., & Tan, K. H. (2018). Specificity in English for Academic Purposes (EAP): A Corpus Analysis of Lexical Bundles in Academic Writing. *3L: The Southeast Asian Journal of English Language Studies*, 24(2), 82-94.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263-286.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Lexical expressions in speech and writing. In D. Biber & G. Leech (Eds.), *Longman grammar of spoken and written English* (pp. 988 -1036). Harlow, Essex: Longman.
- Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, 5(5), 31-64.
- Cheetham, D. (2015). Extensive reading of children's literature in first, second, and foreign language vocabulary acquisition. *Children's Literature in English Language Education*, 3(2), 1-23.
- Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 12(2), 30-49.

- Conrad, S., & Biber, D. (2004). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica*, 20, 56-71.
- Cortes, V. (2002). Lexical bundles in Freshman composition. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 131-145). Amsterdam: John Benjamins Publishing Company.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397-423.
- Gouverneur, C. (2008). The phraseological patterns of high-frequency verbs in advanced English for general purposes: A corpus-driven approach to EFL textbook analysis. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp.223-243). Amsterdam: John Benjamins.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21.
- Karabacak, E., & Qin, J. (2013). Comparison of lexical bundles used by Turkish, Chinese, and American university students. *Procedia-Social and Behavioral Sciences*, 70, 622-628.
- Kashiha, H. & Chan, S. H. (2015). A little bit about: Differences in native and non-native speakers' use of formulaic language. *Australian Journal of Linguistics*, 35(4), 297-310.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nelson, K. M. (2016). *Lexical trends in young adult literature: A corpus-based approach* (Unpublished master's thesis). Brigham Young University, the United States.
- Ong, C. S. B., & Yuen, C. K. (2017). Comparing structural and functional lexical bundles in MUET reading test. *Pertanika Journal of Social Sciences and Humanities*, 25(1), 133-148.
- 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.
- Perez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes*, 14, 84-94.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An Academic Formulas List (AFL). *Applied Linguistics*, 31, 487-512.
- Scott, M. (2012). *Wordsmith Tools* (Version 6.0) [computer software]. Available from <https://www.lexically.net/wordsmith/downloads/>
- Wei, Y., & Lei, L. (2011). Lexical Bundles in the Academic Writing of Advanced Chinese EFL Learners. *RELC Journal*, 42(2), 155-166.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

APPENDIX A

List of lexical bundles present in the 30 selected children's fiction

Number	Frequency	Lexical Bundles	Number of Books	Number of Authors
1.	190	the rest of the	28	11
2.	156	in the middle of	30	11
3.	151	the end of the	28	11
4.	149	the top of the	28	10
5.	135	the middle of the	30	11
6.	130	the edge of the	27	11
7.	110	at the end of	24	10
8.	100	in front of the	20	10
9.	93	out of the window	20	8
10.	89	the back of the	24	10
11.	88	the bottom of the	24	9
12.	88	the other side of	25	9
13.	86	for the first time	22	9
14.	84	at the top of	23	8
15.	81	he was going to	23	7
16.	79	in front of him	19	8
17.	78	out of the way	23	7
18.	72	for a long time	24	9
19.	71	at the bottom of	26	10
20.	70	in a low voice	24	6
21.	67	on the other side	23	7
22.	66	out of the room	20	9
23.	64	at the same time	25	10
24.	63	what are you doing	21	6
25.	61	what do you mean	22	8
26.	60	the foot of the	19	8
27.	58	the door of the	21	9
28.	56	to go to the	23	9
29.	55	are you going to	20	9
30.	54	up and down the	15	6
31.	53	the side of the	19	9
32.	52	at the foot of	16	6
33.	51	as soon as he	22	7
34.	51	it would have been	18	8
35.	50	at the back of	18	5
36.	50	went back to the	19	6
37.	48	as if he were	19	8
38.	48	what do you think	22	6
39.	47	go back to the	20	8
40.	45	on the edge of	19	7
41.	45	to the top of	15	6

42.	44	as if it were	16	8
43.	44	came out of the	21	8
44.	44	in front of them	18	7
45.	43	as soon as they	19	6
46.	43	for a few minutes	21	8
47.	43	the three of them	14	5
48.	43	what was going on	15	7
49.	42	was going to be	22	9
50.	41	as if they were	15	6
51.	41	I don't want to	16	7
52.	40	to be able to	20	8
53.	39	in and out of	15	7
54.	39	it must have been	20	6
55.	38	go down to the	18	7
56.	38	he said in a	17	6
57.	38	it would be a	17	6
58.	38	on the other hand	17	6
59.	38	the back of his	12	6
60.	38	to go back to	19	6
61.	37	into the air and	11	6
62.	37	looked at one another	18	5
63.	36	that it was a	16	6
64.	36	the hole in the	11	5
65.	36	to the end of	17	9
66.	35	in a few minutes	20	8
67.	35	sat down on the	23	10
68.	35	the roof of the	13	6
69.	34	at the same moment	15	7
70.	34	get out of the	17	6
71.	34	I want you to	13	9
72.	34	there was no sign	17	7
73.	34	to the edge of	16	6
74.	34	went down to the	17	6

APPENDIX B

Structural classifications of lexical bundles

Category	Sub-category	Lexical bundle
Noun Phrases	Noun phrases with <i>of</i> phrase fragments	the rest of the the end of the the top of the the middle of the the edge of the the back of the the bottom of the the other side of the foot of the the door of the the side of the the back of his the roof of the the three of them
	Other Noun Phrases	the hole in the there was no sign
Prepositional Phrases	Prepositions + noun phrase fragments containing <i>of</i>	in the middle of at the end of in front of the out of the window at the top of in front of him out of the way at the bottom of out of the room at the foot of at the back of on the edge of to the top of in front of them in and out of to the end of to the edge of
	Other prepositional phrases (fragments)	for the first time for a long time in a low voice on the other side at the same time

		up and down the on the other hand into the air and in a few minutes at the same moment for a few minutes
Verb Phrases	Verb phrase expressions	looked at one another went back to the go back to the came out of the go down to the sat down on the get out of the went down to the
	Anticipatory <i>it</i> + verb phrases	it would be a it would have been it must have been
	Yes-no question fragment	are you going to
	<i>that</i> -clause fragment	that it was a
	<i>-to</i> clause fragments	to go to the to go back to to be able to
	modal/ semi-modal expression	was going to be
	Others	Personal Pronoun + Lexical Verb Phrases
Wh-question fragments		what are you doing what do you mean what do you think what was going on

	Adverbial Clause Fragments	as soon as he as soon as they as if he were as if it were as if they were
--	-------------------------------	---

APPENDIX C

Functional classifications of lexical bundles

1. STANCE EXPRESSION	Lexical Bundle
Attitudinal/Modality Stance	
1A Desire Personal	I don't want to
1B Obligation/Directive Personal	I want you to
1C Intention/Prediction Personal	are you going to
Impersonal	was going to be he was going to
1D Ability Personal	to be able to to go back to to go to the
2. DISCOURSE ORGANISER	
2A Topic Elaboration/ Clarification	on the other hand as soon as he as soon as they as if he were as if they were as if it were there was no sign
3. REFERENTIAL EXPRESSION	
3A Identification/ Focus	that it was a it must have been it would be a it would have been the three of them

3B Time Reference	at the same time for the first time for a long time in a few minutes at the same moment for a few minutes
3C Multi-Functional Reference	the end of the the top of the the middle of the the edge of the the back of the the bottom of the the other side of the foot of the the door of the the side of the the back of his the roof of the the hole in the in the middle of at the end of in front of the out of the window at the top of at the bottom of out of the room at the foot of at the back of on the edge of to the top of to the end of to the edge of in front of him out of the way in front of them on the other side up and down the the rest of the
4 SPECIAL CONVERSATIONAL FUNCTION	

4A Simple Inquiry	What are you doing What do you mean What do you think What was going on
5 OTHERS ACTIONS/HAPPENINGS	
5A Describing simple actions	looked at one another went back to the go back to the came out of the go down to the
	sat down on the get out of the in and out of into the air and in a low voice went down to the he said in a