AN ANALYSIS OF FOUR-WORD LEXICAL BUNDLES: APPLICATION TO ENHANCE LISTENING SKILLS FOR TOEIC PREPARATION COURSE

Wichanon Phongjit
Kamphaeng Phet Rajabhat University Maesot, Thailand
wichanontu@gmail.com

Abstract

The study was conducted to provide frequent four-word lexical bundles to enhance listening skills for Thai undergraduate students who attended the TOEIC preparation course at Kamphaeng Phet Rajabhat university Maesot. The listening scripts from TOEIC preparation commercial books were investigated and were compiled as a corpus. The corpus approximately contained 100,000 words, and it was investigated by the Antconc Application to classify four-word frequency bundles. Then, the study aims to identify the use of the bundles based on the structural and functional taxonomy of lexical bundles by Biber et al. (2004). From the corpus-based investigation, it was found that the lexical bundle structures and functional structures are varied. Lexical bundles that incorporate verb phrase fragments appeared to be the high frequent category. The referential bundles and stance bundle were frequently found in the corpus. The findings are also useful for English teachers to prepare these bundles for their students in teaching listening skills. This paper was intended to discuss the beginning stage of the frequent lexical bundles to be used in material design or course development which might enhance language proficiency in listening skills for TOEIC preparation.

Keywords: Corpus, Lexical bundles, Listening skills

1. Introduction

Learning English as a foreign language can be classified in many areas that students may realize the difficulty in a stage of development. One main area is focused on students learning experiences. It is acknowledged that the students require a certain stage in learning, that is, receptive skills. The receptive skills can be divided into two main skills, which are reading and listening. Those skills share the same learning theoretical concepts for learning which are micro skills and macro skills. One gap that cannot be denied that listening is more interactive in the interactions. Tsai (2001) mentions that in a real situation, students need to understand the conversation to apply language in the classroom to function in English. Morley (2001) observes that "listening is still regarded as the least important skill". Underwood (1989:1) states that listening is the attention process that the listener tries to get from the message we hear. Listening can be viewed as a part of verbal communication that is typically used in daily lives. Therefore, it is necessary to find out more effective teaching methodologies to help students comprehend their listening. Besides, in order to study the effectiveness, testing listening performance should be an aspect that should be considered. Many English standardized tests evaluate the listening skills of the test's takers, which are TOEIC, TOEFL, and IELTS, TOEIC. The contents of these tests are mainly focused on the field of business (Lougheed, 2012).

In this study, one aspect that could enhance the listening skills could be lexical bundles. Lexical bundles are a key component of spoken and written language that is an effective use for processing language (Biber & Barbieri, 2007; Hyland, 2008). The lexical bundles can be found in many parts of language use. This study also focused on the use of language in TOEIC. This is because TOEIC is considered for testing general communication. It could be said that the lexical bundles are frequently found in the listening. It may occur in TOEIC listening parts. Therefore, this study investigated the lexical

bundles used in TOEIC listening scripts to see their frequent forms and functions, enhancing their listening skills for the TOEIC test.

2. Review of literature

2.1 Lexical bundles

Lexical bundles are frequent words in extract form and repeatedly used in a certain register (Biber et al., 1999). They can extract by the computation program from the corpus, which collects all data in the set representing enough amount of the language in each register (McEnery & Wilson, 2001). The lexical bundles can be either idiomatic expression or non-idiomatic expression. The lexical bundles are viewed as "discourse building blocks" (Biber, 2009) which can essentially serve communicative purposes for each register (Conrad & Biber, 2005, p.63). To identify lexical bundles, Biber et al. (2004) provide the theoretical framework that lexical bundles need to be analysed in parts of structural and functional categories.

2.2 Theoretical Framework of Lexical Bundles

In terms of structural categories, Biber *et al*. (2004), according to this classification, there are three main structural types which include 1) lexical bundles that incorporate verb phrase fragments like *that's one of the, is based on the*, 2) lexical bundles that incorporate dependent clause fragments like *that this is a, to come up with*, 3) lexical bundles that incorporate noun phrase and prepositional phrase fragments like *at the end of, at the same time*. Each main structural type entails several sub-types, as illustrated in Table 1.

Table 1. Three Types of Lexical Bundles Based on Structural Classification (Biber et al., 2004)

Structural types	Sub-types	Sample bundles
1. Lexical bundles that incorporate verb phrase fragments	1.a 1st/2nd person pronoun + VP fragment	*I'm not going to*
	1.b 3rd person pronoun + VP	*and this is a*
	1.c discourse marker + VP fragment	*I mean I don't*
	1.d Verb phrase (with non-passive	*have a lot of*
	1.e Verb phrase (with passive verb)	*is based on*
	1.f yes-no question fragments	*are you going to*
	1.g WH- question fragments	*what do you think*
2. Lexical bundles that incorporate dependent clause fragments	2.a 1st/2nd person pronoun + dependent	*I want you to*
	2.b WH-clause fragments	*when we get to*
	2.c If-clause fragments	*if we look at*
	2.d to-clause fragment	*to be able to*
	2.e That-clause fragment	*that this is a*
3. Lexical bundles that incorporate noun phrase and prepositional phrase fragments	3.a Noun phrase with of-phrase fragment	*one of the things*
	3.b Noun phrase with other post-modifier fragment	*the way in which*
	3.c Other noun phrase expressions	*a little bit more*
	3.d Prepositional phrase expressions	*at the end of*
	3.e Comparative expressions	*as well as the*

Table 1 consists of three main columns on the top: structural types, subtypes, and sample bundles. In the first column, Structural types are the main type of the sub-types in the second column. The last column, which names sample bundles, demonstrates the sample of corresponding sub-types.

Another analysis of lexical bundles is functional analysis or discourse functions. Biber *et al*. (2004) define three types which are 1) stance bundles express attitude or assessment, 2) discourse organizers reflect the relationships between different parts of texts, 3) referential expressions refer to physical or abstract entities or other textual parts. These are shown in Table 2.

Table 2. Discourse functions of lexical bundles (based on Biber *et al*., 2004, pp.384-388)

Table 2 consists of three main columns on the top: categories, sub-categories, and sample bundles. In the first column, categories are the main type of the Sample bundles in the second column. The last column, which names sample bundles, demonstrates the sample of corresponding sub-types.

Categories	Sub-categories	Sample bundles
I. Stance bundles	**A. Epistemic stance**	*the fact that the, and I think that*
	B. Attitudinal/ modality stance	
	B1) Desire	*what do you want*
	B2) Obligation/ directive	*it is important to*
	B3) Intention/ Prediction	*it's going to be*
	B4) Ability	*it is possible to*
II. Discourse organizers	**A. Topic introduction**	*in this chapter we*
	B. Topic elaboration/ clarification	*on the other hand*
III. Referential bundles	**A. Identification/ focus**	*is one of the*
	B. Imprecision	*or something like that*
	C. Specification of attributes	
	C1) Quantity specification	*a lot of people*
	C2) Tangible framing	*in the form of*
	C3) Intangible framing	*on the basis of*
	D. Time/ Place/ Text reference	
	D1) Place reference	*in the United States*
	D2) Time reference	*at the same time*
	D3) Text-deixis	*as shown in figure*
	D4) Multi-functional reference	*in the middle of*

2.2 Listening comprehension and Lexical Bundles

Bottom-up processing is used when listeners build the meaning from the sound that he/she hears. Sound is converted to words and from words to sentience, creating meaning (Vandergrift, 1997, p. 387). Richards (2008) indicates that this process plays important roles in understanding utterances from speakers. One process of bottom-up processing mentally breakdown the language into its components which refers to

chunking. He also points out that the core meaning of the message can be understood by using chunking. Besides, Yeldham and Gruba (2014, p. 35) argue that bottom-up processing relates to developing students' micro listening skills solely. It is noted that students come across a problem when they listen to long and complex sentences, and they cannot use this information processing to help their comprehension. On the other hand, Batova (2013, p. 187) argues that this process can also help foreign language learners listen to unfamiliar language patterns.

Top-down processing refers to the listener prior knowledge of the topic, their general knowledge of the world, and how the message generally interacts with their linguistic knowledge to interpret the message. This could be the knowledge about the topic, and learners try to use the knowledge that they already have on listening topics to crate expectation on the topic what they expect to hear and interpret their understandings (Helgesen, 2003, p. 26). According to Rahimi (2012) and Buck (2001), when the listeners try to use prior knowledge on the topics to interpret the listening text, they may sometimes make an incorrect interpretation. This relates to the concept 'schemata', A schema represents our knowledge about concepts; that is, "those underlying objects, situations, events, sequences of events, actions and sequences of actions" (Fang, 2008, p. 23), and, is defined as a mental representation of a typical instance (Cook, 1997, p. 86). It is found that the learners tend to use their schema concerning the topic to construct and guide an incoming language input (bottom–up processing) for making interpretation.

Bottom-up knowledge can refer to the micro marker such as well, ok and now (Chaudron & Richards, 1986, p. 111). Top-down knowledge is a macro marker such as *I' am going to, in the end of*. It is possible to say that a macro marker relates to lexical bundles from its form and function. Chaudron and Richards (1986) studied how these two types of marker contribute to students' listening comprehension. They believed that the micro and makers could contribute to one another in order to help listening comprehension. It was found that the macro markers, defined as lexical bundles found it more useful in listening comprehension than the micro marker.

2.3 Previous Studies

Pervious has been studied on lexical bundles in spoken and written. Since the study is related to spoken lexical bundles, five empirical studies will be discussed. Biber et al. (2004) explored frequency sequences of lexical bundles in four registers, including conversation, classroom teaching, textbook and academic pose. The findings revealed that classroom teaching used more stance organizers than discourse organizers to communicate in the classroom. On the contrary, the more tokens found on classroom teaching using more referential bundles than academic pose. This study is valid when Nesi and Basturkmen (2006) studied lexical bundles in teacher talk. They investigated the use of lexical bundles from two corpora from two monologic university lecturers. The results were in the same line as Biber et al. (2004). Moreover, Kwon and Lee (2014) studied the use of lexical bundles in Korean EFL teacher talk from native English teacher and non-native English teacher. They recorded the teachers' talk and complied with the corpora. The findings found that the functional analysis of the lexical bundles showed that the non-native English teachers relied heavily on a narrow range of functional categories. Besides, Kashiha and Heng (2014) studied the functional categorization of lexical bundles used in academic lectures of two disciplines: politics and chemistry. The findings found that there was some significant variation found in discourse functions of the lexical bundles. Another study by Cortes and Neely (2009) investigated how the use of lexical

bundles affects the listening comprehension of students in the EFL classroom. They focused on the topic-introducing/discourse organizing bundles *if you look at, a little bit about, a little bit of, I want you to, and I would like you* occur in the spoken production of instructors and students in the academic lectures in order to find the pedagogical implication using lexical bundles in listening comprehension lesson. The findings suggested that the use of corpus-based activity in the lesson can prepare students to be familiar with the bundles that they may hear in listening texts.

The previous studies signify the language use in the real spoken language that they are fluently uttered the lexical bundles. So, the students should be trained to know and prepare in advance to be helpful when they face the real listening. Besides, the investigation of lexical bundles in the listening part of TOEIC preparation course research appears to be the least number of recent studies. Therefore, this study aims to identify the structural and functional features of lexical bundles in TOEIC preparation listening scripts

3. Method

3.1 Sample

The sample is selected by the availability of the preparation TOEIC book at Kamphaeng Phet Rajabhat university Maesot. Longman preparation series were selected. The data were obtained from the listening scripts in TOEIC Longman Preparation books. There are two different parts of the data: scripts from the test strategy sections and script from the TOEIC mock listening tests. All of the data was in PDF files.

3.2 Data collection

There were two main criteria to be used in the selection process. The First criteria, the number of the words, should be approximately the same numbers which were 12,000 – 14,000 for scripts from the test strategies and 3,300 – 4,000 for the script from the TOEIC mock listening tests. The second criteria, the books were selected by their beginners, intermediate and advanced levels. Therefore, the eight sets from the TOEIC listening practice tests and four sets from the test strategy sections were compiled in Notepad. Then, the notepad files were loaded into AntConc. (Ver. 3.5.8). The total of words in the corpus was 106,590.

Table 3. Composition of TOEIC Listening Corpus

Source	Word Counts
Test 1	3,526
Test 2	3,476
Test 3	3,728
Test 4	3,850
Test 5	3,952
Test 6	4,104
Test 7	3,387
Test 8	3,471
Test strategies practice 1	18,261
Test strategies practice 2	18,618
Test strategies practice 3	16,531
Test strategies practice 4	19,958
Total	102,862

3.3 Data Analysis

In order to find out the lexical bundles used in TOEIC listening scripts, there were five stages involved in this study. First, the cluster analysis in Antconc was used to classified the data. The program was set to search four-word bundles in the lexical bundles. Second, the top three hundred ranks of bundles in the corpus were used in data analysis. Third, in terms of data analysis, the researcher adopted the structural and functional taxonomy frameworks of lexical bundles by Biber et al. (2004) to analyse the corpus in order to extract four-word structural and functional bundles. Fourth, their categories and frequencies were recorded in a computation program (Excel.). Last, the structural and functional categories obtained from the analysis were counted their frequencies of each-sub categories and were calculated to be the percentage.

4. Findings and Discussion

After analyzing the data, the researcher found the frequency and percentage of the functional features and structures features. These were reported and discussed below.

Table 4. The frequency and percentage of the functional features

Functional category	Freq.	Percentage	Examples
A. Referential expressions			
1. Identification/focus	3	2.4	*It is not a*
2. Imprecision	0	0	
3. Specification of attributes			
Quantity specification	8	6.4	*The rest of the*
Tangible Framing	0	0	
Intangible Framing	0	0	
4. Time/place/Text reference			
Place reference	9	7.2	*On the ground floor*
Time reference	7	5.6	*This time of year,*
Text-deixis	4	3.2	*To the following excerpt*
Multi-functional reference	0	0	
B. Stance Expression			
1. Epistemic stance	5	3.96	*All right, I will*
2. Attitudinal /modality stance			
Desire	58	46.4	*Do you want to*
Obligation/directive	7	5.6	*Don't forget to*
Intention/Prediction	12	9.6	*It is going to*
Ability	2	1.6	*will be able to*
C. Discourse Organizer			
1. Topic introduction	7	5.6	the following news report
2. Topic Elaboration/ Clarification	3	2.4	*It is a very*
Total	**125**	**100**	

Table 3 shows the frequency and percentage of the functional features. As can be seen, the Stance Expression constituted the highest percentage of up to about 50.36 %. This suggests that the Stance Expression was most frequently used in the corpus. Desire (e.g., *do you want to*) was found to be the most common types, with 46.4 in this type. One

possible reason for this study has related to the purpose of the TOEIC test. That aims to test the everyday English skills of individuals who speak English as a foreign language. Therefore, people speak English to express their desires that reflect the authenticity of the language in the TOEIC test. Besides, Table 3 shows the frequency and percentage of the functional features. It can be seen the second common type that was referential expressions, accounting for 26.80 % of the bundles' type. The most significant of this bundles type was Time/place/Text reference (e.g., on the ground floor, this time of year). It suggests that the TOEIC listening also focuses on the place and time. Therefore, students should be prepared to learn about lexical bundles of place and time. Besides, the lesson should promote the use of a top-down process. For example, Teachers provide contents about time and place to prepare students for their listening.

Table 5. The frequency and percentage of the structural features

Structural category	Freq.	Percentage	Example
1. Lexical Bundles that incorporate verb phrase fragment			
a 1st/2nd person pronoun + VP fragment	27	21.42	I am looking for
3rd person pronoun + VP fragment	0	0	
discourse marker + VP fragment	7	5.55	I think I will
Verb phrase (with non-passive verb)	31	24.60	Is going to be
Verb phrase (with passive verb)	0	0	-
yes-no question fragments	7	5.55	Do you want to
WH- question fragments	17	13.49	Why did the woman
2. Lexical bundles that incorporate dependent clause fragments			
a 1st/2nd person pronoun + dependent clause fragment	0	0	
WH-clause fragments	1	0.79	When he says I
If-clause fragments	0	0	-
to-clause fragment	4	3.125	To let you know
That-clause fragment	0	0	-
3. Lexical bundles that incorporate noun phrase and prepositional			
Noun phrase with of-phrase fragment	5	3.96	
Noun phrase with other post-modifier fragment	8	6.34	following conversation with three
Other noun phrase expressions	0	0	-
Prepositional phrase expressions	18	14.28	At the same time
Comparative expressions	3	2.38	As long as you
	126	100 %	

Table 4 lays out the frequency and percentage of the structural features. As can be seen, the verb fragments made up the highest percentage of 66%, suggesting that the verb fragment was used most frequently in the corpus. Verb phrase (with non-passive verb) (e.g. is the most common type found in this category, accounting for 24.60 % of the bundles type. The results are contrasted with the study of Kashiha and Heng (2014). They found "1st/2nd person pronoun + VP fragment" that was their most common types. One possible reason for this study is related to the different sources of language use between TOEIC listening and chemistry and politics lectures. Even though 1st/2nd person pronoun + VP fragment" was not found in the first common types of the corpus, it still

constituted the second-highest percentage of up to about 21.42 %. It could be said that the findings were in line with the study of Kashiha and Heng (2014).

6. Conclusion

The purpose of the study was to study the structural and functional types of lexical bundles in the TOEIC listening scripts. From the most frequent four-word lexical bundles, they were identified to see the two types. The data were derived from the TOEIC listening scripts from TOEIC preparation books. In functional analysis, stance expression was found to be the most frequent use in the entire corpus. *For example, 'Do you want to' was used to express the desire to do something. In structural analysis,* Lexical Bundles that incorporate verb phrase fragment were appeared to be the most frequent use in the corpus. For example, they appeared in the form of *Verb phrase (with non-passive verb), such as* Is going to be to describe an action. The findings of this study would contribute to the pedagogical implication in terms of the frequent words that students can find in their TOEIC listening. It would help students to facilitate their actual listening comprehension.

8. Suggestions

8.1 Suggestions for pedagogical Implication

One main pedagogical implication of this study, the lesson should encourage students to know the frequent forms and functions of lexical bundles. For example, the form of the lexical bundles that consist of Verb fragments (non-passive verbs) appeared to be the most frequent use. It could be implied that the listening lessons should incorporate and focus on the use of verbs in various types. Another aspect of the form that should be focused on is prepositional phrase expressions. The lesson should provide students to listen to the frequent prepositional phrase to help students comprehend about listening for specific information such as location and time. Besides, the frequent function found in this study was to express desires and needs. The repetition activity may help students to link and comprehend, which may enhance their listening skills.

8.2 Suggestions for future studies

The study suggests that more research is needed to explore the usefulness of using teaching materials and lexical bundles lessons. Researchers may study each lexical bundle type to find out the effectiveness of the students' performance. Further research can divide the corpus into each part based on the parts of the TOEIC test and investigate the lexical bundle use in each part. For example, listening to TOEIC consists of four parts: pictures, questions and responses, conversation, and short- talk. Besides, As the present study focused on form and function, it is also recommended that further research would incorporate the meaning, such as looking at the topic where the lexical bundles have appeared in the conversation. This may help students to get familiar with the lexical bundles.

References

Batova, N. (2013) Academic listening: Is there a place for bottom-up processing? *International Journal of Education and Research*, 1(4):1-10.

Buck, G. (2001) Assessing listening. Cambridge: Cambridge University Press.

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. and Barbieri, F. 2007. "Lexical bundles in university spoken and written registers". *English for Specific Purposes*, 26(3), 263-286.

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, 14(3), 275- 311

Cook, G. 1997. Key concepts in ELT. *ELT Journal,* 51(1), 86.

Conrad, S., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica,* 54-71 Retrieved from https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1002&context=ling_fac

Chaudron, C., & Richards, J.C. (1986). The effect of discourse markers on the comprehension of lecture. *Applied Linguistics*, 7(2), 113-127.

Kashiha,H. & Heng, C., (2014). Structural Analysis of Lexical Bundles in University Lectures of Politics and Chemistry. *International Journal of Applied Linguistics and English Literature* 3(1) 224-230.

Lougheed, L. (2012). *Longman preparation series for the TOEIC test*: Advance Course (5th ed.). White Plains, NY: Pearson Education, Inc.

Lougheed, L. (2018). *Longman preparation series for the TOEIC test*: Introductory Course (6th ed.). White Plains, NY: Pearson Education, Inc.

Lougheed, L. (2018). *Longman preparation series for the TOEIC test*: intermediate Course (6th ed.). White Plains, NY: Pearson Education, Inc.

Lougheed, L. (2018). *Longman preparation series for the TOEIC test*: Advance Course (6th ed.). White Plains, NY: Pearson Education, Inc.

Kwon, Y.-E & Lee, E.-J. (2014). Lexical bundles in the Korean EFL teacher talk corpus: A comparison between non-native and native English teachers. *Journal of Asia TEFL*. 11. 73-103.

McEnery, A. M., & Wilson, A. (2001). *Corpus linguistics: an introduction*: Edinburgh University Press.

Morley, J. (2001). "Aural comprehension instruction: Principles and practices." In CelceMurcia, M. (Ed.) *Teaching English as a second or foreign language*. Boston: Heinle & Heinle, 69-85.

Nadig, A. (2013). Listening Comprehension. *Encyclopedia of Autism Spectrum Disorders*, 1743.

Neely, Elizabeth & Cortes, Viviana. (2009). A little bit about: Analyzing and teaching lexical bundles in academic lectures. *Language Value*. 1(1), 17 – 38

Nesi, H. and Basturkmen, H. (2006) Lexical bundles and discourse signaling in academic lecturers. *International Journal of Corpus Linguistics*, volume 11 (3), 283-304.

Richards, JC. (2008). *Teaching listening and speaking*: *From theory to practice,* Cambridge: Cambridge University Press.

Tsai, M. (2001). "Learning is a lifelong process." In Belcher, D. and U. Connor (Eds.) . *Reflections on Multiliterate Lives*. Tonawanda: Cromwell Press Ltd., 135-140.

Helgesen, M. (2003). Teaching listening. In: Nunan, D. (Ed.). *Practical English Language Teaching.* New York: McGraw-Hill.

Rahimi, AH. (2012). On the role of strategy use and strategy instruction in listening comprehension. *Journal of Language Teaching and Research*, 3(3), 550-559.

Vandergrift, L. (1997). The strategies of second language (French) listeners: A descriptive study. *Foreign Language Annals*, 30(3), 387-409.

Yeldham, M & Gruba, P. (2014). Toward an instructional approach to developing interactive second language listening. *Language Teaching Research*, 18(1), 33–53.