Discourse Functions of Lexical Bundles in Pakistani Chemistry and Physics Textbooks

Gulfam Hussain[a]
gulfamhussainqureshi@gmail.com
Air University, Islamabad, Pakistan

Tehseen Zahra[b]
tehsin.azhar@yahoo.com
Air University, Islamabad, Pakistan

Akhtar Abbas
asharabbas83@gmail.com
Air University, Islamabad, Pakistan

ABSTRACT

Lexical bundles (LBs) are indispensable building blocks and essential constituents of academic discourse. The appropriate utilization of the lexical bundle's approach can effectively enhance students' understanding of academic discourse. LBs in various academic genres have extensively been studied concerning written and spoken language. However, less research has been conducted to explore the occurrence, nature, and frequency of LBs in Pakistani academic discourse, especially in textbooks. Therefore, the present study aims to explore four-word common LBs and their functional taxonomies employed in Pakistani Higher Secondary School Certificate (HSSC) level textbooks of Chemistry and Physics. A specialized corpus of these textbooks was built which was run on Antconc software for the identification and extraction of the LBs in the corpus. The classification of the identified LBs was then carried out utilizing Biber functional taxonomies of LBs. The study found 102 LBs occurring in the selected textbooks. In functional categories, there was a dominant use of discourse organizers and referential expressions. The findings related to frequent strings of words which can have significant educational implications for teachers, language material developers, and syllabus designers. The list of LBs with discourse functions provided by this study can significantly be used to enhance students' academic writing and their ability to comprehend different types of scientific texts.

Keywords: Academic discourse; Corpus-based study; English for academic purposes; Lexical bundles; Pakistani science textbooks

INTRODUCTION

Lexical bundles are multi-word strings that generally recur in a text more than expected, such as, *on the other hand*, *at the same time*, *on the surface of,* and *at the end of.* Initially, lexical bundles (Henceforth LBs) were identified and explored by Biber, Johansson, Leech, Conrad, and Finegan (1999) as "bundle of words that show a statistical tendency to co-occur" (p.989) and as "recurrent expressions, regardless of their idiomaticity and regardless of their structural status" (p.990). In the existing literature, many terms have been used referring to LBs, such as lexical chunks (O'Keeffe et al. 2007), lexical phrases (Nattinger & DeCarrico, 1992), prefabricated structures (Yousaf & Shehzad, 2018), n-grams (Stubbs & Barth, 2003), formulaic sequences (Schmitt & Carter, 2004; Wray, 2002), multi-word expressions (Siyanova-Chanturia

& Martinez, 2014b) and lexical bundles (Biber & Conrad, 1998). All the expressions of labelling LBs mentioned above refer to recurring multi-word units known as LBs.

LBs are the most frequent and constituting significant academic discourse units; they are also known as basic building blocks, markers of proficiency, and significant components of fluent linguistic production in academic discourse (Hyland, 2008b; Kashiha & Heng, 2014a; Yousaf & Shehzad, 2018). The frequent utility of LBs illustrates "competent language use within a register to the point that learning conventions of register use may in part consist of learning how to use certain fixed phrases" (Cortes, 2004, p. 398). LBs help to shape meaning in a particular text and context; they also add to the sense of coherence in a text from a specific discourse and its perspective (Hyland, 2008a).

There are specific parameters to identify LBs. These parameters are based on the length of LBs, frequencies of occurrences, corpus size, and number of texts. For three-word LBs, they must appear ten-time per million words (PMWs); for four-word LBs, they must appear five-time PMWs, in more than a single text within a register (Biber et al., 1999). The presence of LBs in more than a single text within a register is essential to guard against peculiar features of the individual writer.

Textbooks are the essence of the academic discourse; they are authentic and valid. Textbooks present "the authorized version of a society's valid knowledge" (Olson, 1989, p. 238). Moreover, textbooks play a vital role in disseminating academic knowledge and are reliable source of knowledge for teachers and students both; therefore, the role of textbooks in the academic world to disseminate academic knowledge is undeniable. The role of academic language used in textbooks plays a critical role in teaching and learning process in any academic setting. LBs in this regard are one of the important academic language resources in making the text's writer responsible and reader-friendly. In the Pakistani context where the English language is officially declared medium of instruction students face several problems in producing effective academic genres within the classroom setting (Manan, Dumanig & David, 2017; Manan, 2019) in English as a second language. These problems include lacking employment of analytical skills through academic writing and inadequate command of the English language (Khan, Majoka & Fazal, 2015). The students mostly rely merely on grammar rules, which are not enough to produce stretches of academic discourse accurately. Furthermore, students also face difficulties while reading to comprehend various types of texts, especially science texts.

To study the issues of composition and comprehension, LBs have extensively been studied concerning spoken and written academic discourse belonging to various genres and registers. For instance, Biber et al. (1999) compared LBs in conversation with the LBs in academic prose. Similarly, Biber et al. (2004) studied LBs in textbooks and classroom teaching; they compared them with their earlier findings (Biber et al., 1999). Moreover, Biber (2006) identified similarities and differences in LBs across various disciplines, genres, and registers. Furthermore, LBs in English as a lingua franca concerning English self-study textbooks have also been explored (Allan, 2017). However, in the Pakistani context, there is little research conducted concerning LBs. For instance, LBs in Ph.D. theses across various disciplines have been identified and analyzed structurally (Yousaf & Shehzad, 2018); and noticeable interdisciplinary variations in LBs concerning frequencies and structures have been found. The understanding and the utility of LBs help learners in reading and producing communicative texts, especially science texts. It also has been established that complexities of scientific English used in science textbooks are not only difficult for the students of ESL but also for learners whose first language is English due to heavy nominalization and denseness (Halliday, 1993). Due to linguistic difficulty, English scientific texts are usually complex, incredibly ambiguous, and challenging for students of all ages to read (To & Mehboob, 2019). In Pakistan, by utilizing the list of LBs, an academic discourse of learners at Higher Secondary School

Certificate level can be made proficient. Students may face complications while dealing with the text of science; they might be incompetent in comprehending the text of science and write for science. Also, students may face problems in comprehending and understanding various phenomena in science textbooks, such as interpreting diagrams, tables, figures, and various procedures stated in textbooks. Developing competence of employing LBs communicatively can help students in producing responsible and reader-friendly academic discourse.

The lexical bundle approach (Conrad & Biber, 2005; Granger, 2014) is one of the approaches to enhance students' academic performance significantly. Pakistani students are neither taught LBs nor LBs are part of their study course. Therefore, the present study, in order to contribute in pedagogy, aimed to explore four-word common lexical bundles and their discourse functions in the selected Pakistani textbooks used at the Higher Secondary School Certificate level. Two-word sequences (lexical bundles) were not considered since these are word associations that do not have a distinct discourse-level function (Conrad & Biber, 2004). Three-word LBs are extremely common, tending to be of not much interest (Hyland, 2008a). Besides, three-word LBs are mostly included in the four-word LBs, such as '*on the other*' in '*on the other hand*'. Five- and six-word LBs were also not investigated because the longer the bundle is, the lower their frequencies will be (Simpson-Vlach & Ellis, 2010; Hyland, 2008a). Since the inquiry was made in specialized corpora, not in general corpora, four-word LBs seem the most suitable for extraction from the specialized corpora. The list of poly-word strings with discourse functions provided by this study can be significantly used to enhance learners' performance concerning academic discourse.

LITERATURE REVIEW

With the help of corpus linguistics and corpus-tools, corpora can be significantly used in Computer Assisted Language Learning (CALL). Learners are introduced to the useful teaching materials for practicing and learning a language from the real instances of language use. In most of the textbooks, the text presented as a resource material is not much useful rather based on intuitions of textbooks' authors or teachers, and the chunks of a language are not contextualized. Learners learn effectively when a language in real-life use is taught. Corpora offer learners understandable examples of language use from real instances (O'Keeffe et al., 2007). Also, through the help of corpora, the gap between what students learn in school and their out-of-school experiences can be minimized if textbooks and material designers are informed via corpora oriented studies to revise textbooks accordingly.

Moreover, with the help of corpora, useful teaching materials for English language teaching can be designed, such as teaching guides, course books such as Touchstone Series (McCarthy, McCarten & Sandiford, 2005), vocabulary books, a list of common phrases, grammar books and many more. Learners can get corpus-based materials for learning a language, such as handouts: various tasks, and activities (Johns, 1994). These activities are known as data-driven learning (DDL) activities, which are available in both print and online (Johns, 2002).

An analysis of corpora plays a vital role in the field of EAP/ESAP. For analysts, the advent and the use of corpora have made it possible and easy to examine language patterns, most frequent words, and frequent phrases in various domains. Moreover, it is easy to have thorough insights into a particular genre to explore its characteristics. The use of corpora in EAP is also known as an evidence-based approach to specific genres to understand what is typical in them. In EAP, this approach is employed to determine certain linguistic features for a particular kind of discourse. Concerning designing teaching-learning materials, useful corpus-informed dictionaries (Major, 2006; Rundell, 2007) and corpus-informed textbooks (Huntley, 2006; Swales & Feak, 2004) are produced through employing corpora.

Previously conducted research studies on LBs generally agree on the pedagogical value of LBs. Many studies not just focusing on the theoretical status/aspects of lexical bundles but also provide particular suggestions for pedagogical implications. As an established fact, pedagogically, the importance of LBs cannot be denied, especially in academic discourse. Similarly, a list of academic formulas for EAP curricula has been suggested (Simpson-Vlach & Ellis, 2010). Besides, the inclusion of LBs in learners' academic reading and writing tasks can improve their consciousness about them to perform well in academic reading and writing activities (Cortes, 2004; Hyland, 2008a). Further, there is a need to include better and more prolonged exposure to LBs in disciplinary writing courses (Cortes, 2006); it has been suggested after measuring the effectiveness of LBs' teaching through an experimental research study. Also, utilizing a set of LBs in academic lectures has been recommended (Neely & Cortes, 2009). Moreover, a list of twenty-one 4-word LBs used in science, arts, commerce, and law has been established for pedagogical implications (Byrd & Coxhead, 2010).

Further, LBs employed in Physics research articles have been structurally explored to generate a list of the most frequent bundles (Farvardin, Afghari & Koosha, 2012). In their results, from the top ten bundles, 3 bundles are also occurring in the selected textbooks for the present study, such as, *on the other hand, with respect to the,* and *in the case of*. Moreover, LBs in the lectures of politics and chemistry have also been investigated and explored structurally; and the explicit teaching of LBs has been emphasized for helping learners to acquire a language (Kashiha & Heng, 2014b). In their findings, very few bundles are similar to the present study, such as, *at the end of*, which is used as a referential expression to refer to time, place, or the text.

Moreover, LBs in various disciplines in university students' writing have been investigated (Durrant, 2017); a corpus has been compiled in several disciplines such as science/technology, humanities/social sciences, life sciences, and commerce. Notable variations were found across these disciplines. Furthermore, differences within the disciplines among writers have also been found. It has also been found that most of the disciplines are relatively internally homogenous (Durrant, 2017). Biber et al. (2004) studied lexical bundles in university lectures and textbooks. In their data, textbooks from natural sciences were included. They found noticeable differences in lexical bundles in classroom lectures and textbooks. Similarly, Farvardin, Afghari & Koosha (2012) studied lexical bundles in Physics research articles. They identified lexical bundles structurally; however, they did not focus on the discourse functions of these bundles. A study has also investigated anticipatory 'it' lexical bundles in Linguistics and Chemistry research articles (Jalali, 2014). Contrary to our expectations, no similarities were found in LBs between the results of his study and the findings of the present study.

Summing up, all the studies discussed above have addressed LBs in various genres and registers, both in spoken and written academic discourse. Considering recommendations of these studies, they emphasize the pedagogical importance of lexical bundles in academic discourses, as LBs are the building blocks of academic discourse. The research studies reviewed here justify the need of the present research study to explore LBs to create a list of bundles that can be utilized while making decisions for EAP/ESAP pedagogy, especially in the Pakistani context. The list of LBs provided by this study may inform the practice of the English language in other ESL settings worldwide.

The present study sought to identify four-word common lexical bundles in Pakistani textbooks of Physics and Chemistry used in Khyber Pakhtunkhwa at an intermediate level. Although several textbook boards are operational in Pakistan, the textbooks of the Khyber Pakhtunkhwa textbook board were selected. Some of the books, such as Physics, Math, and English of Khyber Pakhtunkhwa textbook board, are also preferably implemented in the federal colleges. The rationale for a number of textbooks selected has been given in the methodology

section. The current research study explores the functional taxonomies of LBs used in these textbooks. The present study has the following objectives:

- To identify four-word common core lexical bundles occurring in Pakistani Higher Secondary School Certificate level textbooks of Physics and Chemistry.
- To classify the discourse functions of four-word common core lexical bundles occurring in Pakistani Higher Secondary School Certificate level textbooks of Physics and Chemistry.

THEORETICAL FRAMEWORK

In this study, the functional taxonomies of lexical bundles devised by Biber et al. (2004, pp.381-384) were used as the theoretical basis. Biber et al. (2004) investigated lexical bundles in university lectures and textbooks. The functional taxonomies used in Biber et al. (2004), for the classification of functional categories of lexical bundles were adopted from Biber et al. (2003), in which functional types of lexical bundles developed for conversation and academic prose were discussed.

For this study, the functional taxonomies used in Biber et al. (2004) seem the most suitable as in many research studies focusing on lexical bundles, this framework has been used and found reliable (see, for example, Allan, 2016; Biber et al. 2007; Hyland, 2008; Kashiha & Chan, 2015; Kashiha & Heng, 2014; Simpson-Vlach & Ellis, 2010; Tomankova, 2016). The present study, therefore, adopted the functional taxonomies from Biber et al. (2004), which have been modified due to the new functional categories explored in textbooks.

In Biber et al. (2004), three main functional categories of lexical bundles: stance expressions, referential expressions, and discourse organizers have been identified. Besides these functional categories of lexical bundles, special conversational function groups have also been identified. Referential bundles "make direct reference to physical or abstract entities or to the textual context itself" (Biber et al., 2004, p.384). Furthermore, "stance bundles express attitudes or assessments of certainty that frame some other proposition" (Biber et al., 2004, p.384). Lastly, discourse bundles, according to Biber et al. (2004), negotiate and arrange the flow of discourse by providing links to the previous and coming sections. Including subcategories of lexical bundles, Table 1 presents the functional taxonomies of lexical bundles used in Biber et al. (2004).

TABLE 1. Functional taxonomies of lexical bundles in Biber et al. (2004)

Functional categories of lexical bundles	Examples
1. Stance Expressions	
• Epistemic stance	
➢ Personal	*I don't know if, I think it was*
➢ Impersonal	*are more likely to, the fact that the*
• Modality/attitudinal stance	
➢ Desire	
▪ Personal	*I don't want to, what do you want*
➢ Obligation/directive	
▪ Personal	*you need to know, I want you to*
▪ Impersonal	*it is necessary to, it is important to*
➢ Intention/prediction	
▪ Personal	*I am going to, are we going to*
▪ Impersonal	it*'s going to be, are going to be*
➢ Ability	
▪ Personal	*to come up with, to be able to*
▪ Impersonal	*it is possible to, can be used to*
2. Discourse Organizers	

	• Topic introduction/focus	*take a look at, what to do is*
	• Topic elaboration/clarification	*on the other hand, nothing to do with*
3.	Referential Expressions	
	• Identification/focus	*is one of the, one of the most*
	• Imprecision	*and stuff like that, or something like that*
	• Specification of attributes	
	➢ Quantity specification	*have a lot of, in a lot of*
	➢ Tangible framing attributes	*in the form of, the size of the*
	➢ Intangible framing attributes	*in terms of the, in the case of*
	• Time/place/text reference	
	➢ Place reference	*in the united states, of the united states*
	➢ Time reference	*at the same time, at the time of*
	➢ Text deixis	*as shown in the figure, shown in figure N*
	➢ Multi-functional reference	*at the end of, the top of the*
4.	Special Conversational Functions	
	• Politeness	*thank you very much*
	• Simple inquiry	*what are you doing*
	• Reporting	*I said to him/her*

RESEARCH METHODOLOGY

Corpus Linguistics is a branch of linguistics that provides tools and methods for corpus construction and its analysis (Zahra & Abbas, 2018). In the present study, corpus methods were used to identify lexical bundles with their functions in Pakistani textbooks. Moreover, a mixed-method approach: both quantitative and qualitative, were employed in the present study. A quantitative analysis was performed to explore the number and frequencies of lexical bundles. In addition, a qualitative analysis was carried out to explore discourse functions of lexical bundles.

RESEARCH DATA

Higher Secondary School Certificate level textbooks were selected for this research. In Higher Secondary School Certificate (HSSC) level textbooks, Physics and Chemistry textbooks (Physics part I and II, and Chemistry part I and II) were selected. The selected textbooks have been published by Khyber Pakhtunkhwa (one of the provinces of Pakistan) Textbook Board. Although several textbook boards are operational in Pakistan, the textbooks of Khyber Pakhtunkhwa Textbook Board were selected. Some of the textbooks, such as Physics, Math, and English of Khyber Pakhtunkhwa Textbook Board, are also preferably implemented in the colleges of Federal Capital of Pakistan i.e. Islamabad. Textbooks of various textbook boards are used for academic practices in different provinces of Pakistan; and a decision of the selection of textbooks for schools and colleges in Islamabad is made based on the quality of the textbooks. The textbooks published by Khyber Pakhtunkhwa Textbook Board are not only used in Khyber Pakhtunkhwa province but also preferred over other textbooks for schools and colleges in the capital territory. Moreover, due to the unavailability of soft copies of the selected textbooks and suitable textbooks' corpora for the present study, we felt the need to construct a specialized corpus. Besides, it is time consuming and costly to construct a corpus from those textbooks which are available only in hard form. Therefore, we selected an adequate number of textbooks in hard form. Considering the issues of copyright, the text of these books have only been used for the research purpose only in this study. The text used in the current research cannot be used for any commercial purposes, or it will not be reproduced or published for commercial benefits. Lexical bundles and few extracts have been taken from the textbooks to use in the current study. Moreover, this corpus will not be made publicly available in order to consider ethical issues at maximum.

CORPUS CONSTRUCTION AND ANALYSIS

For corpus compilation, the selected textbooks were scanned to make them digital. The corpus compiled for this study consists of four text-files: Physics Part One, Physics Part Two, Chemistry Part One, and Chemistry Part Two. After scanning the books, Free-OCR software (version 5.41) was used; it is Optimal Character Recognition software freely available on the internet; scanned pages were passed through OCR software in order to get a digital/editable version of them. The digital form of the text was copied from OCR and was pasted in Microsoft office MS Word 2007. Files were passed through Text-fixer software to clean the data by removing unnecessary spaces like line spaces and paragraph spaces. This software is freely available online (https://www.textfixer.com). Once the data was collected and cleaned through the respective procedure, Microsoft word files were converted into text files through a free online file converter (https://www.online-convert.com).

When the files were converted into text-files, they were loaded on AntConc version 3.5.2 software to extract lexical bundles from the text-files for analysis. AntConc software was used for the extraction of LBs. In this software, the Clusters/N-Gram option was used to get the list of the desired lexical bundles. Therefore, cluster minimum, as well as a maximum size, was set on 4. Besides, the minimum frequency was set on 4, and the minimum range was also set on 4. In the list, LBs are common across the textbooks, which signals its occurrence in all the selected textbooks.

CORPUS COMPOSITION

For this study, a specialized corpus named Pakistani Corpus of Science Textbooks (PCST) was constructed in order to achieve specific research objectives. The overall size of the corpus is 275981 words. This corpus consists of four text files: Physics I (68567 words), Physics II (78418 words), Chemistry I (63798 words), and Chemistry II (65198 words). The details of the words in the corpus and of these text-files have been presented in the following Table 2:

TABLE 2. Composition of the Corpus: Pakistani Corpus of Science Textbooks

Textbooks	No. of Texts	Word Types	Word Tokens
Physics Part One	1	4234	68567
Physics Part Two	1	4944	78418
Chemistry Part One	1	4112	63798
Chemistry Part Two	1	5606	65198
Total	**4**	**18896**	**275981**

FINDINGS AND DISCUSSION

After generating the list of LBs by AntConc 3.5.2, manual filtration was carried out; few lexical bundles were excluded because they have no distinctive and clear functions. The rest of the bundles were analyzed functionally. We found 102 common four-word lexical bundles used in the selected textbooks (See Appendix).

In Biber et al. (2004), four functional categories of LBs were proposed: stance expressions, discourse organizers, referential expressions, and special conversational functions. Firstly, as the current study made inquiries into LBs employed in the textbooks; therefore, the study found only one LB *what is the difference* with special conversational function that is simple inquiry. In other words, the findings did not reveal more LB with special conversational functions (e.g. politeness, simple inquiry, and reporting). After a thorough functional analysis of LBs, the findings revealed that there is only one bundle with special conversational function. Secondly, the results of this analysis confirms only one stance

expression i.e. modality/attitudinal stance/ability: impersonal, the examples are: *can be used to, can be used for* occurring in the corpus. Thirdly, an interesting side of findings was that discourse organizers and referential expressions are frequently used in the corpus of the selected textbooks. Out of 102 lexical bundles, majority of them i.e. 52 are employed for discourse organizing functions as the science textbooks include a lot of discussion regarding the various phenomena. In discourse organizer functions, topic introduction/focus such as *the study of the, in this chapter we,* and topic elaboration/clarification functions such as *is known as the, this process is called,* etc.*,* were some of the most typified bundles found. In the discourse organizer category, the analysis confirms the dominant use of LBs for the topic elaboration/clarification function as the science textbooks present the objective reality regarding various phenomena that needs elaboration and clarification of various topics to be made in order to make text reader friendly. The results yielded consist of 12 bundles in textbooks that are used for comparison purposes. To have a clear understanding, it is common practice to compare different ideas, entities in the science text books. 10 bundles were found used for cause and effect function. In Physics and Chemistry textbooks, it is necessary to have a description of various processes that need to discuss things with causes and effects. Fourthly, after discourse organizers, our findings reveal the frequent use of LBs for referential expression functions. The results obtained by the present study include 32 LBs used for making various types of references, such as time, text, and place references. In referential expressions, the results show the dominant use of tangible framing attribute and quantity specification functions. Lastly, the findings confirm that 7 LBs were used for stance expression functions. The analysis proves that stance expressions are used to show impersonal ability such as, *can be used to* and *can be used for*. Contrary to our expectations, the study found no modality/attitudinal stance as the writers of the science textbooks may not take attitudinal or personal ability stance.

In addition to above mentioned discourse functions, the findings of our study contribute some more functions into the discourse organizers' functional category. All the functional categories of lexical bundles with this contribution to the modal have been stated in Table 3 with examples from the textbooks. The new functions (highlighted as bold in Table 3) which have been added to the Biber et al. (2004) functional taxonomies into discourse organizers are cause and effect (e.g. *is due to the, this is because the, as a result the*), procedure/process (e.g. *is passed through a, is added to the*), comparative/comparison (e.g *same as that of, as compared to the*) and description (e.g. *is a measure of, from left to right*). However, surprisingly, we did not find any disciplinary differences in the use of LBs (functions) across the textbooks of Physics and Chemistry indicating closely related conventional practice of discourse.

It is also important to note that the findings of the current study do not make any further functional contribution to Biber et al. (2004) functional taxonomies except for discourse organizers. All referential expressions in Biber et al. (2004) such as identification/focus, specification of attributes, and time/place/text reference, are present in the selected science textbooks except for *Imprecision*. All functional taxonomies of LBs that are present in the science textbooks have been mentioned in the following Table 3 with examples. It is important to state some of the examples of the discourse functions of lexical bundles. Therefore, the examples of each functional category have been taken from the textbooks discussed in the next section of the paper.

Numerous scholars have conducted research on lexical bundles in the area of science concerning Physics and Chemistry research articles. For instance, Farvardin, Afghari & Koosha (2012) have studied lexical bundles in Physic research articles. They explored lexical bundles structurally but not functionally. In their results, from the top ten bundles, 3 bundles are similar to bundles found in our study which are present in the selected science textbooks

such as, *on the other hand, with respect to the,* and *in the case of.* Jalali (2014) investigated anticipatory 'it' lexical bundles in Chemistry and Applied Linguistics research articles, surprisingly, the results obtained from the chemistry articles show no similarities with the results of the present study. Similarly, Kashiha and Heng (2014b) identified lexical bundles structurally in university lectures on Politics and Chemistry. In their study, very few bundles are similar to the bundles found in the present study. For example, *at the end of* is used as a referential expression to refer to time, place, or the text.

TABLE 3. Functional taxonomies of lexical bundles used in textbooks

Functional Categories of LBs in Textbooks	Examples	Numbers of LBs
Stance Expressions		**07**
Modality/attitudinal stance		
Ability Impersonal	*can be used to, can be used for*	07
Discourse Organizers		**52**
Topic introduction/focus	*the study of the, in this chapter we*	03
Topic elaboration/clarification	*on the other hand, this process is called*	21
Cause and effect	*this is because the, is due to the*	10
Process/procedure	*is added to the, is passed through a*	05
Comparative/comparison	*as compared to the, same as that of*	12
Description	*from left to right, is a measure of*	**0**1
Referential Expressions		**32**
Identification/focus	*one of the following, is one of the*	06
Specification of attributes		
Quantity specification	*a large number of, is the amount of*	09
Tangible framing attributes	*in the form of, the size of the*	03
Intangible framing attributes	*the nature of the, in such a way*	14
Time/place/text reference		**10**
Place reference	*on the surface of, the bottom of the*	03
Time reference	*the same time the, at the same time*	02
Text deixis	*as shown in the, shown in the fig*	04
Multi-functional reference	*at the end of*	01
Special Conversational Functions		**01**
Politeness	------------	
Simple inquiry	*what is the difference*	01
Reporting	------------	

STANCE EXPRESSIONS

In stance expressions, ability/impersonal functions such as <u>*can be used to*</u> is one the most prototypical bundles used in the science textbooks. This LB has been used for ability as a modality stance in both Physics and Chemistry. The following examples have been taken from textbooks:

1. *Following methods <u>can be used to</u> detect these elements, in the organic compound* (Chemistry).
2. *The pattern of energy absorption <u>can be used to</u> produce a computer enhanced photograph* (Physics).

Figure 1 shows some of the occurrences of this lexical bundle.

10	of Sn remains -0.14, cell potential can be calculated as follows. Table 12.1	can be used to	predict the reaction and voltage of any cell consisting of an	Chemistry P.
11	gen as well as oxygen, nitrogen, sulphur and halogens. Following methods	can be used to	detect these elements, in the organic compound. As carbon	Chemistry P.
12	s reagent, Fehling's solution and Benedict's solution. These three reactants	can be used to	distinguish aldehydes from ketones as ketones are not oxid	Chemistry P.
13	ups. An example of the class is acid yellow 36 (Metanil yellow). These dyes	can be used to	dye wool or cotton with a mordant but are usually used for	Chemistry P.
14	ongue across the lips. Water-soluble dyes such as green or blue food dyes	can be used to	provide lipstick colouration, but they are, usually, first comb	Chemistry P.
15	i-clockwise sense. Its direction can be determined as, The above Equations	can be used to	find the resultant of any number of Thus by adding the x-co	Physics Part
16	acting. Let F, be the net force acting on ml and is given by: Eq. ii and Eq. iii	can be used to	find the magnitudes of effective forces acting on m1 and m	Physics Part
17	vitational constant, mass and the radius of the orbit. Analyze that satellites	can be used to	send information between places on the earth which are fa	Physics Part
18	rs, energy is absorbed by the molecules. The pattern of energy absorption	can be used to	produce a computer enhanced photograph. Cooking of Foc	Physics Part
19	ance damage. A thermistor with a negative temperature co-efficient (NTC)	can be used to	issue an alarm for excessive temperature of winding of mo	Physics Part
20	e circular motion of an electron shot perpendicularly into a magnetic field	can be used to	determine its charge to mass ratio. Using the equation Prac	Physics Part
21	to secondary voltages. describe how step-up and step-down transformers	can be used to	ensure efficient transfer of electricity along cables. Electricit	Physics Part
22	ds to spectral lines. explain how the uniqueness of the spectra of elements	can be used to	identify an element. analyse the significance of the hydroge	Physics Part
23	as an The result derived above for the energy levels along with postulate 4	can be used to	derive the expression for the wavelength of the hydrogen s	Physics Part
24	ally the same. Thus, similar to the Balmer empirical formula, Bohr's theory	can be used to	compute the energies or wavelength of the transitions invo	Physics Part
25	potential energies are zero at n = \xB0\xB0 level. The energy level diagram	can be used to	illustrate the origin of various spectral series observed in th	Physics Part
26	t crystals gave birth to x -ray crystallography. The Laue diffraction pattern	can be used to	determine the internal structure of the crystals. The spacing	Physics Part
27	cal tool for "welding" detached retains. , The narrow intense beam of laser	can be used to	destroy tissue in a localized area. Tiny organelles with a livi	Physics Part
28	rints. It can be used in telecommunication along optical fibers. Laser beam	can be used to	generate three dimensional images of objects in process ca	Physics Part
29	ality or abnormality of the specific parts of the body. Radioactive particles	can be used to	trace chemicals participating in various reactions one of the	Physics Part

FIGURE 1. A screen shot of LB *can be used to*

DISCOURSE ORGANIZERS

In discourse organizers, *topic introduction/focus* function, such as *the study of the* was found in the science textbooks. The examples are:

3. *The study of the chemistry of carbon containing compounds (organic compounds) is called organic chemistry* (Chemistry).
4. *The study of the light emitted from the sun and from distant stars gives information about their composition infrared* (Physics).

Moreover, in discourse organizers, *topic elaboration/clarification* functions such as *on the other hand* are there in textbooks. Examples are:

5. *These on one hand, are essential for human body, but on the other hand, if their concentrations are greater than about 500 ppm, they make water unfit for drinking, and such water is considered to be polluted* (Chemistry).
6. *The total Kinetic energy, on the other hand, is generally not conserved in a collision because some of the Kinetic energy is converted into internal energy* (Physics).

In addition, *cause and effect* function was found in the textbooks such as *is due to the.* The following examples are from textbooks.

7. *Paramagnetic behavior is due to the presence of one or more unpaired electron in an atom, ion or molecule of the substance* (Chemistry).
8. *This property (i.e inductance) is due to the self-induced e.m.f. in the coil itself by the changing current* (Physics).

Furthermore, in discourse organizers, *process/procedure* functions such as *is passed through a* were also identified (See 9 and 10 below).

9. *When light radiation is passed through a prism, it bends* (Chemistry).
10. *When an electric current is passed through a coil of wire, the coil act like a bar magnet with a north pole at one end and south at the other* (Physics).

In discourse organizer category, *comparative/comparison*, such as *as compared to the* has been found in textbooks (See 11 and 12 below).

11. *The size of the gas molecule is very small as compared to the distance between them* (Chemistry).
12. *It should be noted that the central maximum is of a high intensity and very broad as compared to the other maximum* (Physics).

Moreover, *description* function such as *from left to right* were also found in the textbooks (See 13 and 14 below).

13. *The reaction from left to right, as the equation for the reaction is written is known as the forward reaction and the reaction from right to left is known as the back reaction* (Chemistry).

14. *The current is flowing from left to right i.e. in opposite direction of electron flow* (Physics).

REFERENTIAL EXPRESSIONS

In referential expressions, *Identification/focus* functions such as *one of the following* are identified in textbooks. The following are examples i.e 15 and 16 from the textbooks.

15. *To get two coherent waves from a point source, one of the following two methods is adopted* (Physics).

16. *Which one of the following is not a secondary pollutant?* (Chemistry).

SPECIFICATION OF ATTRIBUTES

In referential expressions, specification of attributes such as *quantity specification* functions such as *a large number of* occur in textbooks. The following examples i.e 17 and 18 are from the textbooks.

17. *Plants have always been a rich source of a large number of organic compounds* (Chemistry).

18. *The liquid is rapidly volatized and a large number of bubbles are formed* (Physics).

In referential expressions, *tangible framing attributes* such as *the size of the* are also found in textbooks. The following examples 19 and 20 are from textbooks.

19. *The value of n is associated with the size of the shell* (Chemistry).

20. *The electric charges qi and are assumed to be point of localized charges, provided the size of the bodies carrying the charges is very small as compared to the distance between them* (Physics).

In the same functional category, *Intangible framing attributes* functions are also explored such as *the nature of the* are also there in textbooks. The following examples 21 and 22 are from textbooks.

21. *The acidities of the carboxylic acid vary considerably with the nature of the substituents present in the molecule* (Chemistry).

22. *In his theory, Plank made two assumptions, which at that time bold and controversial, concerning the nature of the oscillating charges of the cavity walls* (Physics).

TIME/PLACE/TEXT REFERENCE

In textbooks, in referential expressions, *Place reference* such as *the bottom of the* were also found, the following examples 23 and 24 have been taken from the textbooks.

23. *NaOH is collected at the bottom of the cell* (Chemistry).

24. *The air and water together are forced to the bottom of the filter pump* (Physics).

In referential expressions, *time reference functions* such as *at the same time* are also identified in textbooks. Here are some examples from textbooks.

25. *A system may lose energy to the surroundings in the form of heat but at the same time the same amount of energy is absorbed by the surroundings* (Chemistry).

26. *If the whole heart muscle contracted <u>at the same time</u>, there would be no pumping effect* (Physics).

In referential expressions, *text deixis* functions such as <u>shown in the fig</u> are also present in textbooks, here are few examples (27 and 28) from the textbooks given below:

27. *In this way a temporary dipole is created in the atom as <u>shown in the Fig</u> 5.3* (Chemistry).
28. *Consider a thin film of a refracting medium with a thin wedge shape structure and refractive index as <u>shown in the fig</u> 9.11* (Physics).

In referential expressions, *multi-functional reference* also occurs in textbooks, such as <u>at the end of.</u> Here are examples (29 to 31) from the textbooks.

29. *Choose the suitable answer from the following choices given <u>at the end of</u> the question* (Chemistry).
30. *Both theories, Discovered <u>at the end of</u> 19th century* (Physics).
31. *A mass <u>at the end of</u> spring describes S.H.M with T= 0.40s* (Physics).

Moreover, Figure 2 presents some of the occurrences of the respective functions.

1	s units. In this case an extra step is added (either at the beginning or at the end of the procedure) that we have just learned. It is important to recogniz Chemistry Pa
2	ometric amounts are reacted together, there will be no reactants left at the end of the reaction. In actual practice, however, due to one reason or the ot Chemistry Pa
3	consumed in the reaction and are recovered chemically unchanged at the end of the reaction. Very small amount of catalyst is needed for speeding u Chemistry Pa
4	nisms. Choose the suitable answer from the following choices given at the end of the question. Activated complex is a substance which is A reaction is Chemistry Pa
5	e common at the beginning of series and +2 state is more common at the end of series as shown in table 14.3. The highest oxidation state from Sc+3 Chemistry Pa
6	ze ethane to ethyl alcohol. Select the most appropriate choice given at the end of each question? The Transition metals Copper and are best electrical Chemistry Pa
7	alkenes in electrophilic reactions. Alkynes in which the triple bond is at the end of the chain, are refered to as terminal or 1 - alkynes. Terminal alkynes Chemistry Pa
8	one body. Centrifugal force or Reaction Force When we whirl a ball at the end of a string, we transmit this force to the ball by means of string, pulling Physics Part
9	an position, If spring constant of spring is 0.4 N m^1 and its velocity at the end of this displacement be 0.4 m s . Calculate (i) Time period 'T (ii) Freque Physics Part
10	one of the halves, the frequency is '72'\x95 Find out fxf2~l. A mass at the end of spring describes S.H.M with T = 0.40 s. Find out a when the displace Physics Part
11	nd the amplitude, frequency and time period of an object oscillating at the end of a spring , if the equation for its position at any instant t is given by x Physics Part
12	source covers distance V after 1 second and reaches the listener "L". At the end of one second the source covers distance "a" where it gives the last wa Physics Part
13	eory of relativity and the quantum theory. Both theories, Discovered at the end of 19th century, revolutionized physics in A GPS satellite is a satellite u Physics Part
14	e is about 10~*S. What is its uncertainty in energy during this time? At the end of this chapter the student will be able to: describe and explain the orig Physics Part

FIGURE 2. A screen shot of LB *at the end of*

SPECIAL CONVERSATIONAL FUNCTIONS

In this particular functional category, the study found only one lexical bundle occurring in the textbooks i.e. *what is the difference*. This LB comes under the category of simple inquiry. Since the text for analysis of this study was from the science textbooks, that is the reason that there cannot be conversational functions except for simple inquiry. Through this lexical bundle, specific simple inquires in the books have been made.

CONCLUSION AND RECOMMENDATIONS

This corpus-based study has generated a list of 102 four-word common core LBs found across the selected textbooks. All the LBs have been analyzed functionally. Furthermore, examples have been stated from the textbooks in order to state the discourse functions of these common core LBs across the textbooks. The list of LBs with their functions provided by this study offers an insight to inform EAP practitioners in the field of research and pedagogy. Some previous studies, such as Cortes (2006), Jones and Haywood (2004), and Byrd and Coxhead (2010), strongly recommend the utility of LBs as the basis for material design and curriculum development. In our point of view also, the list provided by this study can be significant to enhance learners' skills of writing and reading both. As these common LBs occur in Pakistani textbooks are corpus-informed (list of LBs) material/results, therefore, this corpus-informed

material is likely to be effective if used for pedagogical purposes in syllabus design and curriculum development in the ESL context of the country. Similar to the findings of Cortes (2006), who conducted an experimental study and noticed enhanced awareness and interest of the learns regarding use of LBs in language learning, we also assume LBs can be quite effective in increasing the learners' awareness and interest in the context of Pakistan.

Furthermore, Jones and Haywood (2004) also gauged the effectiveness of teaching and learning LBs; their results showed that LBs are highly effective in improving students' academic discourse. In this way, the LBs occurring in textbooks can be significant for Pakistani students also to enhance their reading comprehension and writing for academic purposes. Keeping in mind the above view, the list of common core four-word lexical bundles developed in the current study can be crucially important to play a vital role in designing syllabus for Higher Secondary School Certificate level. This list would definitely enhance learners' ability to read (in terms of comprehension) and write (writing in general and academic writing) effectively. Moreover, these LBs can also enrich the creative writing of the learners. If students are pedagogically exposed to this list of LBs, it would not just improve their reading and writing but also their speaking (establishing coherence in what they say) and listening (comprehending what they listen to) skills.

By utilizing the list of LBs provided by this study, various assessment tests can be designed to check the students' level of proficiency in English language skills. Furthermore, supplementary materials can also be designed using the list of LBs provided by this study for learners to practice different tasks and activities of reading and writing.

In the Pakistani context, the language issues (such as inadequate command of the English language and analytical skills) faced by the students (Khan, Majoka & Fazal, 2015) can be resolved to greater extent by incorporating corpus-informed pedagogy in general and knowledge of language building blocks i.e LBs in pedagogy in particular.

For enhancing students' discourse competence, LBs can be made part of different classroom tasks and activities to give practical exposure of the understanding and the utility of LBs in academic discourse. Various writing and reading tasks/activities can be designed, such as identifying functions of LBs, using LBs according to the generic conventions, and making use of LBs ESP contexts. Students can be asked to write paragraphs on topics by providing them a list of LBs to be used according to their functions. Akin to proposals of Wray (2002) and Millar (2011), instructions on the identification and the utility of LBs can effectively enhance students' discourse competence.

Similar to the recommendations of Biber (2006) and Swales and Feak (2012), the list provided by this corpus-based study can also be useful for students performing various academic writing tasks in classroom.

The role of LBs in text books of science cannot be neglected as Beng and Keong (2015) state that lexical bundles can help learners to comprehend science-based texts to indicate location, description, and quantity specification. Therefore, the science students of Pakistan can benefit significantly by incorporating LBs in their science textbooks.

Based on the current research, several future research studies can be conducted on lexical bundles. Future researchers could go for structural classifications of the lexical bundles in selected textbooks. Moreover, an experimental research study can also be conducted to measure the effectiveness of the teaching of lexical bundles in academic discourse. The list of lexical bundles with discourse functions obtained from this study can be compared with the results from different genres of the same domain or other genres. In sum, EAP, ESP, ERP especially and ELT in general are the avenues for future research on the role of LBs in various discourses.

REFERENCES

Allan, R. (2016). Lexical bundles in graded readers: To what extent does language restriction affect lexical patterning? *System*, *59*, 61-72.

Allan, R. (2017). From Do You Know to I Don't Know: An Analysis of the Frequency and Usefulness of Lexical Bundles in Five English Language Self-Study Books. *Corpus Pragmatics*, *1*(4), 351-372.

Beng, C. O. S., & Keong, Y. C. (2015). Functional types of lexical bundles in reading texts of Malaysian University English test: A corpus study. *GEMA Online® Journal of Language Studies*, *15*(1).

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam, Netherlands: John Benjamins.

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. (1999). Lexical bundles in conversations and academic prose. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 181–190). Amsterdam, Netherlands: Rodopi.

Biber, D., Conrad, S., & Cortes, V. (2003). Lexical bundles in speech and writing: An initial taxonomy. In A. Wilson, P. Rayson & T. McEnery (Eds.), *Corpus linguistics by the Lun : A festschrift for Geoffrey Leech* (pp. 71–93). 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*(3), 371-405.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge, UK: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*. London, UK: 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,* 31-64.

Conrad, S. M., & Biber, D. (2004). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica*, *20*, 56-71.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, *23*(4), 397-423.

Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education, 17,* 391-406.

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, *38*(2), 165-193.

Farvardin, M. T., Afghari, A., & Koosha, M. (2012). Analysis of Four-Word Lexical Bundles in Physics Research. *Advances in Digital Multimedia*, *1*(3), 2166-2916.

Granger, S. (2014). A lexical bundle approach to comparing languages: Stems in English and French. *Languages in Contrast*, *14*(1), 58-72.

Halliday, M. A. K. (1993). The construction of knowledge and value in the grammar of scientific discourse: Charles Darwin's The Origin of the Species. In M. A. K. Halliday & J. K. Martin (Eds.), *Writing science: Literacy and discourse power* (pp. 86-105). Washington, London: Falmer.

Howarth, P. A. (2013). *Phraseology in English academic writing: Some implications for language learning and dictionary making* (Vol. 75). Walter de Gruyter.

Huntley, H. (2006). Essential academic vocabulary: Mastering the complete academic word list. *The CATESOL Journal, 19*(1), 210-219.

Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4-21.

Hyland, K. (2008b). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, *18*(1), 41-62.

Jalali, H. (2014). A Comparative Study of Anticipatory it Lexical Bundles in Applied Linguistics and Analytical Chemistry Research Articles. *Journal of Foreign Language Teaching and Translation Studies*, *3*(1), 24-42.

Johns, T. (2002). Data-driven learning: The perpetual challenge. In Kettemann, B., & Marko, G. (Eds.), *Teaching and learning by doing corpus analysis* (pp. 105-117). Amsterdam, Netherlands: Brill Rodopi.

Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context. In N. Schmitt (Ed.), Formulaic sequences (pp. 269-291). Amsterdam, Netherland: John Benjamins.

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.

Kashiha, H., & Heng, C. S. (2014a). Discourse functions of formulaic sequences in academic speech across two disciplines. *GEMA Online® Journal of Language Studies*, *14*(2).

Kashiha, H., & Heng, C. S. (2014b). Structural analysis of lexical bundles in university lectures of politics and chemistry. *International Journal of Applied Linguistics and English Literature*, *3*(1), 224-230.

Khan, M. I, Majoka, M. I, & Fazal, S. (2015). Post/graduate academic writing problems: A Pakistan case. In Badenhorst, C., & Guerin, C. (Eds.), *Research literacies and writing pedagogies for masters and doctoral writers* (pp. 389-406). Netherlands: Brill.

Major, M. (2006). *Longman exams dictionary*. Harlow, London, UK: Pearson Education Limited.

Manan, S. A. (2019). Myth of English teaching and learning: A study of practices in the low-cost schools in Pakistan. *Asian Englishes*, *21*(2), 172-189.

Manan, S. A., Dumanig, F. P., & David, M. K. (2017). The English-medium fever in Pakistan: Analyzing policy, perceptions and practices through additive bi/multilingual education lens. *International Journal of Bilingual Education and Bilingualism*, *20*(6), 736-752.

McCarthy, M., McCarten, J., & Sandiford, H. (2005). *Touchstone: Student's book 1*. New York, NY: Cambridge University Press.

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics*, *32*(2), 129-148.

Moon, R. (1998). *Fixed idioms and expressions in English*. Clarendon, London, UK: Oxford University Press.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Clarendon, London, UK: Oxford University Press.

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

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. England, UK: Cambridge University Press.

Olson, D. R. (1989). On the language and authority of textbooks. In S. de Castell, A. Luke, & C. Luke (Eds.), *Language, authority, and criticism: readings on the school textbook* (pp. 233-244). London, UK: The Falmer Press.

Rundell, M. (2007). *Macmillan English dictionary: for advanced learners* (2nd ed.). Oxford, London, UK: Macmillan Education.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Eds.), *Formulaic sequences: Acquisition, processing, and use* (pp. 1–22) Amsterdam, Netherland: John Benjamins.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487-512.

Siyanova-Chanturia, A., & Martinez, R. (2014). The idiom principle revisited. *Applied Linguistics*, *36*(5), 549-569.

Stubbs, M., & Barth, I. (2003). Using recurrent phrases as text-type discriminators: A quantitative method and some findings. *Functions of Language*, *10*(1), 61–104.

Swales, J. M., & Feak, C. B. (2004). *Academic writing for graduate students: Essential tasks and skills*. Ann Arbor, MI: University of Michigan Press.

To, V. T., & Mahboob, A. (2018). Complexity of English textbook language: A systemic functional analysis. *Linguistics and the Human Sciences*, *13*(3), 264-293. DOI: 10.1558/lhs.31905

Tomankova, V. (2016). Lexical bundles in legal texts corpora–selection, classification and pedagogical implications. *Discourse and Interaction*, *9*(2), 75.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, UK: Cambridge University Press.

Yousaf, M & Shehzad, W. (2018) Prevalence of prefabricated structures in academic discourse: A corpus-based study. *International Journal of English Linguistics*, *8*(5), 297-306.

Zahra, T., & Abbas, A. (2018). Pedagogical Implications of Corpus-Based Approaches to ELT in Pakistan. *Journal of Education and Educational Development*, *5*(2), 259-275.

APPENDIX

LIST OF LEXICAL BUNDLES WITH DISCOURSE FUNCTIONS

Rank	Frequency	Lexical Bundle	Discourse Function
1	73	is equal to the	quantity specification
2	56	is defined as the	topic elaboration/clarification
3	51	is said to be	topic elaboration/clarification
4	49	in the form of	tangible framing attributes
5	45	is directly proportional to	quantity specification
6	42	on the other hand	topic elaboration/clarification
7	39	directly proportional to the	quantity specification
8	39	which of the following	identification/focus
9	37	as shown in the	text reference
10	37	one of the following	identification/focus
11	37	shown in the fig	text reference
12	35	what is meant by	topic elaboration/clarification
13	35	which one of the	identification/focus
14	30	is due to the	cause and effect
15	29	can be used to	modality stance/ability
16	28	as a result of	cause and effect
17	24	a large number of	quantity specification
18	20	in such a way	intangible framing attribute
19	20	the nature of the	intangible framing attribute
20	19	at the same time	time reference
21	19	is based on the	intangible framing attribute
22	18	can be represented by	modality stance/ability
23	18	is known as the	topic elaboration/clarification
24	18	is the amount of	quantity specification
25	18	is the number of	quantity specification
26	17	in the absence of	intangible framing attribute
27	17	such a way that	intangible framing attribute
28	16	are said to be	topic elaboration/clarification
29	16	is one of the	identification/focus
30	16	on the surface of	place reference
31	15	this is known as	topic elaboration/clarification
32	14	at the end of	text/place/time reference
33	14	is less than the	quantity specification
34	13	in the case of	intangible framing attribute
35	13	is represented by the	text reference
36	13	is the same as	comparison/comparative
37	13	on the nature of	intangible framing attribute
38	13	phenomenon is known as	topic elaboration/clarification
39	13	the difference between the	comparison/comparative
40	13	the position of the	intangible framing attribute
41	13	the size of the	tangible framing attribute
42	13	this phenomenon is called	topic elaboration/clarification
43	12	is passed through a	procedure
44	12	is the difference between	comparison/comparative
45	12	the surface of the	place reference
46	11	can be used for	modality stance/ability
47	11	it is due to	cause and effect
48	10	as a result the	cause and effect
49	10	is a measure of	quantity specification
50	10	is meant by the	topic elaboration/clarification
51	10	same as that of	comparison/comparative
52	10	the intensity of the	intangible framing attribute
53	10	the other hand the	topic elaboration/clarification
54	10	the study of the	topic introduction/focus
55	10	with respect to the	topic introduction/focus
56	9	in other words the	topic elaboration/clarification
57	9	in this case the	intangible framing attribute
58	9	is an example of	topic elaboration/clarification
59	9	is related to the	comparison/comparative
60	9	is used as a	procedure
61	9	some of these are	identification/focus

62	9	the same as that	comparison/comparative
63	9	this is because the	cause and effect
64	9	this process is called	topic elaboration/clarification
65	8	and is known as	topic elaboration/clarification
66	8	are given in table	text reference
67	8	as compared to the	comparison/comparative
68	8	be defined as the	topic elaboration/clarification
69	8	can also be used	modality stance/ability
70	8	explain the significance of	topic elaboration/clarification
71	8	in this chapter we	topic introduction/focus
72	8	is determined by the	procedure
73	8	the other hand if	topic elaboration/clarification
74	7	as that of the	comparison/comparative
75	7	can be calculated by	modality stance/ability
76	7	can be represented as	modality stance/ability
77	7	explain what is meant	topic elaboration/clarification
78	7	from left to right	description
79	7	greater will be the	cause and effect
80	7	in terms of the	intangible framing attributes
81	7	is considered to be	topic elaboration/clarification
82	7	the same time the	time reference
83	7	what is the difference	questioning
84	6	due to the fact	cause and effect
85	6	is added to the	procedure
86	6	the greater is the	cause and effect
87	6	the greater will be	cause and effect
88	6	to the fact that	intangible framing attributes
89	6	used to determine the	procedure
90	5	an increase in the	quantity specification
91	5	are due to the	cause and effect
92	5	in contact with the	intangible framing attributes
93	5	is in accordance with	comparison/comparative
94	5	is used in the	impersonal epistemic stance
95	5	referred to as the	comparison/comparative
96	5	that there is a	identification/focus
97	5	which has the same	comparison/comparative
98	5	with the increase of	intangible framing attributes
99	4	be used as a	comparison/comparative
100	4	it is called a	topic elaboration/clarification
101	4	the bottom of the	place reference
102	4	the positions of the	tangible framing attributes

ABOUT THE AUTHORS

Gulfam Hussain obtained his M.Phil. Linguistics from Air University, Islamabad. His research interests are ELT, EAP, corpus-linguistics, applied-linguistics, genre-analysis, and discourse-analysis.

Tehseen Zahra is an Assistant Professor at Department of English, Air University, Islamabad. Her research interests are corpus linguistics, discourse studies and pragmatics. She is the Managing Editor of Corprum: Journal of Corpus Linguistics. She is Co-PI of a research project named PakGenText funded by Higher Education Commission, Pakistan.

Akhtar Abbas is an Assistant Professor at Department of English, Air University, Islamabad. His research interests are academic discourse, corpus linguistics and metadiscourse. He is Co-PI of *Pakistan National Corpus of English*, a research project funded by Higher Education Commission, Pakistan.