An Analysis of IELTS Academic Reading Texts through A Corpus-based Approach

Lao Xin Qin
Chiang Rai Rajabhat University, Thailand
Corresponding author's email: holajoper@gmail.com
Received September 22, 2022; revised December 11, 2022;
Accepted January 06, 2023; online June 15, 2023

Abstract

The IELTS academic reading test is an important gatekeeping device for second language learners. However, few studies have compared the language used in the test's reading passages and their accompanying questions. Accordingly, this study explores 12 IELTS reading passages and accompanying questions from a corpus-based perspective. Specifically, I focus on eight types of four-word lexical bundles (LBs) and five sentence patterns (S-patterns) in a corpus of 444 sentences. Results showed that doing/seeing and being S-patterns accounted for 170 and 130 sentences, respectively, where other noun phrases (NPs), other prepositional phrases, NP + of, and be + NP/ Adjectivial phrases were the most prevalent LBs. To improve candidates' use of reading strategies, I highlight the significance of the most frequent four-word LBs found in the texts. Moreover, by drawing on existing work that looks at how student writers revise academic texts, I show how an IELTS exam writer can design questions that are perhaps clearer and more tightly tied to the language of an accompanying reading. I close by further clarifying how an analysis of S-patterns and LBs can be used to guide both the reception and production of answer-question combinations in academic reading tasks.

Keywords: exam question writing, sentence patterns (S-patterns), lexical bundles (LBs), AntConc

The IELTS academic reading is a test designed to assess applicants' reading ability and academic background in preparation for studying abroad. It is an assessment of receptive skills in which readers receive information from an author. The test involves extracting meanings and/or recreating the author's intention (Carrell & Grabe, 2002). The reading passages and their accompanying tasks are chosen from various sources, such as books, magazines, journals, and newspapers (Moore et al., 2012). The tasks use two types of operations to test candidates: The first type is an adventure reading operation, which involves skimming and scanning; the second one is a careful reading operation, which requires higher-level thinking skills, such as distinguishing facts from opinions, identifying the author's attitude, understanding inferred information, and inferring the meaning of unknown words from co-text/context (Alshammari, 2016).

According to Reppen (2010, p. 4), corpus-based analysis can offer "a readily available source of spontaneous or authoritative texts for language study". When specific texts are analyzed linguistically using corpora, the outcomes are typically regarded as "representative of a given language" (Francis, 1982, p. 7). Accordingly, corpus-based analyses has been used to inform the teaching of vocabulary and grammar, and has had various beneficial effects, such as increasing students' comprehension of vocabulary usage and lowering grammatical and collocational errors (Aston, 2001; Hunston, 2002; O' Keeffe & Farr, 2003; Sinclair, 2004). Studies have also provided teachers with some helpful language teaching strategies (Gabrielatos, 2005; O' Keeffe et al., 2007; Pérez-Paredes et al., 2018; Reppen, 2010; Smirnova, 2017). Few studies, though, have explored sentence patterns (S-patterns) and lexical bundles (LBs) in both IELTS academic reading passages and their accompanying questions.

The majority of studies and teaching methods used for IELTS academic reading concentrate on the analysis of vocabulary, LBs, and grammar, rather than S-patterns. To create a variety of sentences with fixed tenses and structures, S-patterns incorporate grammar, vocabulary, and LBs (Verspoor & Sauter, 2000, pp.16-32). However, despite some studies identifying sentence variety as a crucial element of advanced language use (Coombe et al., 2007; Sullivan & Eggleston,

2006), these studies do not show students the full range of sentence diversity that is possible, such as the process of modifying the original sentence, removing one word, or replacing it with another (Stevenson et al., 2006). Therefore, the current study can further such work by analyzing the diversity of sentences found in IELTS academic reading texts.

The current study examines sentence diversity in IELTS academic reading texts by focusing on five typical S-patterns (Verspoor & Sauter, 2000) and four-word LBs (Hyland, 2008a). A corpus-based analysis was performed on 12 reading texts from the *IELTS Academic 15* book. The study focused on the most common types of S-patterns and LBs, as well as the diversity of the nominal and verbal group (NGrp and VGrp, respectively). The research questions are as follows:

1. What are the most frequent S-patterns found in a range of IELTS academic reading texts?

2. What are the most frequent types of LBs found in a range of IELTS academic reading texts?

3. How can IELTS exam writers formulate clearer, more cohesive task questions?

Literature Review

Learner Challenges with the IELTS Academic Reading Text

Articles from various fields of expertise are used in the IELTS academic reading section to evaluate candidates' reading skills. However, when reading passages are on unfamiliar topics, students may exhibit difficulty in extracting meanings from texts. Moreover, according to McKoon and Ratcliff (2018) and Yang et al. (2005), English as a foreign language (EFL) learners have less reading experience in English than first language English speakers and, as a result, they are less adept at reading-related skills such as word retrieval and syntactic analysis. They concluded that poor reading comprehension is a result of these two factors among EFL students.

Another factor that affects the reading proficiency of EFL candidates is the IELTS academic reading task requirement. Indeed, less proficient EFL students are less likely than native English speakers

to adopt an appropriate reading strategy to help them adapt to the demands of a reading task (Cerdán et al., 2011, 2019; De Milliano et al., 2016). For instance, Cerdán et al. (2011) showed that when faced with unfamiliar subject matter, EFL readers are more likely to select information in the text that matches the information given in the task's instructions or questions. Such a strategy does not explore the information needs of the task, which would include finding multiple keywords, attempting to integrate widely dispersed details, and concentrating on the expressed intent, information, or answer form of the text (Gehrer et al., 2013).

Insufficient analysis combined with a lack of vocabulary and syntactic mastery are other key factors that contribute to low reading proficiency among EFL learners. The most significant obstacles to EFL learners, according to researchers Lems et al. (2017), are the syntactic and lexical-level recognition abilities required for decoding. In addition to the fact that these abilities are language-specific, and must be learnt separately in each language, decoding English words is also incredibly difficult. Grabe (2009), for instance, notes how first language English speakers are more likely than EFL readers to have attained tacit language proficiency in terms of syntax, vocabulary, and LBs before the age of six.

S-patterns with Five Typical Structures

The verb in the predicate determines the S-pattern in English syntax. The subject, predicate, object, complement, and adverbial words must be arranged in a recognizable S-pattern to be understood. To understand better the typical characteristics of S-patterns, Verspoor and Sauter (2000) proposed the following five S-pattern categories: running, being, doing/seeing, giving/buying, and making/considering.

The *running* pattern (realized via intransitive verbs) is a sentence made up of a subject and a predicate that are typically followed by one or more adverbs. This pattern uses an intransitive verb to indicate the action of the main participant. The sentence structure is S (subject) + P (predicate) + A (adverb). Typical verbs that realize a running pattern are swim, speak, ride, listen, where the adverb further specifies some quality of the action.

The second pattern, *being*, uses a copula verb to show that the object provides details about the subject and not the predicate. The sentence structure is S + P + SA (subject attribute) + A, and some examples of verbs are appear, grow, seem, look, be, make, smell, sound, become, prove, taste, feel, remain, and turn.

The third pattern, which is *doing or seeing*, is a statement that describes the shape and purpose of an action or mental experience using a monotransitive verb. The sentence structure is S + P + DO (direct object) + (A), and some example verbs are hold, count, build, and kick, or see, feel, hear, believe, and think.

The fourth pattern, *giving or buying*, is a ditransitive verb that describes a process involving one to three participants. This pattern of sentences is made up of a subject, a predicate, an indirect or benefactive object, and a direct object. The sentence structure is S + P + IO (indirect object)/ BO (benefactive object) + DO + A. Some example verbs include give, pass, send, tell, make, buy, and offer.

The last pattern is *making or considering*, which is realized by complex transitive verbs of making and thinking. This pattern includes a subject, predicate, direct object, and its attributes, wherein the direct object refers to a particular concept. The sentence structure is S + P + DO + OA (object attribute) + A, and some example verbs are make, wipe, drive, call, crown, name, and elect, or consider, assume, prove, declare, certify, regard, and deem.

In S-patterns, the verb predicate licenses how the other sentence components interact, but these sentence components can also organize the structure of LBs. To show any elements or characteristics in this pattern that may affect the meaning or genre of the sentence, S-patterns are presumptively based on real-world discourse cases (Kuzar, 2012). Most of the time, the element or feature that influences a sentence's meaning can be viewed as an object, such as LBs that lack perceptual prominence, recognize a partial structural unit, or do not conform to idiomatic usage in terms of meaning (Biber & Barbieri, 2007). As a result, incomplete phrases and clauses can be supported by LBs in S-patterns. Due to LBs structural inadequacy and their oft lack of idiomatic clarity, they usually fall outside the purview of vocabulary in traditional EFL classrooms. Despite this, it is an essential component

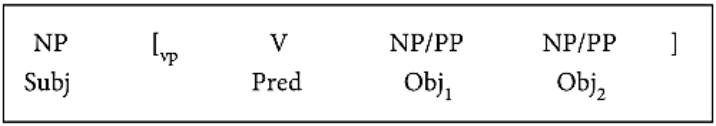
of sentence diversity and plays a significant role in S-pattern usage. Merging LBs with a typical S-pattern can demonstrate the variety of a sentence with multiple meanings.

Although S-patterns are initiated by a predicate verb, LBs can provide diversity in the S-pattern's subject, predicate, object, and adverbial components. LBs are considered "clusters" or "chunks" in sentences (Hyland, 2008a, 2008b), which provide extended collocations to typical S-patterns (Verspoor & Sauter, 2000). They assist in the shaping of different meanings in S-patterns, and help learners in discovering different communicative functions expressed in sentences. Liao (2019) for example, investigated the impact of vocabulary and task demands on IELTS reading comprehension and discovered that the scope of content and task questions in the text were comparable in terms of sentence structure, vocabulary, and LB coverage.

A predicate can also be a type of LB in an S-pattern that can be composed of more than one word unit. According to Verspoor and Sauter (2000), S-patterns determine the order of sentence components and inform readers about the agent (e.g., subject), and any direct or indirect objects. As a result, predicate verbs typically require one or two words to realize a type of S-pattern, while LBs typically require at least two to four words in the form of "subject + predicate" or "predicate + preposition phrase". Thus, according to Kuzar (2012), all possible structures of S-patterns are summarizable from SV, SVO, and SVOO patterns.

As shown in Figure 1, S (subject) and O (object) can be replaced by NP (noun phrase) and PP (prepositional phrase), resulting in NP^V (SV), NP^V^NP/AdjP (SVO), and NP^V^NP^PP (SVOO). Hyland (2008a), on the other hand, used the LBs of Passive + PP, Anticipatory it + verb, and be + NP/AP to replace the forms of SV and VO in the V (predicate) of S-patterns. This directly demonstrates that S-patterns and LBs have the same relationship in terms of usage and predicate function.

Figure 1
The V S-pattern in English (adapted from Kuzar, 2012, p. 17)



Similarly, there is a complementary relationship between S-patterns and LBs in terms of subject, object, attribute, and adverbial. Kuzar (2012) explains that S-patterns arrange the subject, predicate, object, attribute, and adverbial components according to category affiliation and function. In S-patterns, the number and identity of these components correspond primarily to predicates and their parameters. To understand better these structural and functional characteristics, researchers have proposed LBs as components of S-patterns (Biber & Barbieri, 2007; Hyland, 2008a). The four main categories of these LBs are verb phrases, noun phrases, prepositional phrases, and others.

The following section will go into greater detail about the structures and functions of LBs.

LBs in Academic Reading Texts

LBs have been used in the analysis of corpus-driven results, where researchers have focused on repeated sequences of consecutive words using examples of rules (Biber et al., 1999). According to Biber and Barbieri (2007), LBs are multi-word lexical units that frequently appear in texts. LBs also frequently appear as incomplete in terms of meaning because they are the result of frequency-based inquiries. In essence, some identified LBs are not fully functional because they are fragments of much larger grammatical constructions (Biber, 2009).

It is explained in LB studies that IELTS academic reading test results remain unsatisfactory even after candidates focus on acquiring a large vocabulary. As an example, Chen and Liu (2020) found that IELTS reading scores correlated positively with vocabulary breadth and depth. The results of a multiple regression analysis showed that once words reach a certain level (6000 words in this study), the comprehension

of IELTS academic questions is affected. This is because candidates fail to consider that LBs are much more substantial in grammatical and contextual fragments than a single word.

Many academic articles are largely composed of eight structural types and three genre functions of LBs, which have instructional value when comparing texts from various disciplines (Hyland, 2008a, 2008b). The primary focus of such LB research has been on the features associated with these structures and functions, and how their examination can reveal disparities in phraseology. Hyland (2008a, 2008b) also contended that LBs have many potential benefits for teaching language forms, but that research assumptions about the universality or frequency of academic bundles should be carefully examined. This is because LBs in academic texts from various disciplines represent unique situationally specific language features in their respective fields. These LBs manifest in various forms, and their specific patterns must be accumulated and separated in analysis. For example, Hyland (2008a) analyzed academic papers from four disciplines and discovered the following eight structural types of LBs: (1) *Noun phrase (NP) + of* (the end of the, the nature of the, the beginning of the, a large number of); (2) *other noun phrases (NP)* (e.g., the fact that the, one of the most, the extent to which); (3) *prepositional phrase (PP) + of* (e.g., at the end of, as a result of, on the basis of, in the context of); (4) *other prepositional phrases (PP)* (e.g., on the other hand, at the same time, in the present study, with respect to the); (5) *passive + prep phrase (PP)* fragment (e.g., is shown in figure, is based on the, is defined as the, can be found in); (6) *anticipatory it + verb/adj* (e.g., it is important to, it is possible that, it was found that, it should be noted); (7) *be + noun/adjectival phrase (NP/AP)* (e.g., is the same as, is a matter of, is due to the, be the result of); and (8) *others* (e.g., as shown in figure, should be noted that, is likely to be, as well as the).

Hyland also divided these LBs into three orienting functions, the first of which is the research-oriented function. This function includes LBs of subcategories like location (e.g., at the beginning of, at the same time), procedure (e.g., the use of the, the role of the), quantification (e.g., the magnitude of the, a wide range of), description (e.g., the structure of the, the size of the), and topic (e.g., in the Hong Kong).

These bundles support authors in developing research activities and actual experiences, expressing scientific ideologies, and emphasizing experiences rather than explanations, all of which are advantageous to "strong" scientific claims (Hyland, 2008a). The second function, text-oriented, deals with how text is organized and how important it is as information or as a point of contention. This function is realized via transition signals (on the other hand, in addition to the), resultative signals (as a result of, it was found that), structuring signals (in the next section, as shown in figure), and framing signal (in the case of, with respect to the). Hyland (2008a, 2008b) claims that these LBs often present more detailed content, provide proof, connects ideas, guide readers to read the key information of the text, and contain the meanings of any limitations. The third participant-oriented function has two subcategories: standpoint feature (are likely to be, may be due to) and stickiness feature (it should be noted that, as can be seen). These concentrate on the attitude and viewpoint that the author wishes to convey or to directly transmit critical information to the readers, respectively (Hyland, 2005).

Reading Strategies used by IELTS Candidates

Candidates will employ various reading strategies in the IELTS academic reading test to achieve their desired score. Holi et al. (2020) found that candidates used a series of adaptive reading skills to respond to the reading test questions. Such reading skills include assuming and emphasizing but also underscoring main concepts, ignoring, forecasting, using verbal cues, stressing, and using non-textual cues. In studies such as these, effective reading comprehension is correlated with critical reading abilities and proper strategy choice (Marzban & Barati, 2016). The connection between metacognition and cognitive reading strategies in reading tests was also defined by Ahmadian et al. (2016) as follows:

1. Cognitive reading strategies: underscoring and emphasizing, stopping and thinking, skimming, forecasting, paraphrasing, taking notes, deductive reasoning, re-reading, and so forth.

2. Metacognitive reading strategies: self-monitoring, strategizing, self-questioning, self-evaluation, and concentrating.

Reading comprehension is a goal for most language learners, irrespective of their level of language proficiency. The fundamental skills needed for reading comprehension include developing effective, appropriate, and reliable comprehension strategies (Nation & Norbury, 2005; Zare, 2012). Moreover, to understand the meanings of a text, effective readers employ various strategies, switching between them as they read (Cohen, 1998; Phakiti, 2003; Rogers & Harley, 1999).

Top-down and bottom-up reading comprehension strategies are described by scholars as follows (Cohen, 1998; Phakiti, 2003; Rogers & Harley, 1999): (a) bottom-down, which assists students in improving their ability to read sentences line by line to fully understand a paragraph; (b) top-down, which assists in the comprehension of larger pieces of information, such as an essay. These strategies are said to assist students and readers in understanding how a sentence, or even a set of phrases, fits into a larger meaningful whole.

Test-taking strategies, on the other hand, are "language use strategies but also test wiseness strategies" (Cohen, 1998, p. 219). Language usage strategies are deliberate actions students take to enhance their use of a foreign or second language to complete testing tasks. Researchers who looked at reading assessments given to students learning English as a second language found that test-takers use a range of strategies to deal with reading tasks and to find solutions (Bachman & Palmer, 1996; Cohen & Upton, 2007; Nation & Norbury, 2005). Rogers and Harley (1999) outlined the following strategies for taking tests to distinguish further between them: carefully reading question types and answering directives, scheduling your time for problem-solving, using keywords in questions, postponing answers to complex and challenging questions, and re-reading and evaluating tasks to ensure correct answers.

Corpus-Based Studies on Academic Texts

According to some scholars, a corpus refers to a high-level representation of language that is comprised of a collection of spoken or written texts used for language assessment, or a representative and balanced sample of (written or spoken) texts produced in natural communication environments (Gries, 2016; McEnery & Hardie, 2011;

Weisser, 2016). Many corpora are meant to be representative and balanced in terms of a language variable or an academic or literary genre, and they are subsequently used for corpus-based research.

Academic text analysis is commonly used with a specialized corpus, wherein analysis of such corpora has revealed that the expressive power of language is restricted by the level of terms or vocabulary contained in it (Biber & Barbieri, 2007). By conducting research on specific fields and disciplines in corpora, researchers are directed to focus on the reference and feasibility of data in this field (Simsek & Gün, 2021). In this way, corpus technology can be used to explore aspects of language description and solve numerous linguistic or lexicological problems (see Aston, 2001).

The positive output of linguistic data in the study of academic writing processed by corpus tools is very useful in the field of linguistics. Such valuable studies are primarily found in lexicology, for example, wherein corpus-based comparisons of the use of hedges and boosters in academic writing are made (e.g. Akbas & Hardman, 2018; Hyland & Milton, 1997; Taymaz, 2021); Corpus-based LBs in research articles (Gezegin-Bal, 2019; Romer & Arbor, 2009); authorial self-expression in research articles written by corpus-based study (Hryniuk, 2018); and LBs in master's and doctoral dissertations based on corpus research (Hyland, 2008a, 2008b). There is no doubt that these studies focus on corpus analysis of vocabulary in academic texts, but the distinction is in the various objects investigated, such as the class of words, the structure of words, and the function of words.

In some studies, researchers have paid closer attention to the expression of LBs in a variety of linguistic domains (Hyland, 2008a, 2008b). LBs were first proposed by Biber et al. (1999), and defined as "recurrent expressions" regardless of their idiomatic or structural states in the texts (p. 990). Following this, structural types and functions of LBs have been proposed and classified into three categories: stance expression, discourse organization, and reference expression (Biber et al., 2004; Cortes, 2002, 2004). These researchers and others believe that LBs serve multiple functions in terms of discourse organization and structure, such as expressing, clarifying, and expounding the theme. Hyland (2008a), for instance, categorized LBs into eight structural types

of four-word bundles for corpus analysis of master's and doctoral theses from four disciplines. He proposed three orienting functions that are more frequently found in these types of academic texts.

Many researchers have also recognized and referenced research-oriented, text-oriented, and participant-oriented LBs to other academic aspects of the research field (e.g., Ang & Tan, 2018; Panthong & Poonpon, 2020; Wei & Lei, 2011). The present study examines IELTS academic reading texts from the perspective of Hyland's (2008a) eight structures and three orienting functions to investigate their significance in sentence diversity.

Method

Research Setting

Every year, Cambridge University Press (CUP) releases a new IELTS academic book that reflects the most recent IELTS task types and subject trends, ensuring that candidates are up to date on the most recent IELTS tests. I selected the academic reading texts in this series of test books for corpus-based analysis in this study. The sentence diversity mentioned by Coombe et al. (2007) and Sullivan and Eggleston (2006) is the focus of this study, wherein I explore S-patterns and LBs in IELTS academic reading texts.

Research Sample

The sample was composed of 12 reading texts from four IELTS academic reading students' drilling tests. These texts, which offer candidates actual simulation tests and answers, are taken from the Cambridge English IELTS 15 academic book (Cambridge English, 2020).

The list of subjects and disciplines from the 12 readings is categorized in Table 1. Four tests, each with three academic passages and 40 questions, make up the real IELTS 15 Academic book examination papers. Two to three different tasks and 13 to 14 quizzes are included for each academic passage. Each reading contained between 776 and 956 words, while word tokens and word types total 10,594 and 2,775 words, respectively.

Table 1
List of disciplines and subjects of 12 IELTS academic reading texts

Discipline	Title of subject *(word count)*
Biology	Nutmeg - a valuable spice *(868 words)*
Engineering	Driverless cars *(790 words)*
History	What is exploration? *(929 words)*
Engineering	Could urban engineers learn from dance? *(927 words)*
Biology	Should we try to bring extinct species back to life? *(956 words)*
Psychology	Having a laugh *(950 words)*
History	Henry Moore (1898-1986) *(893 words)*
Physics	The Desolenator: producing clean water *(823 words)*
Linguistics	Why fairy tales are really scary tales *(965 words)*
Biology	The return of the huarango *(913 words)*
Linguistics	Silbo Gomero - the whistle 'language' of the Canary Islands *(776 words)*
Business Studies	Environmental practices of big businesses *(804 words)*

Research Procedures

Because AntConc could not automatically retrieve the number of S-patterns, the first S-pattern theory's text mining had some noise and retrieval issues, necessitating manual computation and differentiation of the five most common S-patterns and the number of sentences. The three steps in the text analysis process are as follows:

1. I used AntConc's concordance and file view to analyze the predicate verb attribute in a sentence, and identified the five typical S-patterns in the sentences of the IELTS academic reading texts. This allowed me to track how S-patterns are distributed throughout the article.

2. I used the N-grams functions in the AntConc tool to examine the proportion of the eight structural types of four-word bundles in the IELTS academic readings.

3. I extracted the S-patterns and four-word bundles from the original sentences of the IELTS academic reading texts. This allowed me to observe sentence diversity, and combine the question sentence analysis of reading tasks to give readers or candidates an understanding of the question through the exam writer's transformational activity.

Data Analysis

In this study, AntConc was used to code and count the structural types of S-patterns and LBs in IELTS academic reading texts. Then, the S-pattern framework was applied to explain the functions of S-patterns and LBs in task questions and original texts.

Results

Analysis of S-patterns in IELTS Academic Reading Texts

The predicate is inextricably linked to the main structure of an S-pattern. According to Kolln and Funk (2011), S-patterns include all English sentences, from the simplest to the most complex, in addition to summarizing all grammatical functions. S-patterns differ from LBs in that they place more emphasis on the predicate and the attributes that go along with it. Due to this, the analysis of S-patterns necessitates a retrieval analysis of verb attributes of predicates. This is done through the functions of Concordance and File View to distinguish five typical S-patterns.

Utilizing the S-patterns identified by Verspoor and Sauter (2000), I analyzed a total number of 444 sentences in the 12 readings for the IELTS test. The investigation found that the running pattern was represented 0 to 5 times per reading and 25 times (5.63%) across all 12 readings. The being pattern, meanwhile, was represented 6 to 16 times per article or 130 times (29.28%) overall; the doing/seeing pattern was represented 8 to 22 times per article or 170 times (38.29%) overall; the giving/buying pattern was represented 0 to 6 times per article or 35 times (7.88%) overall, and the making/considering pattern was represented 2 to 14 times per article or 84 times (18.92%) overall.

The results show that the second and third S-patterns are more frequent in these IELTS academic reading, making up more than 60% of all patterns. However, the running pattern is either never used or used only occasionally in each the IELTS academic reading texts.

Structural Types of 4-word LBs in IELTS Reading Texts

This study analyzed the proportions of eight structural LBs types in IELTS academic reading texts using AntConc's N-grams function. I used Hyland's (2008a, 2008b) eight structural types of 4-word bundles, because research has shown that 4-word bundles are more prevalent than 5-word bundles, and offer a clearer grammatical structure and function than 3-word bundles.

The essence of LBs, according to Hyland (2008a), comprises the frequency of occurrence and the breadth of extended collocations, but the amount of actual frequency truncation is arbitrary. A comparison with the data corpus of Hyland (2008a, 2008b) or Biber et al. (1999) is impossible because of the relatively small corpus (10,594 words with 2,775 word types) in the current study. Hence, I take a conservative approach in my study to assist EFL learners in developing and acquiring appropriate and subject-sensitive LBs.

The findings show a noticeable trend in the distribution of LBs in the sampled IELTS academic reading texts. AntConc's N-Grams analysis of 4-word bundles revealed the following results: LB structural types consisting of NP + of, Other NP, Other PP, passive + PP, and be + NP/AP appeared 182 times (12.48%), 432 times (29.63%), 335 times (22.98%), 108 times (7.41%), and 169 times (11.59%), respectively, making up 84.08% of all LB structural types in the sample. This suggests that the distribution of these five LB structural types in the sampled IELTS academic readings is fairly broad. Among them, NP + of, Other NP, and Other PP 4-word structures appear in the greatest proportion, confirming Hyland's (2008a, 2008b) findings that most clusters in academic writing are components of nouns or prepositional phrases. In contrast, PP + of, Anticipatory it, and Others are the three structural types in LBs in IELTS reading that are used the least, accounting for 73 times (5.01%), 71 times (4.87%), and 88 times (6.04%), respectively, for a relatively low combined ratio of 15.92%.

Utilizing S-Patterns to Guide Reading Strategies

The third goal of this paper was to evaluate the efficacy of a corpus-based S-pattern framework as a guide to answering questions. For a more detailed explanation of this step, I show how IELTS writers

might extract task questions from the original text using relevant S-patterns and LBs, and then, by recognizing the structural types and functions, students can answer the questions. As a result, the examples will represent pragmatic functions from two perspectives: transformational activity for task questions (from the perspective of the IELTS exam writer), and the S-pattern framework for the original text (from the perspective of the reader).

This third step, which aims to provide guidance for answering questions, used examples of seven task types from the IELTS corpus (Moore et al., 2012). These task types are as follows: (1) True/False/ Not given, (2) section-summary match, (3) gapped summary, (4) short answer, (5) information-category match, (6) multiple choice, and (7) other (e.g., sentence completion, information transfer, etc.).

Candidates can find relevant content in the reading using the keywords provided in the task, such as "In the Middle Ages," "most Europeans," and "nutmeg" in the task question shown in Example (1).

(1) **Task type:** True/False/Not given
Statement: In the Middle Ages, most Europeans knew where nutmeg was grown.
Original text:
Nutmeg was a highly prized and costly ingredient in European cuisine in the Middle Ages, and was used as a flavouring, medicinal, and preservative agent. Throughout this period, the Arabs were the exclusive importers of the spice to Europe. They sold nutmeg for high prices to merchants based in Venice, but they never revealed the exact location of the source of this extremely valuable commodity.
(adapted from Cambridge English, 2020, p. 16)

The 4-word bundles in this text can be marked as "structuring signals" in Hyland's (2008a) text-oriented function. This LB refers to extended text information in organized discourse and directs readers to the following content. As a result, the corresponding sentence appears in the second sentence in the original text of Example 1 after "in the Middle Ages".

This means that readers must read three sentences in the original text to determine whether the question is correct or incorrect based on the information provided by the question.

As shown below, the answer is in italics and bold of Table 2, which is "False" because the original text only mentioned "the Arabs", not "most Europeans". The exam writer transformed "they" into "most Europeans," "never revealed" into "knew," "the exact location of" and "the source of this" into "where", and "extremely valuable commodity" into "nutmeg".

Table 2

Exam writer's transformational activity in composing the task question

Original Sentence	They sold nutmeg *for high prices to* merchants based in Venice, but	they	***never revealed***	*the exact location of*	*the source of this* extremely valuable commodity.
Task statement	*In the Middle Ages,*	most Europeans	***knew***	where	nutmeg was grown.

Because this study is primarily about the S-pattern framework's potential for answering question guidance, Table 3 is the key object to explain. The S-pattern framework distinguishes between S-patterns and LBs by using bold, italic and underline, as well as the method of marking types and functions. Thus, the predicate is marked by a "giving/buying pattern," and the LBs of "for high prices to" is marked by the research-oriented function's "quantification", which refers to the quantity/value of something. Furthermore, the LBs "the exact location of" and "the source of this", are the "location" of the "research-oriented" subcategory to locate something in the text. These LBs can remind the reader to pay attention to the location information in reading.

Table 3
S-pattern framework in original text from the reader's perspective

S-pattern framework	S (subject) + **sold** + N (noun) + ***for high prices to*** + SB (somebody) + based in + P (place), but + S (subject) + **never revealed** + ***the exact location of*** + ***the source of this*** + adv. + adj. + noun.
Mark	**The giving/buying pattern; Quantification (LBs: Other PP); Location (LBs: NP + of)**

To answer the question in Example (2), candidates must read a large amount of information, which makes for a very difficult task. The keyword provided in the task question is "appropriate vehicle". However, because there is little key information and the word "vehicle" appears in almost every section, finding the corresponding sentence in the original text is extremely difficult. Hence, the LBs function provides candidates with a better perspective to consider the question; namely, selecting "the opportunity of choosing" in italics in Table 4's task question, which belongs to the "procedure" of research-oriented function and refers to the process of using and operating.

(2) **Task type:** Section-summary match
Question: Reading Passage 2 has seven sections, **A-G**. Which section contains the following information? Reference to the opportunity of choosing the most appropriate vehicle for each trip
Original text:
E section
Specialised vehicles may then be available for exceptional journeys, such as going on a family camping trip or helping a son or daughter move to university.
(adapted from Cambridge English, 2020, p. 21)

Table 4
Exam writer's transformational activity in composing the task question

Original Sentence	**Specialised vehicles** may then	***<u>be available for exceptional</u>***	**journeys**, such as going on *a family camping* ***trip*** or **helping a son or daughter move to university**.
Task Question	Reference to	***<u>the opportunity of choosing</u>***	the most **appropriate vehicle** for **each trip**

This section of text demonstrates the process of selecting a vehicle, and candidates are guided by the LB's function to read the section on selecting a vehicle. "Be available for exceptional" is marked as a "being pattern" and "stance features" in the S-pattern framework of the original text in Table 5. These LBs correspond to the task question's LB, "the opportunity of choosing". The "a family camping trip" is labeled as the "description" of the research-oriented subcategory, which refers to describing "each trip" in the task question. It is worth noting that "each trip" in the task question corresponds to the noun "journeys" in the original text's S-pattern framework, but it also corresponds to the "such as + doing sth or doing sth" in the sentence's latter part to describe the various journeys in detail.

Table 5
S-pattern framework in original text from the reader's perspective

S-pattern framework	S + **may then** ***<u>be available for exceptional</u>*** + **journeys, such as + doing** + prep. + ***<u>a family camping trip</u>*** + **or +doing** + a/an + noun + or + noun + verb + to + place.
Mark	**The being pattern + Stance features (LBs: Be + AP)**; **Description (LBs: Other NP)**

Gapped summary and short answer tasks are combined in Example (3). The main difference is that the former is an incomplete sentence, and the latter is a question, but the task requirements are the same.

(3) **Task type:** Gapped summary; 4. Short answer
Choose **ONE WORD ONLY** from the passage for each answer.
Question: The 2 ..... surrounds the fruit and breaks open when the fruit is ripe
Original text:
The fruit is encased in a fleshy husk. When the fruit is ripe, this husk splits into two halves along a ridge running the length of the fruit.
(adapted from Cambridge English, 2020, p.16)

Candidates who can pull keywords from the question "when the fruit is ripe" are more likely to locate the answer in the original text. In Table 6, the transformational activity was used to rephrase the IELTS exam writer's question. In the task question, the writer uses the word "and" to connect information from two sentences in the original text.

Table 6
Exam writer's transformational activity in composing the task question

Original Sentence	The fruit	***is encased in a***	fleshy **husk**. When the fruit is ripe, this **husk** splits ***into two halves along*** a ridge running ***the length of the*** fruit.
Task Question	The ~~**husk**~~	**surrounds**	the fruit ***and*** **breaks open** when the fruit is ripe

In Table 7 below, "is encased in a" is labelled as a "description" function, which is used to describe a state of something, along with "splits" as a "doing/seeing pattern". It responds to the "surrounds" part of the original question (Table 6), which is used to determine the answer location. Moreover, the LBs of "into two halves along" is marked by the "procedure" of the research-oriented subcategory, which refers to describing the process of something, corresponding to the "breaks open" part of the task question. In a second assessment, these LBs can also be used to determine the answer in the original text. The "description" function marks the last LBs "the length of the". Thus, using the S-pattern framework in the IELTS academic reading test, candidates can quickly

read the overall text, comprehend important information, and locate answers.

Table 7
S-pattern framework in original text from the reader's perspective

S-pattern framework	S + ***is encased in a*** + adj. + noun. When S (the + noun) + **is** + adj, S + **splits** + ***into two halves along*** + a + noun + adj. + ***the length of the*** + noun.
Mark	**The doing/seeing** and **being patterns** + **Description (LBs: Passive + PP fragment**; **NP + of)**; **Procedure (LBs: other PP)**

In Example (4), the information-category match task type of question requires candidates to read an entire paragraph or text to determine the answer, which is difficult because the name of the same person may appear in each paragraph. This would be especially difficult for weaker readers, who would have to read the entire text.

(4) **Task type:** Information-category match
Match each statement with the correct person, **A, B** or **C**.
Question: It is important to concentrate on the causes of an animal's extinction.
List of People
A Ben Novak
B Michael Archer
C Beth Shapiro
Original text:
"Many of the technologies that people have in mind when they think about de-extinction can be used as a form of 'genetic rescue'," explains Shapiro. She prefers to focus the debate on how this emerging technology could be used to fully understand why various species went extinct in the first place, and therefore how we could use it to make genetic modifications which could prevent mass extinctions in the future.

(adapted from Cambridge English, 2020, p. 41)

The question's positioning keyword is "animal extinction", which appears in almost every paragraph. However, readers with weak reading ability may have no place to start. Nevertheless, as shown in

Table 8, by adopting an S-pattern classification and LB function to label the information in the question, this strategy makes it easier for readers to answer the question. For example, in Hyland's (2008a, 2008b) "participant-oriented" function, "It is important to" is a "stance feature" subcategory that conveys the author's perspective. Here, by paying close attention to the sentences in which the author attempts to convey important information, the predicate phrase "concentrate on" is identified by "making/considering pattern". This phrase refers to the relationship between the direct object and the object attribute that means "animal extinction" and "causes". That is, the S-pattern identifies "causes" as the key information in the article as the object attribute. This is also confirmed by the LBs of "the causes of an", which is marked as "resultative signals" of the "text-oriented" function, referring to the reasoning or causal information of "animal's extinction". The original sentences are located after the functional analysis of S-patterns and LBs mentioned in this study.

Table 8

Exam writer's transformational activity in composing the task question

Original Sentence	She prefers	**to focus**	**the debate** on **how** this emerging technology *could be used to* fully understand why various species went extinct *in the first place*, and therefore **how** we could use it to make genetic modifications which could prevent mass extinctions in the future.
Task Question	*It is important to*	**concentrate on**	*the **causes** of an* animal's extinction.

In Table 9, the S-pattern framework responds to the information provided by the question to determine the location of the answer. The "prefers to focus" is marked as a "making/considering pattern", followed by the noun "debate" as the direct attribute corresponding to the question's "an animal's extinction" prompt or keyword. Then, as the object attribute corresponding to the "causes", the "on how + (sentence), and therefore how + (sentence)" serves as a detailed description or

inference of "debate". Furthermore, "could be used to" and "in the first place" are marked as "procedure" and "location" subcategories of the "research-oriented" function.

Table 9

S-pattern framework in original text from the reader's perspective

S-pattern framework	S + **prefers to focus** + the + noun + on + **how** + (sentence) ***could be used to***; ***in the first place*** +, and therefore + **how** + (sentence).
Mark	**The making/considering pattern**; **Procedure (LBs: Passive + PP fragment)**; **Location (LBs: Other PP)**

Example (5) is a task question with multiple choices. The positioning information "in the first paragraph" is given directly in the question, with the keywords "laughter" and "the writer emphasizes".

(5) **Task type:** Multiple choice
Choose the correct letters, **A, B, C,** or **D**.
Question: When referring to laughter in the first paragraph, the writer emphasises
A. Its impact on language.
B. Its function in human culture.
C. Its value to scientific research.
D. Its universality in animal societies.
Original text:
Laughter is universal across all human cultures and even exists in some form in rats, chimps, and bonobos. Like other human emotions and expressions, laughter and humour provide psychological scientists with rich resources for studying human psychology, ranging from the development of language to the neuroscience of social perception.
(adapted from Cambridge English, 2020, p. 44)

In Table 10, "in the first paragraph" is marked as "structuring signals" of the "text-oriented" function, which refers to the extended information in the text that the reader is directed to read and specifies the text limits (Hyland, 2008a, 2008b). According to the task question's perspective, the IELTS exam writer deleted the answer "scientific research" presumably to test the candidates' reading ability. The word

"laughter" is directly positioned in the first sentence of the original text S-pattern framework in Table 10. Hence, readers with poor reading skills may mistakenly believe that the author wishes to emphasize their own perspective, while failing to recognize that "is" belongs to the "being pattern". It refers to the subject-object relationship.

Table 10

Exam writer's transformational activity in composing the task question

Original Sentence	*Like other human emotions* and expressions, laughter and humour **provide psychological scientists** *with rich resources for* studying human psychology, ranging *from the development of* language to *the neuroscience of social* perception.
Task Question	When referring to laughter ***in the first paragraph***, the writer emphasises ~~**C. Its value to scientific research.**~~

In the S-pattern framework analysis, shown in Table 11, the phrase "in some form in" refers to the establishment of additional links between elements and is marked "transition signals" of the "text-oriented" function (Hyland, 2008a, 2008b). In contrast, "laughter" is mentioned again, with "humour" as a subject in the second sentence. Moreover, the predicate "provide" is marked as the "giving/buying pattern", in this case, "laughter and humour" is what the author is trying to emphasize. The phrase "with rich resources for" is tagged "framing signals" of "text-oriented", which refers to a limiting condition. Then, the phrases "from the development of" and "the neuroscience of social" are marked as "procedure" and "description" subcategories of the "research-oriented" function.

Table 11
S-pattern framework in original text from the reader's perspective

S-pattern framework	S is	SA (subject attribute). ***in some form in***
Mark	The being pattern	**Transition signals (LBs: Other PP)**
S-pattern framework	***Like other human emotions*** and + noun (s), S + **provide** + adj. + SB (somebody) + ***with rich resources for*** + doing sth. (something), doing + ***from the development of*** + noun + to + ***the neuroscience of social*** + noun.	
Mark	**Resultative signals (LBs: Other PP)**; **The giving/buying pattern**; **Framing signals (LBs: Other PP)**; **Procedure (LBs: NP + of)**; **Description (LBs: NP + of)**	

The five examples given above are the most commonly encountered in IELTS academic reading tests. In their corpus analysis of IELTS academic reading tests, Moore et al. (2012) classified some task types that appear less frequently in tests as other task types. As a result, the final example is from the table type in the IELTS test.

The keywords provided by the question in Example (6) is "Middle Ages". This keyword can directly assist candidates in determining pertinent information about the reading content.

(6) **Task type:** Other
(e.g. sentence completion, information transfer etc.)
Choose **ONE WORD ONLY** from the passage for each answer.
Question: Middle Ages Nutmeg was brought to Europe by the 8.....
Original text:
Nutmeg was a highly prized and costly ingredient in European cuisine in the Middle Ages, and was used as a flavouring, medicinal, and preservative agent. Throughout this period, the Arabs were the exclusive importers of the spice to Europe.
(adapted from Cambridge English, 2020, p. 16)

The exam writer's transformational activity is shown in Table 12, where the predicate "were" and the LB, "the exclusive importers of", have been transformed into "was brought to Europe". The resultant 4-word LB is a "resultative signal", belonging to the text-oriented function.

The phrase "throughout this period" was then transformed to "Middle Ages" by the IELTS exam writer in the task question.

Table 12

Exam writer's transformational activity in composing the task question

Original Sentence	**Throughout this period,**	**the Arabs**	**were**	***the exclusive importers of***	**the spice** to Europe.
Task Question	**Middle Ages**	**Nutmeg**	***was brought to Europe***		by the **~~8. Arabs~~**.

In Table 13, in the S-pattern framework from the reader's perspective, "were" is marked as a "being pattern", indicating that the subject and object are related in some way. The LBs "the exclusive importers of" is designated as a "description" of research-oriented function, and it refers to the Arabs.

Table 13

S-pattern framework in original text from the reader's perspective

S-pattern framework	Throughout this period,	S	**were**	***the exclusive importers of***	the noun + to + place.
Mark			**The being pattern**	**Description (LBs: NP + of)**	

Discussion

A corpus-based approach was used in this study to create a targeted corpus composed of IELTS academic reading passages and questions. The primary aim was to explore sentence diversity, as well as to analyze the distribution of S-patterns and LBs found in 12 IELTS academic reading texts. The following are responses and discussions to the three research questions raised at the start of this study.

The study's first research question was "What is the most frequent S-patterns found in a range of IELTS academic reading texts?" According to the corpus analysis results, the two most common S-patterns were *doing/seeing* and *being* patterns. They accounted for

170 (38.29%) and 130 (29.28%) sentences, respectively, and together, they account for 67.57% of the five typical S-patterns found in the sampled texts.

Based on the analysis of the 10,594 words in the IELTS academic reading corpus, the exam writers seem more likely to use these two S-patterns when writing about various subjects. To begin with, the most common pattern in the corpus is the *doing/seeing* pattern, which refers to using a single transitive verb to describe an action or sense perception (Verspoor & Sauter, 2000). For example, consider a perception involving two participants: one who acts or experiences and one who is acted upon or perceived. Table 7 showed an S-pattern framework of S + splits + into two halves along + a + noun + adj. + the length of the + noun.

The *being* pattern, on the other hand, was the second most frequent pattern in the corpus. This pattern refers to the use of copula verbs to introduce grammatical objects that provide details about the subject. The primary function of this S-pattern is to indicate the relationship between the subject and an attribute or category of the subject, which is equivalent to the mathematical symbol "=" (Verspoor & Sauter, 2000). Table 11, meanwhile, showed the first S-pattern framework of S + is + SA (subject attribute). With this kind of pattern, the author emphasizes an objective statement regarding the relationship between the subject and the attribute of the subject.

Overall, the corpus-based results showed that the text writing construction of IELTS academic reading tasks tends to use these two patterns. This finding has some value and importance in teaching IELTS academic reading, because teachers may want to tell their students to focus on understanding declarative structures. However, each typical S-pattern investigated in this study will express multiple meanings and a diverse range of communicative functions.

Because there is little literature on S-patterns in IELTS texts, the findings of this study are difficult to compare. However, the analysis of the five S-pattern's occurrence frequencies of Verspoor and Sauter (2000) provides some means of comparison. The most common form of S-patterns encountered in academic books and articles differs from spoken language in that the writers carefully consider the content of the sentence before writing it, and thus written texts lack the blank

and unfinished sentences that may appear in spoken language. Another reason is that longer sentences in more formal written forms are frequently used to demonstrate the writers' creativity (that is, sentence diversity), and thus written sentences are often more complex than spoken sentences (see Verspoor & Sauter, 2000).

The study's second research question was "What are the most frequent types of LBs found in a range of IELTS academic reading texts?" In this study, Hyland's (2008a, 2008b) eight structural types of four-word bundles were used to explore the syntactic frames used in the IELTS academic reading texts. The results show that "other NP", "other PP", "NP + of", and "be + NP/AP" are the most frequently used structural types, appearing 432 times (29.63%), 335 times (22.98%), 182 times (12.48%), and 169 times (11.59%), respectively. These findings differ slightly from those of the four most common structural types of LBs discussed in Hyland (2008a). In my corpus analysis of IELTS academic reading texts, for example, the writers used "other NPs" the most. In contrast, the structural type "noun phrase + of" is the most frequently used in Hyland's (2008a) academic theses corpus. In other words, the use of LBs differs between IELTS academic reading texts and academic theses texts. Furthermore, in Hyland (2008a), the "passive + PP" ranked second in the highest frequency of occurrence in the academic theses corpus of four disciplines, whereas the "passive + PP" ranked fifth in the IELTS academic reading corpus in this study with a relatively small number of occurrence at 108 times (7.41%). By comparing the data of these two corpora, there are noticeable differences in the structural types of LBs used in the same overall language variety (academic English), and these differences seem to be driven by the different academic purposes of each corpora. As a result, the findings of this study address the needs of a specific group of students and researchers, rather than being generalizable to a larger population.

Nevertheless, based on a frequency analysis of the IELTS academic reading corpus, these four structural types of LBs have instructional significance in the teaching of IELTS students and in the training of IELTS teachers. The analysis of "be + NP/AP" is premised on the same theory as that of "being pattern" in the previous

step's data analysis. However, the reason for the higher frequency of the former is that the collected data are sentence fragments, whereas the latter is collected for complete sentences. This emphasizes the significance of the being pattern in IELTS academic reading and provides candidates with these two frequencies of S-pattern and LBs as references, which should be worth noting for its structural collocation forms and functions.

LBs found in IELTS academic reading (based on corpus analysis) are a language-based means of representing IELTS writers' communicative ability. These LBs appear to have a significant impact on academic writing and reading, as they organize texts, persuade readers, and moderate positions (see Bowen & Thomas, 2020). The use of these structural types of four-word bundles varies greatly across disciplines and cultures. It can be seen that LBs are central to discourse construction and textual genre use. The third research question elaborates on their functions by combining S-patterns and four-word bundles to form an S-pattern framework for analysis. As a reminder, the study's third research question was "How can IELTS exam writers formulate clearer, more cohesive task questions?" To provide a more detailed response, the third step of this study's results is interpreted from two perspectives: IELTS academic reading exam writers and readers.

First, the results can be used to help exam writers write clear and consistent task questions. In transformational activities—activities where the exam writer takes sections of the original text and transforms the content for inclusion in questions—this is primarily reflected in S-patterns and LBs. For example, in the transformation of LBs to S-patterns in Table 6, the original four-word bundle, "is encased in a", was transformed by the IELTS exam writer into a *doing/seeing* pattern, "surrounds". The other components in the S-pattern to LBs involved "the exact location of" and "the source of this" in the original text in Table 2, which were transformed by the IELTS exam writer into "where". Moreover in the original text, "the debate" and the two "hows" in Table 8 were transformed by the IELTS exam writer into "the causes of an". It is not difficult to conclude from these six examples that these transformational activities are primarily noun phrases, be + noun/adjective phrases, and S-patterns. Compared to other studies,

these findings are similar to Bowen's (2019) findings regarding the utility of noun groups in academic writing. Thus, a corpus-based approach to academic writing and reading that focuses on noun phrases, prepositional phrases, and be + noun/adjective phrases may be useful in preparing for the IELTS test. This step also explains why the data from the second step shows a higher frequency of structural types in other NP, other PP, NP + of, and be + NP/AP. The examples of transformational activity from the IELTS exam writers, on the other hand, can further explain the use of various four-word bundles in academic writing across disciplines. Some academic writing studies on LBs based on corpus analysis have not demonstrated this in detail (e.g., Ang & Tan, 2018; Hyland, 2008a, 2008b; Panthong & Poonpon, 2020; Wei & Lei, 2011).

Second, this analysis helps us to interpret the functions of S-patterns and LBs in the original text and question in response to the information from the perspective of the readers. In the original sentences, the S-patterns and LBs of Verspoor and Sauter (2000) and Hyland (2008a, 2008b) are used to mark structural types and functions. Hyland (2008a, 2008b) proposed three oriented functions based on academic LBs, explaining how these four LBs help readers establish the connection of reading comprehension. For example, the six examples provided in the study's third step all detail how the task question's LBs and S-patterns locate and judge the corresponding sentence contained in the original text.

Moreover, to make it easier for readers to identify the information in the original sentences, as shown in Table 7, the categories and functional attributes of the LBs "is encased in a", "into two halves along", and "the length of the" are highlighted by *description* (LBs: passive + PP fragment; NP + of) and *procedure* (LBs: other PP). The types of S-patterns realized by "is" and "splits", meanwhile, are *being* and *doing/seeing* patterns, and the remaining sentence components were replaced by single words. The three oriented functions of the LBs make it easier for the readers to understand the overall meaning using an S-pattern framework. Furthermore, the results of this step can be used to show the IELTS reader how to use the functions of S-patterns and LBs to understand the question better. These S-patterns and LBs

are based on corpus analysis and have guidance significance in teaching readings and answering task questions for teachers and test takers.

Conclusion

In conclusion, this study explored the sentence diversity created by the structural types and functions of S-patterns and LBs in 12 IELTS academic reading texts. The results showed that *being* and *doing/seeing* patterns have the highest frequency among the five typical S-patterns and among the eight structural types of four-word bundles, other NP, NP + of, other PP, and be + NP/AP have the highest frequency. In terms of the frequency of structural types of LBs, there are differences from the results of Hyland (2008a) because the present study is based on a corpus analysis of IELTS academic reading texts. Moreover, in this study, the IELTS academic reading corpus was established to serve specific student groups, and the S-pattern framework provided a real discourse example (such as discourse function and grammatical structure) for observation and learning. As a result, this research provides a method for teachers to introduce EFL learners or IELTS candidates to the target language patterns through the AntConc tool, facilitating the retrieval of LBs and S-patterns via corpus analysis.

The AntConc tool does not directly analyze S-patterns, which is a limitation of this study. When analyzing S-patterns, the functions of Concordance and File View must classify and observe the context of the text. Furthermore, similar to other researchers' findings, there may be some minor differences in the data collected in this study, but these differences have little bearing on the overall analysis results and discussion. The analysis of sentence diversity in other English language tests, such as TOEFL, GMAT, and GRE, can be explored and compared with the findings of this study to see if there are any differences so that innovative teaching methods, reading strategies, and writing suggestions can be provided to teachers, EFL students, and test writers.

References

Ahmadian, M., Poulaki, S., & Farahani, E. (2016). Reading strategies used by high scoring and low scoring IELTS candidates: A think-aloud study. *Theory and Practice in Language Studies*, *6*(2), 408-416. https://doi.org/10.17507/tpls.0602.25

Akbas, E., & Hardman, J. (2018). Strengthening or weakening claims in academic knowledge construction: A comparative study of hedges and boosters in postgraduate academic writing. *Educational Sciences: Theory & Practice*, *18*, 831-859. http://dx.doi.org/10.12738/estp.2018.4.0260

Alshammari, M. M. (2016). IELTS academic reading module test: Validity and reliability. *British Journal of English Linguistics*, *4*(2), 26-37. http://doi.org/10.37745/bjel.2013

Ang, L. H., & Tan, K. H. (2018). Specificity in English for Academic Purposes (EAP): A corpus analysis of lexical bundles in academic writing. *3L, Language, Linguistics, Literature*, *24*(2). https://doi.org/10.17576/3L-2018-2402-07

Anthony, L. (2019). *AntConc* [Computer Software]. Waseda University. http://www.laurenceanthony.net/software/antconc/

Aston, G. (2001). Learning with corpora: An overview. *Learning with corpora*, 1-39. http://digital.casalini.it/10.1400/34706

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English*. Longman.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371-405. https://doi.org/10.1093/applin/25.3.371

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, *26*, 263-286. https://doi.org/10.1016/j.esp.2006.08.003

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. https://doi.org/10.1075/ijcl.14.3.08bib

Bowen, N. E. J. A. (2019). Unfolding choices in digital writing: A functional perspective on the language of academic revisions. *Journal of Writing Research*, *10*(3), 465-498. http://doi.org/10.17239/jowr-2019.10.03.03

Bowen, N. E. J. A., & Thomas, N. (2020). Manipulating texture and cohesion in academic writing: A keystroke logging study. *Journal of Second Language Writing*, *50*, 1-15. https://doi.org/10.1016/j.jslw.2020.100773

Cambridge English. (2020). *Cambridge IELTS 15 academic: Authentic practice tests*. Cambridge University Press.

Carrell, P. L., & Grabe, W. (2002). Reading. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 233-250). Arnold. https://doi.org/10.4324/9780429424465-15

Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2011). Selecting information to answer questions: Strategic individual differences when searching texts. *Learning and Individual Differences*, *21*(2), 201-205. https://doi.org/10.1016/j.lindif.2010.11.007

Cerdán, R., Pérez, A., Vidal-Abarca, E., & Rouet, J. F. (2019). To answer questions from text, one has to understand what the question is asking: differential effects of question aids as a function of comprehension skill. *Reading and Writing, 32*(8), 2111-2124. https://doi.org/10.1007/s11145-019-09943-w

Chen, C., & Liu, Y. (2020). The role of vocabulary breadth and depth in IELTS academic reading tests. *Reading in a Foreign Language*, *32*(1), 1-27. https://doi.org/10125/66574

Cohen, A. D. (1998). *Strategies in learning and using a second language*. Longman. https://doi.org/10.4324/9781315833200

Cohen, A., & Upton, T. (2007). *Strategies in responding to the new TOEFL reading tasks*. Monograph Series: ETS. https://doi.org/10.1002/j.2333-8504.2006.tb02012.x

Coombe, C., Folse, K., & Hubley, N. (2007). *A practical guide to assessing English language learners.* The University of Michigan Press. https://doi.org/10.3998/mpub.170430

Cortes, V. (2002). Lexical bundles in freshman composition. *Using corpora to explore linguistic variation, 9*, 131-145. https://doi.org/10.1075/scl.9.09cor

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, *23*(4), 397-423. https://doi.org/10.1016/j.esp.2003.12.001

De Milliano, I., Van Gelderen, A., & Sleegers, P. (2016). Types and sequences of self-regulated reading of low-achieving adolescents in relation to reading task achievement. *Journal of Research in Reading*, *39*(2), 229-252. https://doi.org/10.1111/1467-9817.12037

Francis, W. N. (1982). Problems of assembling and computerizing large corpora. In S. Johansson (Ed.), *Computer corpora in English language research* (pp. 7-24). Norwegian Computing Centre for the Humanities.

Gabrielatos, C. (2005). Corpora and language teaching: just a fling or wedding bells? *Teaching English as a Second or Foreign Language*, *8*(4), 1-32. http://www.tesl-ej.org/ej32/a1.html.

Gehrer, K., Zimmermann, S., Artelt, C., & Weinert, S. (2013). NEPS framework for assessing reading competence and results from an adult pilot study. *Journal for Educational Research Online*, *5*(2), 50-79. https://doi.org/10.25656/01:8424

Gezegin-Bal, B. (2019). Lexical bundles in published research articles: A corpus-based study. *Journal of Language and Linguistic Studies*, *15*(2), 520-534. https://doi.org/10.17263/jlls.586188

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press. https://doi.org/10.1017/CBO9781139150484

Gries, S. T. (2016). *Quantitative corpus linguistics with R: A practical introduction*. Routledge. https://doi.org/10.4324/9781315746210

Holi, A. I., Al Washahi, Q., & Alhassan, A. A. (2020). Unpacking the challenges and accommodation strategies of Omani English-major students on IELTS academic reading tests. *Journal of Language and Linguistic Studies*, *16*(3), 1621-1636. https://doi.org/10.17263/jlls.803922

Hryniuk, K. (2018). Expressing authorial self in research articles written by Polish and English native-speaker writers: A corpus-based study. *Studies in Second Language Learning and Teaching*, *8*(3), 621-642. http://dx.doi.org/10.14746/ssllt.2018.8.3.5

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524773

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language writing*, *6*(2), 183-205. https://doi.org/10.1016/S1060-3743(97)90033-3

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, *7*(2), 173-191. https://doi.org/10.1177/1461445605050365

Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4-21. https://doi.org/10.1016/j.esp.2007.06.001

Hyland, K. (2008b). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, *18*(1), 41-62. https://doi.org/10.1111/j.1473-4192.2008.00178.x

Kolln, M., & Funk, R. (2011). *Understanding English grammar* (9th ed.). Longman.

Kuzar, R. (2012). *Sentence patterns in English and Hebrew*. John Benjamins Publishing. https://doi.org/10.1075/cal.12

Lems, K., Miller, L. D., & Soro, T. M. (2017). *Building literacy with English language learners: Insights from linguistic*s. Guilford Publications.

Liao, L. (2020). A comparability study of text difficulty and task characteristics of parallel academic IELTS reading tests. *English Language Teaching*, *13*(1), 31-42. https://doi.org/10.5539/elt.v13n1p31

Marzban, A., & Barati, Z. (2016). On the relation between critical thinking, languages learning strategies, and reading comprehension of male and female intermediate EFL university students. *Theory and Practice in Language Studies, 6*(6), 1241-1247. https://doi.org/1016/j/tpls.0606.14

McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

McKoon, G., & Ratcliff, R. (2018). Adults with poor reading skills, older adults, and college students: The meanings they understand during reading using a diffusion model analysis. *Journal of Memory and Language, 102*, 115-129. https://doi.org/10.1016/j.jml.2018.05.005

Moore, T., Morton, J., & Price, S. (2012). Construct validity in the IELTS academic reading test: A comparison of reading requirements in IELTS test items and in university study. *Studies in Language Testing, 11*(4), 120-211. https://www.ielts.org/for-researchers/research-reports/volume-11-report-4

Nation, K., & Norbury, F. (2005). Why reading comprehension fails: Insights from developmental disorders. *Topics in Language Disorders, 25*(1), 21-32. https://doi.org/10.1097/00011363-200501000-00004

O' Keeffe, A., & Fair, F. (2003). Using language corpora in initial teacher education: Pedagogical issues and practical applications. *TESOL Quarterly, 37*(3), 389-418. https://doi.org/10.2307/3588397

O' Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom. Language use and language teaching*. Cambridge University Press. https://doi.org/10.1017/CBO9780511497650

Panthong, P., & Poonpon, K. (2020). Lexical bundles in Thai medical research articles. *Journal of Studies in the English Language, 15*(1), 59-106.

Pérez-Paredes, P., Ordoñana Guillamón, C., & Aguado Jiménez, P. (2018). Language teachers' perceptions on the use of OER language processing technologies in MALL. *Computer Assisted Language Learning, 31*(5-6), 522-545. https://doi.org/10.1080/09588221.2017.1418754

Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing*, *20*(1), 26-56. https://doi.org/10.1191/0265532203lt243oa

Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge University Press.

Rogers, W. T., & Harley, D. (1999). An empirical comparison of three-and four-choice items and tests: Susceptibility to testwiseness and internal consistency reliability. *Educational and Psychological Measurement*, *59*(2), 234-247. https://doi.org/10.1177/00131649921969820

Simsek, R., & Gün, M. (2021). A corpus-based proposal for the vocabulary to be taught at A1 Level in the teaching of Turkish as a foreign language. *Educational Policy Analysis and Strategic Research*, *16*(1), 356-373. https://doi.org/10.29329/epasr.2020.334.19

Sinclair, J. M. (2004). *How to use corpora in language teaching*. Benjamins. https://doi.org/10.1075/scl.12

Smirnova, E. A. (2017). Using corpora in EFL classrooms: The case study of IELTS preparation. *RELC Journal*, *48*(3), 302-310. https://doi.org/10.1177/0033688216684280

Stevenson, M., Schoonen, R., & de Glopper, K. (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, *15*(3), 201-233. http://dx.doi.org/10.1016/j.jslw.2006.06.002

Sullivan, K. D., & Eggleston, M. (2006). *The McGraw-Hill desk reference for editors, writers, and proofreaders*. McGraw-Hill.

Taymaz, N. (2021). A corpus-based comparison of use of hedges and boosters by Turkish ELT MA and PhD students. *Journal of Language and Linguistic Studies*, *17*(1), 33-49. https://doi.org/10.17263/jlls.903302

Verspoor, M., & Sauter, K. (2000). *English sentence analysis: An introductory course*. John Benjamins Publishing. https://doi.org/10.1075/z.100

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC journal*, *42*(2), 155-166. https://doi.org/10.1177/0033688211407295

Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. Wiley-Blackwell. https://doi.org/10.1002/9781119180180

Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2005). Less skilled comprehenders' ERPs show sluggish word-to-text integration processes. *Written Language & Literacy*, *8*(2), 157-181. https://doi.org/10.1075/wll.8.2.10yan

Zare, P. (2012). Language learning strategies among EFL/ESL learners: A review of literature. *International Journal of Humanities and Social Science, 2*(5), 162-169. https://doi.org/10.30845/ijhss