



Contents lists available at ScienceDirect

Journal of English for Academic Purposes

journal homepage: www.elsevier.com/locate/jeap

Bundles to beat the band in high-stakes tests: Pedagogical applications of an exploratory investigation of lexical bundles across band scores of the IELTS writing component

Ali Saadatara^a, Gholamreza Kiany^a, Hossein Talebzadeh^{b,*}

^a Tarbiat Modares University, Tehran, Iran

^b Department of Foreign Languages, Faculty of Literature and Humanities, Kharazmi University, Tehran and Karaj, Iran

ARTICLE INFO

Keywords:

Genre

High-stakes tests

IELTS

Lexical bundles

Mock exams

Writing component

ABSTRACT

Lexical bundles (LBs) in ESL/EFL examinees' writings and the effectiveness of their instruction on the learners' performance are under-researched as far as international high-stakes examinations, particularly IELTS (International English Language Testing System) writing genre(s), are concerned. To fill these gaps, the present two-phased study, in its exploratory phase, aimed at investigating the form, functions, and distribution of LBs in approximately 1000 mock academic IELTS writing tasks across three levels of writing competence. For the next phase, a pedagogical intervention was implemented and its impact on improving the performance of the IELTS candidates was evaluated. The results demonstrated that the higher the proficiency of the examinees, the higher the frequency and range and the more varied the structures and functions of deployed LBs. Additionally, the majority of the LBs were equally dispersed in all parts of the texts although there were some instances which characterized introduction, body, or conclusion sub-sections. Finally, the adopted explicit functional pedagogical approach was found to be effective, especially in boosting cohesion and coherence and lexical resources scores. Highlighting the nature of IELTS texts as distinct genres, we conclude with the implications of studying (and teaching) LBs for writing instruction and examination.

1. Introduction

The last decade has witnessed an accumulating body of research in EAP/ESP writing studies on lexical bundles, which has been supported by continued refinements in computer-assisted methods of discourse analysis. Also called clusters (Hyland, 2008; Schmitt, 2004), n-grams (Banerjee & Pedersen, 2003), phrasal lexemes (Moon, 1998), formulaic sequences (Li & Schmitt, 2009; Schmitt & Carter, 2004), recurrent word combinations (Altenberg, 1998; De Cock, 2000), chunks, multiword expressions (such as collocations and idiomatic expressions), prefabricated chunks, ready-made utterances, multiword constructions, formulas (Wood & Appel, 2014; Wray, 2002), sentence stems (Syder & Pawley, 2014), chains and bundles, these structures appear to occur in discourse—both spoken and written—more frequently than has been previously understood (Biber et al., 1999; Hyland, 2008; Shirazizadeh & Amirfazlian, 2021; Simpson-Vlach & Ellis, 2010; Stubbs & Barth, 2003). Wray (2002, p. 9) defines a formulaic sequence as “a sequence, continuous

* Corresponding author.

E-mail addresses: ali.saadatara@modares.ac.ir (A. Saadatara), kiany_gh@modares.ac.ir, rezakiany_gh@modares.ac.ir (G. Kiany), talebzadeh@khu.ac.ir, htmu88@yahoo.com (H. Talebzadeh).

<https://doi.org/10.1016/j.jeap.2022.101208>

Received 15 July 2021; Received in revised form 7 December 2022; Accepted 12 December 2022

Available online 15 December 2022

1475-1585/© 2022 Elsevier Ltd. All rights reserved.

or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar.” The hypothesis concerning the holistic storage and retrieval of formulaic sequences is claimed to be supported by mixed and inconclusive evidence, which questions the psycholinguistic validity of formulaic sequences as single wholes (e.g. Biber et al., 1999; Ellis et al., 2008; Jiang & Nekrasova, 2007; Nekrasova, 2009; Tremblay et al., 2007; Tremblay et al., 2011, among others); nevertheless, it is argued that L1 users of English process these structures faster than L2 users of English due to the ample exposure they have to input as compared to L2 users’ impoverished input (Ellis et al., 2008).

As Staples et al. (2013) argue, a notable proportion of the available literature has focused on the differences between (English) L1 and L2 users’ application of LBs, while only a few studies have explored LBs across proficiency levels (e.g., Vo, 2019). Furthermore, even the findings of such few studies appear to be contradictory, with some concluding that learners with lower proficiency use more LBs than those with higher proficiency and vice versa (Boers et al., 2006; Forsberg, 2010; Myles et al., 1998; Vo, 2019). Additionally, despite their significant consequences for language learners’ future, international high-stakes examinations (including IELTS and TOEFL) are not aptly investigated in the literature, and they remain to be what Hyland (2003) calls *Cinderella genres* (one exception can be Staples et al., 2013; see Pearson, 2019 on the increasing power and impact of IELTS as a gate-keeping institution as well as its being a distinct written genre). Finally, many studies conducted on lexical bundles have been predominantly descriptive in nature, without making ample efforts to be more explanatory or reveal the potential effects that instruction of LBs might have on the students’ actual performance and writing development (e.g., Li & Schmitt, 2009).

To fill these gaps, the present two-phased study aimed at investigating LBs in a corpus of writing Tasks 1 and 2 of IELTS (Academic) mock exams in its first, exploratory phase. Specifically, in the first phase, we intended to identify the most common LBs in the whole corpus, in each of the three score categories—below (yet not including) 5; 5 to 7 (excluding 7); above 7—of IELTS writing tasks; we also identified the most common LBs in relation to the scores received for the selected IELTS scoring criteria (i.e., Lexical Resource (LR), Grammatical Range and Accuracy (GRA), and Cohesion and Coherence (CC)) (hence our first research question). Additionally, we sought to examine the LBs across the three score categories of mock IELTS writing tasks in terms of their frequencies, structures, and functions (hence the second research question). Moreover, the distribution of the LBs, namely, where each LB with a particular function occurred in the written texts (i.e., Introduction, Body, and/or Conclusion) was investigated (namely, the third research question). In the second phase of the study, the researchers examined whether instruction has any significant effects on the participants’ Tasks 1 and 2 texts. Particularly, we wanted to see whether the teaching of high achievers’ LBs could have significant effects on the lower achievers’ overall IELTS Academic writing scores and their selective criteria (i.e., CC, LR, and GRA) scores (therefore, our last research question).

2. Literature review

2.1. Lexical bundles: definition, processing, structure and function

The advancement of technology and developments in corpus-related studies have enabled researchers in language pedagogy and corpus linguistics to get a clearer grasp of both spoken and written discourse. The last decade has witnessed increased empirical research that attends to disciplinary, genre, population, and discourse variations of lexical bundles. As noted above (see 1. Introduction), Lexical bundles (LBs) have been defined and named differently in the literature. For Nattinger and DeCarrico (1992) lexical phrases are “chunks of language of varying lengths” (p.1) with remarkable formal, functional, and statistical attributes. Lexical bundles and formulaic sequences (most often used interchangeably) are regarded as fixed formulas, processing and retrieving of which have been argued to be holistic and top-down.

This formulaicity and holistic processing are best captured in Wray’s definition of formulaic sequences (see 1. Introduction above) as well as Palmer’s definition of collocations, accentuating the holistic nature of LBs, as “successions of words that (for various, different and overlapping reasons) must or should be learned, or is best, or most conveniently learned as an integral whole or independent entity, rather than by the process of placing together their component parts” (Palmer, 1933, p.4). Consequently, their processing and retrieval are much easier for both L1 and L2 users (Wray & Perkins, 2000); still, L1 users of a language are believed to have an advantage in comprehension and production over L2 users thanks to the differential quality and quantity of input (Ellis et al., 2008; see also Adel & Erman, 2012; Karabacak & Qin, 2013). As noted above, there is nevertheless still not a consensus on “the question of whether formulaic sequences are stored and processed holistically in the mind” (Tremblay et al., 2007, p. 260), and there are some conflicting results (depending on the task and experiment types) which cast doubts on taking LBs as valid psycholinguistic units retrieved from memory as single wholes (see Jiang & Nekrasova, 2007; Nekrasova, 2009; Tremblay et al., 2007; Tremblay et al., 2011; for examples and reviews).

LBs play some structural and functional roles in discourse, so a number of taxonomies have been proposed to help researchers get a clearer grasp of those structures and their functions. Among the common taxonomies adopted for this purpose, we can refer to Biber et al.’s (1999) classification of the structures of LBs as well as Hyland’s (2008) and Goźdz-Roszkowski’s (2011) classifications of functions of LBs. Biber and Conrad’s (1999) taxonomy includes structures like, *inter alia*, *NP + of*, *prepositional phrase + of*, *passive + prepositional phrases*, *anticipatory it + verb/adjective*. The three functions of Hyland’s (2008) taxonomy include research-oriented (including *location*, *procedure*, *quantification*, *description*, and *topic*), text-oriented (including *transition signals*, *resultative signals*, *structuring signals*, and *framing signals*), and participant-oriented functions (including *stance features* and *engagement features*). Additionally, Biber et al.’s (2004) study has a functional taxonomy including *stance*, *discourse organizing*, and *referential* functions with their own subparts. In fact, when it comes to investigating the structure and function of LBs, Biber and his colleagues (e.g., Biber et al., 1999;

Biber et al., 2003, 2004) have proposed highly influential structural and functional taxonomies of lexical bundles.

Biber et al. (2004) compared the functional use of LBs in university classroom teaching and in academic textbooks, finding that LBs appeared significantly more frequently in university classrooms than in textbooks, since classroom discourse yielded four times as many LBs as textbooks did. More importantly, a significant aspect of Biber et al.'s (2004) study is using the above-mentioned taxonomy, i.e. classifying LBs into stance bundles (which show attitudes), referential bundles (which denote "physical or abstract entities,") and discourse LBs (which are conducive to the organization of discursive moves). In general, the functional taxonomies differ considerably across genres and registers where more fine-grained functional distinctions are recorded among the three broad categories of referential, discourse-organizing and stance LBs (see Biber & Barbieri, 2007; Biber & Conrad, 2019; Biber et al., 2004).

Nattinger and DeCarrico (1992) suggest that there are four criteria characterizing prefabricated chunks. The first criterion they propose refers to their length and grammatical level, so LBs can be of word level (e.g. all in all) or sentence level (e.g. Can I help you?) (as far as the word level is concerned, it is worth noting that some LBs constitute neat form-to-meaning mappings; therefore, they act as if they were single words expressing a given sense in a particular context of use). The second criterion concerns whether the LBs have canonical (e.g., on the other hand) or non-canonical (e.g., by and large) shapes. The third criterion refers to whether an LB is a variable construction (e.g. this is an X) or a fixed one (e.g. what is this?). And the last one concerns whether the LB is continuous or discontinuous, i.e., whether they are adjacent or separated by other lexical fillers (cited in Strunkyte & Jurkūnaite, 2008).

Lexical bundles in the current study refer to bundles of words which are rarely complete structurally and are functionally complete and/or incomplete, with or without any idiomatic meaning, and with or without semantic transparency (see Biber & Barbieri, 2007; Biber & Conrad, 1999 for the finer distinctions between and among collocations, idioms, and LBs). In this study, three to six-word LBs are analyzed on account of their prevalence in the literature and their high frequency (Appel & Wood, 2016; Li & Volkov, 2018). It should be emphasized that although we have initially identified both complete and incomplete bundles, the priority is given to either complete sequences or the most relevant, useful incomplete ones which could best align with our ultimate pedagogical incentives (see also Section 3 below).

2.2. LBs and generic, disciplinary, register, and discourse variation

Different situational contexts, production circumstances, and communicative purposes necessitate the use of certain words and LBs. This *linguistic variation* is captured in terms of the versatility in the use of lexical items in different situational contexts as a result of the constraints these situations impose on the speakers or writers (Biber, 2006; Biber & Conrad, 2001, 2019). Hence, by gathering corpora from these diverse situational contexts, researchers have revealed the multiplicity and versatility of lexical items (Grabowski, 2015). The underlying assumption of variation of LBs has been scrutinized and substantiated through investigation of different genres (like research articles, MA theses, Ph.D. dissertations, university textbooks, and EAP textbooks, among others), disciplines (e.g. biology, applied linguistics, history, engineering), writing materials of L1 versus L2 users (of English), expert writers versus novice writers, spoken discourse versus written discourse, and academic discourse versus non-academic discourse (e.g. Biber et al., 1999; Brazil, 1995; Hyland, 2008; Leech, 2000; Shirazizadeh & Afzal, 2021; Vo, 2019; Wood & Appel, 2014).

Biber et al. (1999) note that "lexical bundles are identified empirically, as the combinations of words that in fact recur most commonly in a given register" (p. 184). So, the assumption is that LBs are theoretically linked to specific registers. For instance, Byrd and Coxhead (2010), using a 3.6-million-word corpus consisting of four academic disciplines, namely science, law, art, and commerce, analyzed the disciplinary variation existing among these disciplines. Thereby, they found 73 LBs common to all disciplines and provided the frequency of each in all the disciplines.

Cortes (2004) compared student papers with published research articles, revealing that students' use of LBs were comparatively few and far between. Likewise, Hyland (2008) compared LBs used in published academic discourse with postgraduate student papers, the results of which indicated that, to display their competence, postgraduate students went to extremes in using formulaic expressions compared to L1 users of English. The inconclusive findings of such studies necessitate further research with larger and different corpora to examine (and compare) different applications of LBs by student writers (and L1 writers).

Johnston (2017), with regard to the potential differences in the use of LBs by professional writers and learners, attempted to find the differences of LBs in terms of frequency, structure, and function in the texts written by both professionals and learners in the two fields of applied linguistics and literature. She concluded that professional writers in both disciplines used LBs with different frequencies and type/tokens due to the rhetorical requirements of each discipline. Also, learners in both disciplines tended to use the same LBs repetitively as compared to the wider range of LBs used by professional writers of each discipline. In general, there is a myriad of studies which scrutinized LBs across different genres, registers, or language users; nevertheless, as will be noted later (Sub-section 2.4.), there are still some major communities of practice—e.g., people like L2 learners and teachers whose life can be affected by high-stakes international language examinations (see Pearson, 2019)—who might quantitatively and qualitatively resort to LBs differently, for whom mastery and non-mastery of LBs might have significant implications, and whose needs and practices have been underrepresented in the existing literature on LBs (for more on communities of practice see Eckert & McConnell-Ginet, 1998 as well as Wenger et al., 2002).

2.3. LBs across proficiency levels

Studies which have specifically focused on LBs across proficiency levels are comparatively few. One study which delved into the analysis of LBs across proficiency levels was conducted by Li and Volkov (2018). Their corpus, which was a collection of email tasks written by L2 students across three proficiency levels, was divided into three subcorpora, assigning each subcorpus to one proficiency

level. Using AntConc software, they identified the patterns of two to six-word LBs in the emails and then manually coded the functions of each LB. They concluded that as proficiency level increased, the use of LBs in terms of both types and tokens increased. Furthermore, the writing samples of those students at the highest level of proficiency shared a large number of common LBs.

Likewise, [Chen and Baker \(2016\)](#) studied four-word LBs in the writing samples of Chinese L2 learners of English derived from Longman Learner Corpus. Again, they categorized the corpora into three subcorpora across three proficiency levels according to [Cambridge University Press \(2019\)](#) (Common European Framework of Reference), namely B2, B1, and C1. They found that students with lower proficiency levels made use of speech-like LBs in their passages, while students with higher proficiency levels displayed a more academically sophisticated style. Furthermore, they proposed that some of the LBs used by lower-level students were not academically refined and lacked general appropriacy.

Among the few studies of LBs in a writing assessment context, [Appel and Wood \(2016\)](#) analyzed 4-to-7-word LBs from the written texts of lower-level students and higher-level ones in the Canadian Academic English Language Assessment. They explored the differences of functional types of LBs, namely stance, referential, and discourse organizing in the samples of argumentative essays written by L2 learners of English. They postulated that referential LBs were used by higher level students more than lower-level ones, while lower-level students used more of the other two functional types of LBs, namely stance and discourse organizing.

Another noteworthy study of LBs across proficiency levels is that of [Staples et al. \(2013\)](#) (see also the next sub-section for the significance of this study) which examined the frequency and use of LBs using participants' written responses of TOEFL iBT as the corpus. While it might be postulated otherwise, they found that lower-level students generally used more bundles. Additionally, the use of stance and discourse organizing LBs was revealed to be the same among all the proficiency levels, while only a few instances of referential LBs had been used by learners across all proficiency levels. In general, the results pertaining to LBs across proficiency levels tend to be varied based on the genre and context, appear to be less conclusive, and indicate the need for further studies of LBs in L2 texts across levels of (general language and writing) competence levels.

2.4. LBs and pedagogical approaches

As argued by [Wray \(2002\)](#), formulas can facilitate idiomatic production and also signal that a speaker or writer holds an *insider* status in a discourse community (see also [Pearson, 2019](#) and [Shirazizadeh & Afzalian, 2021](#)). With this in mind, we should notice the paramount importance of focusing on teaching, learning, and researching LBs within the context of every discipline, genre, or discourse. So, the question raised here is what LBs students and L2 researchers need in order to perform better in the target language. Moreover, although research on the register and disciplinary variation of LBs abounds (see our brief review above), there is a dearth of research assuming a pedagogical perspective for the teaching of LBs. In fact, as observed by [Byrd and Coxhead \(2010\)](#), there still remains little empirical evidence as to the efficacy of different approaches of teaching LBs, leaving instructors with unclear guidance for exploiting this important component of language teaching (see [Reppen, 2018](#) for a more recent review).

In her book, [Salazar \(2014\)](#) takes a step toward filling this gap by first providing the readers with a list of useful LBs, and the techniques which can be employed by teachers in the subsequent chapters. Like [Biber et al. \(1999\)](#), [Salazar \(2014\)](#) uses three-word, four-word, and five-word LBs. Similarly, [Nattinger and DeCarrico \(1992\)](#) brought lexical bundles in language teaching into the fore; in fact, notwithstanding the ambiguity of their methodology in detecting formal and functional disparities among different registers, the results of their work have been significant and have attended to pedagogical dimensions of LBs with reference to the four skills. Still another line of research which brought LBs into the forefront of language teaching was that of [Lewis's \(1993\)](#) Lexical Approach. Following [Sinclair's \(1991\)](#) idiom principle, the lexical approach was a reaction to traditional views on grammar and vocabulary and their applications in pedagogy. By allocating minimum attention to individual words, [Lewis \(1993\)](#) underscores the formulaicity and chunkedness of natural language.

Investigating the impact of explicit teaching of LBs on students' papers in an intensive history class, [Cortes \(2006\)](#) incorporated into the history class some mini-lessons in order to raise the students' awareness of such LBs. Surprisingly, the results of pre- and post-tests of students' texts showed that there was not any remarkable improvement in the LBs' frequency of use. Nevertheless, Cortes argued that the mini-lessons augmented students' awareness of and motivation to use LBs in their future writing assignments. Likewise, [Spöttl and McCarthy \(2003\)](#) and [O'Keeffe et al. \(2007\)](#) point to the necessity of conscious linguistic analysis during the process of learning. They argue that a language classroom is exactly the place that such kind of conscious analysis should be advocated. Similar to grammatical structures and single lexical items, they claim that phraseological knowledge should be acquired through repeated exposures.

Adopting a holistic approach to investigate teaching and learning of academic lexical phrases, [Cai \(2016\)](#) devised a genre-based intervention study of Chinese Masters students. The study results highlighted the efficacy of LB teaching in their appropriate receptive and productive mastery in tandem with the growth of genre knowledge and move-specific lexical phrases. [Ghafarsamar et al. \(2018\)](#), likewise, investigated the effects of teaching LBs on the writing performance in an EFL context (while teaching LBs to medical students). Their aim and criteria for the measurement of writing proficiency in their study were partially comparable to the ones utilized in the present study, i.e. TR, CC, LR, and GRA. However, their study differs from ours in terms of the genre of interest, not to mention being constrained by the limited number of participants and lack of a control group. Furthermore, only a limited number of four-word LBs were drawn upon in their principally implicit instructions. Their results indicated that the effects of instruction were mostly manifested in the scores participants achieved on the LR measure. Furthermore, statistically significant differences were also observed for TR and CC while the participants did not show any improvement on GRA scores (which was putatively attributed to the short duration of the intervention).

Our brief review of the pertinent studies divulging the disciplinary and register variations of LBs (2.2. above), their variation across

proficiency levels (2.3. above), and the effectiveness of their instruction (2.4.) underlines that some disciplines, registers, and genres are under-represented or not represented in the literature. Among the neglected genres are the “genres of high-stakes writing assessments” (Beck & Jeffery, 2007, p. 60) including the writing module of international language examinations across proficiency levels. This paucity is in spite of the burgeoning population of language learners whose personal, academic, and/or professional future can be contingent on their performance in important exams such as IELTS or TOEFL (Test of English as a Foreign Language) (Pearson, 2019). Apart from Apple and Wood (2016), another rare counterexample is Staples et al. (2013) which—as explained above—examined the frequency and use of LBs across proficiency levels using participants’ written responses of TOEFL iBT as the corpus. Notwithstanding, the genre of the writings of IELTS in relation to LBs (and their instruction) has been under-represented (for a review of formulaic language in other assessment contexts see Bestgen, 2017). Consequently, the present study attempts to fill this gap by investigating the frequency, type, and distribution of LBs dominant in a corpus of mock IELTS writing tasks written by EFL examinees. Moreover, the effect of LBs’ instruction on a similar sample of learners is examined.

3. Methodology

The study was implemented in two major phases. The first phase, which delved into the features of the LBs characterizing EFL learners’ IELTS writings across three levels of writing competence, was descriptive and exploratory in nature; however, the second phase, which employed the findings of the first phase to assess the efficacy of teaching LBs to IELTS low achievers, was experimental and pedagogical in nature.

3.1. Corpus description

The corpus of IELTS writing tasks was obtained from the database of one of the locally-renowned IELTS teaching centers, which claims a considerably high correlation between the scores obtained by the students on its mock IELTS exams and their actual IELTS performance. Apart from the teaching approach, this correlation is maintained to be due to their employing certified IELTS examiners in evaluating their mock examinations. Although there are some published actual IELTS writing answer sheets (e.g., Cambridge IELTS series), the available data was not sizeable enough to meet the requirements of the first exploratory phase of the study across three levels of writing competence. Plus, given the demography of the target participants of the second pedagogical phase of the study (who were Iranian IELTS candidates), it was decided to focus on the papers by writers who resembled the target learners; by so doing we were trying to control for as many extraneous variables as possible to gear the instruction accordingly. Consequently, the present study utilized a corpus of 999 IELTS mock writing Tasks 1 and 2 answer sheets written by candidates with different levels of writing competence; they were scored by the institutions’ qualified examiners (just one rating per each paper by one of the several anonymous raters of the institution who were required by the institutional policy to meet regularly to ensure scoring consistency) based on the four scoring criteria of Task Response/Task Achievement (TR), Cohesion and Coherence (CC), Lexical Resource (LR), and Grammatical Range and Accuracy (GRA) as well as an overall band score. The rubrics for Task 1 included a wide gamut of topics most of which contained graphs and charts, while the rubrics for Task 2 mostly encompassed essay questions that had required the candidates to state their own personal opinion or argument on a given topic.

The corpus contains around 200,000 words, almost 66,000 words from each of the three band score categories of 7–9, 5–7 (excluding 7), and below 5 (excluding 5). One (practical) rationale for this categorization was that the desired overall (and writing) band score for most academic purposes the participants are pursuing is above 6.5. Moreover, it can be compared with the CEFR standards and distinguish basic users with extremely limited to limited competence (below 5 band-scores and roughly comparable to A1, A2, and preliminary B1 levels) from independent users with modest to competent ability (5–6.5 band scores roughly comparable to B1 and B2 levels), and from proficient users with good to expert proficiency (7–9 band scores and comparable with C1 and C2 levels) (see Arefsadr, 2022; Cambridge Assessment English, 2022). Consequently, 333 written texts from each band score category were selected and analyzed to provide us with a balanced corpus (in terms of number of texts and not necessarily word count). Since Task 1 and Task 2 answer sheets are required to be restricted within an average of 150 and 250 words, respectively, there were no marked

Table 1
Characteristics of the corpus based on task type and across band scores.

BAND SCORE	TASK 1				TASK 2			
	Number of Texts	Word Count (Mean)	Task 1 Total words	SD	Number of Texts	Word Count (Mean)	Task 2 Total words	SD
Below 5 (excluding 5)	167	117.1	19546	51.8	166	193.1	32058	70.6
5 to 7 (excluding 7)	167	157.9	26379	42.8	166	273.6	45421	58.3
7–9	167	173.1	28909	34.8	166	286.7	47596	47.3
Total	501	149.4	74834		498	251.2	125075	

variations necessitating data normalization. As shown in Table 1, in spite of the differential length of texts across the three competence levels (i.e., lower proficiencies used fewer words), it can be generally observed that the word counts of the Task 1 and Task 2 texts included in the corpus (totaling around 75,000 and 125,000 words, respectively) reflect the proportionate recommended word count for each task type (i.e., Task 1: 150 words; Task 2: 250 words).

3.2. Participants

The participants who took part in the second phase of the study were 60 Iranian EFL learners with varying (writing) proficiency levels, ages, and educational backgrounds. The participants were randomly assigned to either the experimental group or the control group. Based on their pretest scores (see 3.3. and 3.4. below), both groups were further divided into two groups: one group consisted of those participants who had gained an overall band score of 5–7 (excluding 7) and the other group had gained an overall band score of below 5 (excluding 5). Therefore, the participants were assigned to two groups and four separate classes (each including 15 candidates) who received instruction either with or without a focus on lexical bundles (see 3.4.).

3.3. Instrumentation

The instruments (for phase 2) included a pretest and a posttest. Both the pretest and the posttest were representative samples of IELTS writing task rubrics selected from IELTS Practice Test Series (2019). The questions in Task 1 in both exams consisted of a bar chart representing some statistics required to be compared, contrasted, and reported on, and those of Task 2 asked the participants to write an argumentative essay on a potentially controversial topic.

3.4. Data collection

In the exploratory phase of the study, the researchers transcribed the scored writing answer sheets one by one verbatim (as they were handwritten), fed them to the AntConc concordancing software (Anthony, 2006), and analyzed them (see 3.5 below). The results from the first phase were used for instruction in the second phase of the study (which is elucidated step-by-step below). In particular, the preferred LBs for inclusion in the pedagogical intervention phase of the study were those frequent items and functions which had predominantly appeared in the texts of high-achievers (7–9 band score).

All the participants who took part in the instructional treatment designed for the second phase of the study were given the pretest. The pretests were collected, transcribed, and sent to a certified IELTS ex-examiner to be scored in terms of the four criteria of TA, CC, LR, and GRA. In conjunction with the scores, the examiner corrected the grammatical, stylistic, lexical, and task-related errors and provided the participants with some insightful comments. As was the case with the first phase ratings, the logistic considerations made it impossible to recruit raters to reexamine the texts; however, it was ensured that the papers had been scored by qualified examiners (including certified IELTS ex-examiners) with relevant, extended job experience in a top-tier IELTS preparation institution for phase one and were rated by an ex-examiner of a certified IELTS administration center for phase two.

The instructions for both groups (i.e., four classes) included a total of 10 sessions with the duration of 30–45 minutes except for an additional 2-hour session for the experimental group, in which an explicit functional focus was paid to the lexical bundles obtained from the first phase of the study. This explicit functional approach was adopted since, as Cortes (2004) argued, learners are less likely to accidentally notice the functions LBs serve in different contexts. As a consequence, in such situations where the functions are not elaborated on, learners tend to resort to avoidance strategies when they are not familiar with the functions of LBs, which can lead to their written texts being peculiar and redundant. Also, working with the AntConc concordancing tool and learning to search for the common patterns of LBs in any sort of text input distinguished the two experimental classes from their counterparts.

The ten-session instructions were all asynchronous. Two online classrooms were created on Google Classroom platform and all the instructional videos together with the activities for each group were uploaded on their respective classrooms. The lessons encompassed a wide array of topics including some preliminary ones on general writing issues in the first sessions prior to the teaching of the points related to IELTS writing techniques.

Table 2
Frequency of LBs in the Whole Corpus and in each Band Score.

Frequency of LB tokens in corpus					Frequency of LB types in corpus				
3-word LBs	4-word LBs	5-word LBs	6-word LBs	TOTAL	3-word LBs	4-word LBs	5-word LBs	6-word LBs	TOTAL
3864	966	78	5	4913	291	106	13	1	411
Frequency of LB tokens across band scores									
BAND SCORE			3-word LBs	4-word LBs	5-word LBs	6-word LBs	TOTAL		
Below 5 (excluding 5)			923	213	12	1	1149		
5 to 7 (excluding 7)			1342	383	30	3	1757		
7 to 9			1599	370	36	1	2007		

Table 3

Range and Frequency of Structures in the Whole Corpus and in each Band Score.

Structure	Example from the corpus	Frequency (i.e. number of items) in the whole corpus	Frequency in band scores below 5	Frequency in band scores between 5 and 7	Frequency in band scores between 7 and 9	Total Range (i.e. number of tasks containing LB) in the whole corpus
Other prepositional phrases	<i>As we know</i> , by developing the human knowledge and the technology since 1957 till 2007, the needs of people are changed.	1509	350	540	619	1410
Noun phrase + of	People migrate to other countries for a variety of incentives.	1086	246	411	429	928
Other noun phrases	The lowest percentage in the table is 1% which is for country C.	595	134	242	219	547
Anticipatory it + verb/adjective	It can be seen that almost one-third of income went on food in 1957.	469	104	173	192	440
Others	I strongly believe that the benefits of holding this event outweigh its shortcomings.	262	75	79	108	219
Be + noun/adjective phrase	One of them is related to sand filtering and the other is used to filter and store water in outdoor.	125	37	40	48	105
Prepositional phrase + of	Students mostly go home and sit in front of the TV, browse internet or play video games.	58	13	25	20	61
Passive + prep phrase fragment	More than 40% of household energy is used for heating.	9	5	2	2	7

To design pedagogical activities with the objective of raising the participants' LBs awareness, the researchers predominantly drew from Salazar (2014), which proved to be a valuable resource that contained a wealth of useful learning activities. A compare-and-contrast activity was used for the experimental group which not only engaged the learners in the analysis of the LBs, but also prompted them to generate LBs in their own sentences, hence helping them improve the production of LBs in tandem with their comprehension. Besides, provision of metalinguistic knowledge to make the learners cognizant of the nature of language formulaicity was another technique exploited. Moreover, while noticing plays an indispensable role in the teaching of LBs, the most common and useful LBs may not always be the most salient ones for the learners. Therefore, in line with the other studies which employed underlining, italics, highlighting, and boldface to draw the learners' attention to the less salient bundles (Meunier, 2012; O'Keeffe et al., 2007), the present study underscored an explicit focus on such structures in the activities prepared for the experimental group.

After bringing the instructions to a conclusion and completing the activities, the learners were given one hour to do the posttest. All the participants took the posttest within a maximum period of two weeks after the completion of their instructions.

3.5. Data analysis

After preparing the corpus (see 3.4. above), the transcriptions for every single writing in the three band score categorizations, i.e., 7 to 9, 5 to 7 (excluding 7), and below 5 (excluding 7), were analyzed in terms of their macro-structure (i.e., introduction, body, and conclusion) and were saved on a separate word file. Later, all the separate plain texts of the transcriptions of the three band scores were entered into the concordancing tool. Initially, in order to look for the patterns of lexical bundles in the whole corpus, the Clusters/N-Grams tool of the software was used. Being incentivized by pedagogical applications of the analyses, we started with qualitative exploration, sorting and purging of the bundles and reported them based on their frequency, function, structure, and range in each band score. For the instructional purposes of the second phase, we focused on the most frequent bundles utilized by high-achievers (consequently, disfavoring the semantically and structurally meaningless, incomplete bundles or the essentially too context-specific sequences) (see also Appel & Trofimovich, 2017; Grabowski, 2018). Additionally, the Concordance Plot Tool demonstrated the

Table 4
Range and Frequency of Functions in the Whole Corpus and in each Band Score.

Function	Example from the corpus	LB Token Frequency in the whole corpus	Frequency in band scores below 5	Frequency in band scores 5 to 7	Frequency in band scores 7 to 9	Range in the whole corpus
(academic-) referential (quantification)	<i>We can also see an increase in the proportion of second and third language speakers over the 10-year period.</i>	768	154	329	285	650
Participant-oriented (stance features)	<i>In my opinion, focusing on sports facilities is too narrow an approach and would not have the desired results.</i>	694	198	256	240	674
Text-oriented (framing signals)	<i>First of all, the spread of modern communication devices and people has led to less face-to-face contact.</i>	674	135	250	289	634
Text-oriented (transition signals)	<i>The tables compare the percentages of British university students who spoke different languages in addition to English in two separate years.</i>	568	118	192	258	536
(academic-) referential (location)	<i>You should open our eyes to the fact that holding some International games makes people from all over the world get together as their teams' fans.</i>	461	137	158	166	417
Participant-oriented (engagement features)	<i>As can be seen food and transport are indicated as the most important priority of families to spend money in 1957 and 2007 respectively.</i>	381	97	134	150	363
Text-oriented (resultative signals)	<i>Some people claim that the colors used for painting might damage the environment due to the fact that the chemicals used to decrease the walls emit toxic gases to the air.</i>	258	67	76	115	233
Text-oriented (structuring signals)	<i>In this essay I try to give a balanced view on these inevitable consequences of this modern age tool.</i>	155	38	57	60	148
(academic-) referential (procedure)	<i>First of all, based on the "Rich dad, poor dad" book, I confess that parents have to teach their children how to use their money.</i>	136	28	33	75	124
(academic-) referential (description)	<i>However, the importance of ancient monuments and constructions should not be overlooked.</i>	128	27	49	52	104
(academic-) referential (topic)	<i>As an enthusiastic student I have initiated my education in the field of electrical engineering at an State University.</i>	10	2	2	6	10

location of each LB in each band score separately; it should be mentioned that, as the macro-structural boundaries of each text were determined by the researchers in advance, the concordance results were visually inspected to indicate whether each LB had appeared in the introduction, body, or conclusion sections of the text.

Given our concern with more varied LBs which could be applicable for teaching purposes, we followed the example of some of the previous studies (e.g., Appel & Wood, 2016; Li & Volkov, 2018) and set the N-Gram size for this study to a minimum of 3 words and a maximum of 6 words. It should be noted here that we defined frequency as the number of items in the whole corpus, while range was defined as the number of tasks containing LB in the whole corpus (see Table 3 below). Therefore, for a string of words to be calculated as an LB, the cut-off points for frequency and distribution range were set at 5 times per 200,000 words and 3 tasks, respectively; it means that once a chunk appeared five times or more in at least three texts or more, it was identified as a frequent LB (compare with Biber et al., 2019 that would consider a combination of words which appeared 10 times per 1 million words and at least in 5 texts or more to qualify as an LB in their study).

When the concordance analysis was completed, drawing on Hyland's (2008) modified taxonomy of LBs' functions—derived from Biber (2006), and Biber et al. (2004—as well as Biber et al.'s (1999) revised taxonomy of structures of LBs by Hyland (2008), the researchers manually labeled each LB accordingly; they were chosen since these two taxonomies are best suited to the analysis of written genres (see review section above). To achieve higher reliability, multiple discussion sessions were held by the researchers until they could reach a consensus about the appropriate label for the LBs as well as the accuracy of macro-structural categorizations; plus, the researchers labeled around 10 per cent of the data on different occasions to ensure rater consistency. It should be noted that in cases where more than one function could be attributed to an identified chunk, just one of the discursual functions which was perceived to be the dominant function (e.g., according to the placement of LB in the text) would be reported. For instance, upon exploration of the identified LBs, it was revealed that there existed just three LBs with two distinct functions separated with comma in the essays (namely, *In conclusion, I agree; In conclusion, I believe; and In conclusion, I think*) with the same frequency and range (five instances of each LB in only 15 different essays, namely, each LB in 5 different texts). Given their placement at the conclusion macro-move of the texts and their salient role in bringing the essay to an end, only the dominant perceived function (here, Text-oriented resultative signal) was reported for such LBs, and the Participant-Oriented (stance feature) function of the phrase after comma was not reported; instead, it was reserved for the chunks that predominantly expressed stance and furthered the argumentation (e.g., *Some people believe that; Some people think; I agree with this; etc.*).

For functional analysis of the bundles, the principal classifications of research-oriented, text-oriented, and participant-oriented (Hyland, 2008) were initially adopted (see 2.1., 4.1. and Table 4 for details); nevertheless, in light of the nature of the academic texts in our dataset (IELTS writing essays) which are not—strictly-speaking—research-based academic texts, instead of using the too restrictive label of research-oriented bundles, we opted for the (academic-)referential label, primarily inspired by Biber et al.'s (2004, p. 270) referential markers (while sticking to the sub-categories of *location, procedure, quantification, description, and topic*) (Hyland, 2008). On the other hand, the principal classifications of structural LBs (Biber et al., 1999) included *noun phrase + of, other noun phrases, prepositional phrase + of, other prepositional phrases, passive + prep phrase fragment, anticipatory it + verb/adjective, Be + noun/adjectival phrase, and others*, among others (see 2.1., 4.1. and Table 3 for details).

For the purpose of data analysis of the instructional intervention in the second phase of the study, an experienced IELTS examiner was asked to score the written texts of the participants in the pretest and the posttest. The examiner scored the texts with regard to the four criteria of TA (Task Achievement)/TR (Task Response), CC (Cohesion/Coherence), LR (Lexical Resource), and GRA (Grammatical Range and Accuracy).

After calculating all the participants' scores on the measures of pretest and posttest, a (distribution) normality test was conducted. The results indicated that the data were not normally distributed (see Table 1 in the appendix). Consequently, given the small number of participants and the lack of normality, the non-parametric equivalent of paired-sample *t*-test, i.e., Mann-Whitney *U* Test was run to find the differences between the two control and experimental groups on the measures of TA, CC, LR, GRA, and Overall Band Score for the pretest and the posttest separately. Furthermore, in order to observe the within-group differences between the two control and experimental groups at Time 1 (pretest) and Time 2 (posttest) on the same above-mentioned measures, the non-parametric Wilcoxon Signed-Rank Test was conducted.

4. Results and discussion

4.1. First phase: lexical bundles in IELTS writings across the levels of competence

The initial research questions aimed at identifying and comparing the frequencies, functions, and structures of the LBs. In this study, we had a balanced corpus of 999 IELTS texts across three band scores (333 candidates in each band score), and all the written passages contained approximately 75,000 and 125,000 words for Tasks 1 and 2, respectively (see Table 1). As noted above (3.5), for a string of words to be calculated as an LB, the cut-off points for frequency and distribution range were set at 5 times per 200,000 words and 3 texts, respectively. In terms of frequency, each LB was investigated in the whole corpus and in each band score. LBs with various degrees of frequencies were found in the corpus ranging from the most frequently occurring item which was observed 132 times (*on the other*) in the whole corpus, to the least frequently occurring ones with an overall frequency of only 5 times (e.g., *will be able to*). Also, the LBs with the highest level of frequency were three-word LBs, followed by four-word, five-word, and the rarely occurring expressions of six-word LBs (see Table 2).

As shown in Table 2, the three-word LBs appeared 3864 times in the whole corpus, followed by four-word LBs that occurred 966 times. These numbers dramatically diminished to 78 and 5 for five-word and six-word LBs, respectively. These tokens could be

categorized in terms of 411 LB types (291 three-word; 106 four-word; 13 five-word; and one six-word LB types). The most frequent three-word LB was “on the other,” a structurally incomplete one which was embedded in the most frequently occurring four-word LB of “on the other hand.” Interestingly, “on the other hand the” was the five-word LB with the highest frequency. Finally, the only six-word LBs that were discovered (i.e., “as can be seen from the”) occurred quite infrequently in the whole corpus (predominantly used by more proficient writers). Therefore, a total number of 4913 LBs were found in the corpus, an observation which reinforces the argument in favor of their pervasiveness in discourse and the chunkedness of the language. Thus, IELTS teachers might be recommended to focus more on the most frequent LBs, i.e., three and four-word LBs in their instruction, of course with an eye on the studies which disapprove of (solely) teaching highly frequent LBs. For example, Boers et al. (2010) have argued that medium-frequency LBs should be given more priority provided that they have less chances to be assimilated by learners. In this study, one criterion for inclusion in the instructional phase of the study, apart from LB frequency, was its representation in the written texts of those with the highest band scores.

With reference to the LBs’ distribution across band score categories, three-word bundles had been used 1599 times in the writings of those who gained a band score between 7 and 9, followed by 1342 times in the texts of candidates with a band score between 5 and 7, and finally only 923 instances were observed in band scores below 5. Nevertheless, the pattern was not the same for four-word LBs. Candidates with an overall band score of 5–7, with 383 instances, had used four-word bundles more frequently than those with band scores below 5 (213 instances) and between 7 and 9 (370 instances) (see Table 2).

Generally, regardless of the number of words in each LB, band scores between 7 and 9 had by far the largest proportion of LBs totaling 2007 in the whole corpus, followed by band scores between 5 and 7 and below 5 which had a total frequency of 1757 and 1149, respectively (Table 2). Therefore, these observations corroborate the assumptions underlying the first phase about the potential differences between the LBs in the three band score categories with respect to their frequencies. The pedagogical implication of this finding is that to help students achieve better overall band scores, and particularly CC and LR scores (see the findings below), IELTS teachers should encourage them to use more LBs in their written texts. This might have resulted in and/or contributed to the observation that the texts of more competent writers were generally longer than those of their less competent peers (compare the length of both Task1 and Task 2 texts across the three groups in Table 1 above). However, frequency should not be exclusively concentrated on to the detriment of LB types both structure- and function-wise. Upon detailed analysis of the items, it was obvious that some LBs were much more frequently used by candidates with the lowest band scores (e.g., “a lot of” and “all of the”). Hence, this finding indicates that not always the most frequent LBs are pedagogically useful and careful screening should take place prior to the selection of LBs.

With regards to the frequency and range of IELTS lexical bundle structures across the levels of writing competence, as illustrated in Table 3, “other prepositional phrases” had the highest frequency and range of occurrence in the whole corpus. “noun phrase + of” with the total occurrence of 1086 times in the whole corpus and the range of appearance across 928 texts was ranked the second. Next, it can be observed that the rate of both frequency and range gradually diminished from 595 to 547 for the structure “other noun phrases” to the least frequent structure, i.e., “Passive + prep phrase fragment” with an overall frequency of merely 9 and a meagre range of 7 calling for less prioritization compared to the more frequent structure types. In between, we can notice “anticipatory it + verb/adjective,” “others,” “be + noun/adjective phrase,” and “prepositional phrase + of” ranging from the wide gamut of 469 and 440 to 58 and 61 for frequency and range, respectively.

It should be noted that since the adopted structural and functional taxonomies (Hyland, 2008) had been originally proposed for four-word LBs only, they accounted much better for four-word LBs compared to the three-word and five-word LBs (hence, the discrepancy between the observed tokens and types of LBs and their structural and functional categorizations in Tables 3 and 4).

Regarding the functions of LBs in IELTS writing texts, as can be seen from Table 4, text-oriented bundles (including all its subdivisions) are predominantly higher in terms of the frequency and range of function types in comparison with (academic-)referential and participant-oriented bundles. A total of 1655 occurrences were observed for text-oriented bundles, with framing signals and transition signals at the top (674 and 568 occurrences respectively) followed by resultative and structuring signals (258 and 155 occurrences for each respectively). With 1503 and 1075 instances, (academic-)referential and participant oriented are in the second and third place in terms of overall frequency in the whole corpus. Based on their higher frequencies and ranges, it is suggested that higher priorities be placed on the teaching of text-oriented, (academic-)referential, and participant-oriented bundles, in that order. Since textual organization and conveying the underlying message are highly underscored in IELTS writing, we can see that text-oriented bundles have been more frequently used to achieve this goal. Framing signals, whose purpose is to situate arguments within a text, are more likely to occur in Task 2, due to the fact that Task 2 rubrics are more argumentative in nature. Not only framing signals, but also transition signals—whose goal is to connect elements in a text by adding to them or contrasting them—have been among the most frequent text-oriented bundles used. On the contrary, the reason for the dearth of structuring signals in the corpus can be attributed to the short length of IELTS writing tasks in which the opportunity to direct the reader elsewhere in the text may not arise. Not surprisingly, the same pattern can be observed for the range of function types across the writing answer sheets.

Notably, with the occurrence of text-oriented bundles (including the subdivisions) across 1551 texts, they are at the top in terms of range. With the rate of 1305 and 1037, (academic-)referential and participant-oriented bundles were in the second and the last place when it came to range. Similarly, framing signals and transition signals (634 and 536 respectively) were more prevalent across candidates’ texts than resultative and structuring signals (233 and 148 respectively). Although they are the subcategories of the same function type, there is an enormous gap between both the frequency and range of stance features and engagement features under the rubric of participant-oriented bundles. Additionally, regardless of (academic-)referential bundles of quantification and location (which were used with noticeable frequencies) the other three subcategories (i.e., procedure, description, and topic) are at the bottom of the table—with the excessively sparing use of topic bundles with a mere frequency and range of 10. The rationale for the high frequency of

quantification and location bundles appears to be that in Task 1, the candidates are more likely to be required to write a report on some given figure, graph, or diagram; therefore, familiarizing them with LBs to express quantification would be pedagogically advisable. Although there are task rubrics that present a process and ask the candidates to elaborate on the different stages in the process requiring them to use procedure LBs, such LBs were less frequent in our corpus. The justification is that most of the Task 1 answer

Table 5

Text-part specific bundles prevalent in one or two text parts of IELTS writing texts.

Lexical Bundles	Example from the corpus	Introduction	Body	Conclusion
In this essay	In this essay , I shall appeal to crime statistics to argue that a person is at a far greater risk being shot if there is a gun in a household. (T2-B9-Intro.)	*		
In recent years	In recent years we have seen great improvement in TV programs and fantastic visual achievement in TV industry. (T2-B7.5-Intro.)	*		
Some people think	Although some people think educational centers should filterize the students by which ability they have, ... (T2-B4-Intro.)	*		
I strongly believe that	I strongly believe that the allowance should be paid to children in longer periods like on a weekly basis. (T2-B7-Intro.)	*		
One of the most important	Saving languages and cultures have become one of the most important issues if not the most important issue of human studies. (T2-B6.5-Intro.)	*		
The given table	[t]he given table shows that how much money is spent by students annually in universities of three different countries. (T1-B6-Beginning)	*		
In the following	In the following , we are going to give you some information about the average annual spending of universities in country A, country B and country C. (T1-B5.5-Beginning)	*		
The chart below	The chart below illustrates information regarding materials used for the purpose of packaging in three parts of the world in 2002. (T1-B7-Beginning)	*		
I personally believe	I personally believe that paying money regularly to kids will be beneficial for their future. (T2-B7.5-Intro.)	*		
The chart illustrates	The chart illustrates the amount of money spent on five consumer goods (cars, computers, books, perfume and cameras) in France and the UK in 2010. (T1-B9-Beginning)	*		
As can be seen from the	As can be seen from the pie charts, Mexico received a high percentage of exports which belong to USA with 87 percent. (T1-B5.5-Beginning)	*		
Some people claim that	It is true that classified classes could have some advantages as some people claim that . (T2-B7-Intro.)	*		
The diagram illustrates	The diagram illustrates the amount of consumed materials for packaging per each person in USA Japan and Europe in 2002. (T1-B7-Beginning)	*		
The table shows	The table shows information about University Students' expenditure in three different countries. (T1-B6-Beginning)	*		
First of all	First of all , I will consider how a person's choice of clothes reveals their culture. (T2-B7-Intro)	*	*	
Some people believe that	Some people believe that they shouldn't squander their time on language learning. (T2-B3.5-Body)	*	*	
It is believed that	It is believed that paying for old people by authorities is just a way of losing the capital and cost of living. (T2-B5.5-Body)	*	*	
According to the chart	According to the chart , glass has been another material consuming approximately in each country with high rates. (T1-B5.5-Middle)	*	*	
It is shown	It is shown that Europe had the most exports in Chile with 26% whereas ... (T1-B5.5-Middle)	*	*	
To begin with	To begin with , everyone needs to learn how to manage their financial resources (T2-B7.5-Body)		*	
On the contrary	As grown up/adults we are not allowed to prevent the youngsters from experiencing these sports. On the contrary , it would be better to help them to play in safer conditions. (T2-B7 -Body)		*	
As a result	However, for other two countries would rise up to 21 percent for Chile and 41 percent for Argentina. As a result , USA has more shares of these three countries exports and Europe and other Latin American countries have less amount of imports from these three countries. (T1-B5.5-Middle)		*	*
On the other hand	For example, there is a chatting language which is produced by our smart phones. On the other hand , some languages are vanishing these days. (T2-B5-Body)		*	*
Last but not least	And last but not least is that feeling the risk and excitement in the mentioned sports give the youth motivation to continue living a world full of stress. (T2-B7-Body)		*	*
To sum up	To sum up , from all charts it is obvious that the major consumer of these three countries products was USA. (T1-B7-End)			*
All in all	All in all , to put the issue in contrast, I am on the thought that it is better for all of us to choose what we want to hear or see through the internet, (T2-B8-Conclusion)			*
In conclusion	In conclusion , giving money to children regularly has a lot of advantages (T2-B7-Conclusion)			*
Can be concluded	By and large, it can be concluded that the family budget is going to be dedicates to other more than the past years. (T1-B4.5-End)			*
In conclusion I agree	In conclusion, I agree that any child can be taught particular skills, but ... (T2-B9-Conclusion)			*
In conclusion I believe	In conclusion, I believe that people are suffering from the harmful impact of the social networking systems rather than the positive ones. ... (T2-B7.5-Conclusion)			*
In conclusion I think	In conclusion, I think this is a very good idea, and I hope this program will be put into action for high schools/colleges shortly. ... (T2-B8-Conclusion)			*

Note: *T* refers to Task Number → T1: Task 1 & T2: Task 2).

B refers to the assigned Band Score → e.g., B8: Band 8.

sheets transcribed were more of charts and tables and less of process description types. Finally, the unexpectedly low frequency of topic LBs in the corpus can be attributed to their function best characterizing research articles (Cortes, 2004; Hyland, 2008) and not necessarily the genres of IELTS writings. Thus, one implication is that IELTS teachers do not need to exclusively and/or predominantly focus on such LB types inasmuch as the chances are slim the candidates would need them in their writing tasks.

Furthermore, similar to the high-achievers' superior variety and frequency of employed LB structures, the most proficient applicants had the most functionally diverse texts as well. From among the 4233 occurrences of different types of functions in the whole corpus, the highest band scores had the lion's share of 1696 occurrences. Except for (academic-)referential bundles of quantification and participant-oriented bundles of stance features, the writings of the third band score category (i.e., between 7 and 9) outnumbered those with lower band scores. Not only were candidates' texts with the highest band scores more functionally varied, but they also had the largest number of function ranges across the texts (Table 4).

Similarly, the more frequent stance features were participant-oriented bundles as compared to engagement features; one explanation is that IELTS candidates are more likely to impart their own personal attitudes and arguments (as they are required to do so most of the times in Task 2) than to directly address the reader in their texts, which is more typical of letter writing in Task 1 of IELTS General (not Academic) Module. While our findings are characterized by notable instances of participant-oriented bundles, it is noted in the literature that L2 users of English use such bundles in their written texts much less than L1 users of English, notwithstanding the indispensable role the participant-oriented bundles play in academic discourse (Salazar, 2014). Hyland (2012) attributes this lack of participant-oriented bundles in the texts of L2 users of English, especially Asian EFL learners, to their cultural predilections for a noninterventionist stance in argumentation. The use of hedging devices and depersonalized stance features are an indivisible part of academic writing (Salazar, 2014). As suggested by Hyland (1996), participant-oriented bundles are used to protect academic writers against unjustified generalizations and enable them to stay on the side of caution when there are no categorical findings. Based on these arguments and the findings of the corpus analysis in the present study, therefore, it can be recommended that participant-oriented bundles be given more attention and time in EFL (academic) contexts.

To address the third research question, we utilized the concordancing plot tool to determine the location in the text for each LB with respect to its corresponding structure and function. Essentially, the majority of the LBs have been found to be rather equally dispersed in different sections, i.e., introduction, body, and conclusion; however, there are a few instances that occurred only in specific parts like either in the introduction, body, or conclusion or a combination of two of them. The following table (Table 5) shows some of these LBs that were typical of only one or two text parts (for a more elaborate list of such location-specific resources see Saadatara, 2020). It should be noted that the introductory and concluding sub-sections were noticeably shorter than the body part of both tasks (with the beginning and ending of Task 1 texts generally even shorter than the introduction and conclusion of Task 2 essays).

4.2. Second phase: LB instruction and improvement of IELTS writing

To see the effects of instruction of LBs on the participants' overall band scores as well as CC, LR, and GRA measures, their scores in the pretest and posttest were analyzed. First of all, to see the between-group differences on the above-mentioned scoring criteria, a Mann-Whitney *U* Test was conducted whose results are shown in Table 6. All the participants' scores on TA, CC, LR, and GRA together with overall band scores in both Tasks 1 and 2 in the pretest and posttest were compared separately. For instance, the overall band scores of Task 1 in pretest were compared between the two experimental and control groups. As shown in the table, the statistical significance is not below the significance level of 0.05 for any of the scoring criteria in Task 1 or 2, neither in the pretest nor in the posttest; therefore, we can conclude that the distribution of all the measures is quite similar across the participants in both experimental and control groups as assessed in the pretest and posttest.

To get a more accurate picture of possible differences, Wilcoxon Signed-Rank Test was conducted with the purpose of within-group comparisons and the results are presented in Tables 7 and 8. As illustrated in Table 7, there is a statistically significant difference between the performance of the control group participants in terms of overall band score in Task 1 when comparing their pretest and posttest scores (Sig 0.004). In contrast, there is no statistically significant difference for the same scoring criterion in the performance of the participants in Task 2 (Sig 0.060). The reason for such a lack of improvement on LR and GRA in Task 2 appears to be the absence of lexical and grammatical coverage in the treatment for the control group. Also, the participants frequently reported that they found task 2 more difficult and challenging. Furthermore, this lack of efficiency in Task 2 can be linked to the observation made earlier as to

Table 6
Mann-whitney *U* test results (between-group comparison).

Scoring criteria	Task	Pretest	Posttest
Overall band score	1	.387	.237
	2	.423	.245
Task Achievement	1	.319	.308
	2	.401	.437
Cohesion and Coherence	1	.431	.074
	2	.505	.067
Lexical Resource	1	.325	.196
	2	.526	.125
Grammatical Range and Accuracy	1	.418	.324
	2	.463	.363

Table 7

Wilcoxon signed-rank test results for the control group (within-group comparison).

Scoring criteria	Task	Sig.
Overall band score	1	.004
	2	.060
Task Achievement	1	.004
	2	.009
Cohesion and Coherence	1	.026
	2	.018
Lexical Resource	1	.004
	2	.199
Grammatical Range and Accuracy	1	.017
	2	.132

Table 8

Wilcoxon signed-rank test results for the experimental group (within-group comparison).

Scoring criteria	Task	Sig.
Overall band score	1	.000
	2	.001
Task Achievement	1	.006
	2	.001
Cohesion and Coherence	1	.000
	2	.000
Lexical Resource	1	.000
	2	.001
Grammatical Range and Accuracy	1	.005
	2	.011

the lack of participant-oriented bundles in the writing performance of the participants (which substantiated their difficulty with expressing their personal opinions and arguments).

Generally, it is clearly understood from the statistics that, despite lacking an explicit functional focus on LBs, the IELTS exam instruction provided for the control group was very effective and helped the participants to achieve relatively better post-test scores in all the four criteria in both tasks except for lexical resource (LR) and grammatical accuracy (GRA) in Task 2, for which no significant difference is witnessed. Furthermore, it seems that the instruction for the control group has been effective in terms of task achievement. During the treatment, many of the key conventions of IELTS writing were covered and elaborated on, including points on how to enhance scores on the criterion of task achievement and explaining the expectations of IELTS examiners. Even though LBs were not the primary focus of instruction for the control group, the participants showed great improvements on scores of cohesion and coherence and lexical resource. The reason for such unexpected results could be the fact that, although they were not exclusively elaborated upon and their functions were not explicated as in the instructions for the experimental group, some linking words were touched upon in each session.

The results of the Wilcoxon test for the experimental group are provided in Table 8. It can be deduced from the table that the pedagogical intervention, the activities, and the enhanced input presented to the participants have had a significant impact on their IELTS writing performance with respect to all the scoring criteria. As can be seen, the most statistically significant difference between the scores of the pretest and posttest can be found in the scores for cohesion and coherence, lexical resource, and overall band score.

Even though LB instruction has had positive impacts on all the scores in the experimental groups, its impacts on CC and LR are the most noticeable. Additionally, in contrast to Ghafarsamar et al. (2018) and Eidian et al. (2013), who found no statistically significant difference with respect to GRA after instruction on LBs (a finding attributed to the short duration of their interventions), we can notice meaningful improvements even on that measure in the present study. The authors of the afore-mentioned studies believe that more time is needed for participants to take in the grammar instruction through such a lexical collocational teaching approach. Therefore, we can contend that an explicit functional approach to the teaching of LBs which includes awareness-raising and metalinguistic information (along with other instructions provided during the course) can be pedagogically effective especially for cohesion and coherence (CC) and lexical resource (LR).

Our findings are in line with Wray's (2002) argument that using LBs can help language users identify with a disciplinary community and reduce the burdens on production; likewise, we observed that the written texts of the participants in the experimental group were more typical of IELTS writing in terms of their LBs. In fact, it was noticed that LBs were more appropriately used in the posttest texts with regard to the functions they serve and the place they should appear (compare also with Cai, 2016). For instance, almost all the experimental group participants had used LBs typical of the conclusion paragraph in Task 2, like *in conclusion*, *to sum up*, *to recapitulate*, among others.

Cortes (2004) has suggested that mere exposure to LBs in reading passages alone could not result in learners' active use of such structures. She further posits that unless a formal and explicit instructional component is involved, learners might not be able to actively use LBs in their text production. Therefore, the findings of the present study are aligned with and substantiate Cortes' (2004,

2006) contentions concerning the need for explicit teaching of LBs and its superiority over exposure alone (which was somehow the case with our control group). On the contrary, the present study might question Cortes' (2006) emphasis that for the acquisition and active use of LBs to take effect, learners need to be exposed to them and use them more as a long-term project (compare also with Li & Schmitt, 2009; Vo, 2019).

5. Conclusion and implications

The present study consisted of two phases, a corpus analysis phase, with a focus on the frequency, structure, function, range, and location of LBs in the writing texts of IELTS candidates, and an instructional phase where the effects of teaching of select LBs used by high achievers in the previous phase were investigated. One major outcome of the first exploratory phase of the study is its reiterating the assumption that (academic) IELTS writing module, as an instance of high-stakes writing examination genres, might be argued to be distinct from other written genres (Beck & Jeffery, 2007; Hyland, 2003; Pearson, 2019; Staples et al., 2013; Swales, 2004); in particular, we focused on the LBs (and their idiosyncratic structure, frequency, function, and distribution) and suggested that, in spite of the similarities with other comparable genres, IELTS writing texts differentially deployed these resources both quantitatively and qualitatively.

As the analysis of the LBs in the whole corpus and in each separate band score category revealed, three-word LBs were the most frequent ones both in the whole corpus and in the texts written by the most proficient candidates (Table 2). It was further shown that as the LBs' length increased, their frequencies diminished. Overall, the higher the writing proficiency of the candidates, the higher the frequency and range of LBs used in their writings (Tables 2–5). It was not only in terms of frequency and range of LBs that the proficient candidates' texts were richer, but they also surpassed with regards to structural and functional variety and range. Considering their locations in the texts, the bulk of LBs had been used in all parts of the texts. However, there were instances that were only found in either introduction, body, or conclusion (Table 5). Consequently, IELTS teachers (and materials developers—not to mention the learners themselves—as instructed in our study) can identify such location-specific LBs using concordancing tools and elaborate on the structures and the functions they serve in those parts of the texts (for a more elaborate list of such resources see Saadatara, 2020). Overall, Meunier (2012) makes a distinction between authentic language as evidenced by corpus and the content learners may encounter in commercial textbooks (see also Talebzadeh & Khazraie, 2021) in terms of their frequency and didactic value; therefore, corpora and corpus tools can greatly help all stake-holders (including IELTS and other high-stakes exam teachers, teachers of General English, EAP/ESP teachers, high-stakes examination candidates, material writers, and textbook evaluators) in getting and providing snapshots of more natural language in use and the ensuing consequences.

The findings of phase two highlight the effectiveness of explicit IELTS instruction, in general, and teaching selected lexical bundles, in particular. Generally, the effects of instruction of LBs to the experimental group were found to be much more noticeable across the board, although the control group showed signs of improvement on Task 1 in the posttest thanks to the beneficial effect of IELTS instruction. All in all, it can be argued that the teaching of LBs with the approach adopted and the activities designed for the present study have had perceptible impacts on scores of CC and LR. Therefore, notwithstanding the limitations of our study including the rather limited number of participants, we can conclude that an explicit functional awareness-raising approach while providing metalinguistic information on LBs can be pedagogically beneficial. No doubt such instructional practices, if implemented judiciously, can utilize the exploratory and pedagogical findings of LB studies (like the current one) as a point of departure in sequencing and prioritizing a (syllabi) content, designing various courses, and even evaluating their effectiveness (see Harwood, 2010; Macalister & Nation, 2020).

It appears that LBs can still remain a thriving and promising research area given the constraints and untapped potentials of LBs' investigation within the genres of high-stakes examinations. In particular, we should invite caution in making generalizations based solely on our findings inasmuch as our dataset was limited to mock IELTS writing tasks; although they were taken from a credible IELTS teaching center with a high success rate and dependable scores reflecting the applicants' scores on actual IELTS session, future investigations can examine LBs either in mock exam data from a wealth of other sources/centers or in substantial numbers of general and/or academic tasks written for actual IELTS. Due caution should be exercised in overgeneralizing current findings to similar writing tasks, as a differential focus on each task of IELTS (especially Task 1 in the General and Academic modules), which can be different in organization, purpose, and word limit, might render comparatively different results and implications. Moreover, while we had to delimit our study to the exploration and instruction of LBs in the writing section of IELTS, other test modules (such as listening and speaking) are equally worth examination and the stake-holders can benefit from their identification and teaching.

Another research area which appears to deserve closer scrutiny in further studies of LBs and high-stakes exams instruction would be comparing the efficacy of a multitude of pedagogical techniques, strategies, and technological affordances (e.g. user-friendly online and offline concordance tools for both teachers and students); in particular, studies which can recruit a much larger and more diverse sample size while employing varied (i.e. qualitative, quantitative, and mixed) data collection and analysis methodologies are more likely to shed a better light on many of the unexamined dynamics of LB instruction. Finally, rather than being limited to IELTS and/or TOEFL, there are other national and international examinations whose results would affect the lives of countless individuals worldwide; they deserve to be adequately researched, too.

Acknowledgement

We would like to express our gratitude to Dr Zak Lancaster, associate editor of *Journal of English for Academic Purposes*, and the anonymous reviewers for their insightful suggestions and comments. We also sincerely thank the study participants and the IELTS

center officials who went out of their way and provided us with the data we needed.

Appendix

Table 1
Results of Normality Test

Dependent Variables	Test	task	group	Sig.
Overall Band Score	Pretest	1	Control	.020
		2	Control	.093
		1	Experimental	.007
		2	Experimental	.000
	Posttest	1	Control	.042
		2	Control	.148
		1	Experimental	.011
		2	Experimental	.010
Task Achievement/Task Response	Pretest	1	Control	.176
		2	Control	.157
		1	Experimental	.094
		2	Experimental	.000
	Posttest	1	Control	.138
		2	Control	.158
		1	Experimental	.011
		2	Experimental	.070
Cohesion and Coherence	Pretest	1	Control	.039
		2	Control	.033
		1	Experimental	.004
		2	Experimental	.001
	Posttest	1	Control	.190
		2	Control	.200
		1	Experimental	.052
		2	Experimental	.038
Lexical Resource	Pretest	1	Control	.037
		2	Control	.059
		1	Experimental	.035
		2	Experimental	.002
	Posttest	1	Control	.033
		2	Control	.081
		1	Experimental	.031
		2	Experimental	.021
Grammatical Range and Accuracy	Pretest	1	Control	.100
		2	Control	.088
		1	Experimental	.005
		2	Experimental	.006
	Posttest	1	Control	.090
		2	Control	.146
		1	Experimental	.024
		2	Experimental	.046

Appendix 2. the first 50 most frequent 3-word lexical bundles overlapping with four-, five, and six-word bundles

	3-word bundles (frequency ≥ 10)	3-word bundles (frequency ≥ 10)	4-word bundles	4-word bundles	4-word bundles	4-word bundles	5-word bundles	5-word bundles	6-word bundles
1	On the other		On the other hand	On the other side			On the other hand the		
2	A lot of		A lot of money	A lot of time					
3	One of the		Is one of the	One of the most	One of the best	One of them is	One of the most important		
4	In my opinion		In my opinion it	In my opinion the					
5	The amount of		The amount of money						
6	It can be		It can be seen				It can be seen that		

(continued on next page)

(continued)

	3-word bundles (frequency ≥ 10)	3-word bundles (frequency ≥ 10)	4-word bundles	4-word bundles	4-word bundles	4-word bundles	5-word bundles	5-word bundles	6-word bundles
7	As a result		As a result of						
8	The fact that		To the fact that	Despite the fact that	The fact that the		Due to the fact that		
9	Around the world		All around the world	People around the world					
10	It would be		It would be a	It would be better					
11	It is clear		It is clear that				It is clear from the		
12	Be able to		To be able to	Not be able to	Will be able to				
13	We can see		As we can see	We can see that	We can see the	We can see a			
14	Due to the						Due to the fact that		
15	Point of view		My point of view				In my point of view	From my point of view	
16	Most of the		Most of the time						
17	The most important						One of the most important		
18	Would like to								
19			I am writing to						I am writing to inform you
20	People believe that		Some People believe that	Believe that it is			People believe that it is		
21	Can be seen		As Can be seen	It Can be seen	Can be seen that	Can be seen from	As can be seen from	It can be seen that	As can be seen from the
22	It is obvious		It is obvious that				It is obvious that the		
23	In comparison with		In comparison with other						
24	According to the		According to the chart						
25	At the end		And at the end	At the end of					
26	I agree with		I agree with this						
27	Learn how to		to Learn how to						
28	Of the most		One Of the most				One of the most important		
29	Over the world	All over the	All over the world						
30	Some people believe		Some people believe that						
31	To the fact that		Due to the fact that						
32	As soon as		As soon as possible						
33	It seems that		It seems that the						
34	Believe that it	Believe that the	Believe that it is				People believe that it is		
35	I think it		I think it is						
36	The end of		The end of the	At the end of					
37	All around the		All around the world						
38			I am writing this				I am writing this letter		I am writing this letter to
39	In touch with		Keep in touch with	Be in touch with					

(continued on next page)

(continued)

	3-word bundles (frequency ≥ 10)	3-word bundles (frequency ≥ 10)	4-word bundles	4-word bundles	4-word bundles	4-word bundles	5-word bundles	5-word bundles	6-word bundles
40	Looking forward to	Look forward to	I look forward to				I am looking forward to		Look forward to hearing from you
41	To know about		To know about						
42	Be seen that		Can be seen that				It can be seen that		
43	The highest amount		The highest amount of						
44	The highest percentage		The highest percentage of						
45	The other side		On the other side						
46	Huge amount of		Huge amount of money	A huge amount of					
47	As we can		As we can see						
48	Strongly believe that		I strongly believe that						
49	To learn how		To learn how to						
50	Vast majority of		The vast majority of						

References

- Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31(2), 81–92. <https://doi.org/10.1016/j.esp.2011.08.004>
- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 101–122). Oxford: Oxford University Press.
- Anthony, L. (2006). *AntConc 3.1.302 (windows)*. Freeware: Waseda University.
- Appel, R., & Trofimovich, P. (2017). Transitional probability predicts native and non-native use of formulaic sequences. *International Journal of Applied Linguistics*, 27(1), 24–43. <https://doi.org/10.1111/ijal.12100>
- Appel, R., & Wood, D. (2016). Recurrent word combinations in EAP test-taker writing: Differences between high- and low-proficiency levels. *Language Assessment Quarterly*, 13(1), 55–71. <https://doi.org/10.1080/15434303.2015.1126718>
- Arefasadr. (2022). *Exploring the challenges and feasibility of incorporating world Englishes in assessing language practices*. Unpublished PhD Dissertation: Kharazmi University.
- Banerjee, S., & Pedersen, T. (2003). Extended gloss overlap as a measure of semantic relatedness. In *Proc. Of IJCAI-03* (pp. 805–810).
- Beck, S. W., & Jeffery, J. V. (2007). Genres of high-stakes writing assessments and the construct of writing competence. *Assessing Writing*, 12(1), 60–79. <https://doi.org/10.1016/j.asw.2007.05.001>
- Bestgen, Y. (2017). Beyond single-word measures: L2 writing assessment, lexical richness and formulaic competence. *System*, 69, 65–78.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers* (Amsterdam: Benjamins).
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263–286. <https://doi.org/10.1016/j.esp.2006.08.003>
- Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language and Computers*, 26, 181–190.
- Biber, D., & Conrad, S. (2001). *Register variation: A corpus approach. The handbook of discourse analysis*.
- Biber, D., & Conrad, S. (2019). *Register, genre, and style* (2nd ed.). Cambridge University Press.
- Biber, D., Conrad, S., & Cortes, V. (2003). Towards a taxonomy of lexical bundles in speech and writing. In A. Wilson, P. Rayson, & T. McEnery (Eds.), *Corpus linguistics by the lute: A festschrift for Geoffrey Leech* (pp. 71–92) (Frankfurt/Main: 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., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Grammar of spoken and written English*. Harlow: Longman.
- Boers, F., Deconinck, J., & Lindstromberg, S. (2010). Choosing motivated chunks for teaching. In S. D. Knop, F. Boers, & A. D. Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 239–256). Berlin: De Gruyter Mouton.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a Lexical Approach to the test. *Language Teaching Research*, 10(3), 245–261. <https://doi.org/10.1191/1362168806lr1950a>
- Brazil, D. (1995). *A grammar of speech*. Oxford: Oxford University Press.
- Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, 5(5), 31–64.
- Cai, (L.) J. (2016). An exploratory study on an integrated genre-based approach for the instruction of academic lexical phrases. *Journal of English for Academic Purposes*, 24, 58–74. <https://doi.org/10.1016/j.jeap.2016.09.002>
- Cambridge Assessment English. (2022). *Comparing scores to IELTS*. Retrieved (September 16, 2022) from <https://www.cambridgeenglish.org/Images/461626-cambridge-english-qualifications-comparing-scores-to-ielts.pdf>.
- Cambridge University Press. (2019). *IELTS practice tests*. Cambridge: Cambridge University Press.
- Chen, Y., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, 37(6), 849–880. <https://doi.org/10.1093/applin/amu065>
- 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>
- Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17(4), 391–406. <https://doi.org/10.1016/j.linged.2007.02.001>
- De Cock, S. (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair, & M. Hundt (Eds.), *Corpus linguistics and linguistic theory* (pp. 51–68). Amsterdam: Rodopi.

- Eckert, P., & McConnell-Ginet, S. (1998). Communities of practice: Where language, gender, and power all live. In J. Coates (Ed.), *language and gender: A reader*. Oxford: Blackwell.
- Eidians, F., Gorjian, B., & Aghvami, F. (2013). The impact of lexical collocation instruction on developing writing skill among EFL learners. *International Journal of Learning and Applied Linguistics World*, 4(3), 273–283.
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic Language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *Tesol Quarterly*, 42(3), 375–396. <https://doi.org/10.1002/j.1545-7249.2008.tb00137.x>
- Forsberg, F. (2010). Using conventional sequences in L2 French. *IRAL - International Review of Applied Linguistics in Language Teaching*, 48(1). <https://doi.org/10.1515/iral.2010.002>
- Ghafarsamar, R., Shokrpour, N., & Nasiri, E. (2018). The applicability of teaching lexical bundles on medical students' writing proficiency in an EFL context. *Teaching English Language*, 12(2), 27–44.
- Goźdz-Roszkowski, S. (2011). *Patterns of linguistic variation in American legal English: A corpus-based study*. Frankfurt am Main: Lang.
- Grabowski, L. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, 38, 23–33. <https://doi.org/10.1016/j.esp.2014.10.004>
- Grabowski, L. (2018). Fine-tuning lexical bundles: A methodological reflection in the context of describing drug-drug interactions. In J. Kopaczkyk, & J. Tyrkkö (Eds.), *Applications of pattern-driven methods in corpus linguistics* (pp. 57–80). Amsterdam: John Benjamins.
- Harwood, N. (Ed.). (2010). *English language teaching materials: Theory and practice*. Cambridge and New York: Cambridge University Press.
- Hyland, K. (1996). Nurturing hedges in the ESP curriculum. *System*, 24(4), 477–490. [https://doi.org/10.1016/s0346-251x\(96\)00043-7](https://doi.org/10.1016/s0346-251x(96)00043-7)
- Hyland, K. (2003). Dissertation acknowledgements: The anatomy of a Cinderella genre. *Written Communication*, 20(3), 242–268.
- Hyland, K. (2008). 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. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, 32, 150–169. <https://doi.org/10.1017/s0267190512000037>
- Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(3), 433–445. <https://doi.org/10.1111/j.1540-4781.2007.00589.x>
- Johnston, K. M. (2017). *Lexical bundles in applied linguistics and literature writing: A comparison of intermediate English learners and professionals*. Unpublished Ph.D. Dissertation, Portland State University.
- Karabacak, E., & Qin, J. (2013). Comparison of lexical bundles used by Turkish, Chinese, and American university students. *Procedia - Social and Behavioral Sciences*, 70, 622–628.
- Leech, G. (2000). Grammars of spoken English: New outcomes of corpus-oriented research. *Language Learning*, 50(4), 675–724. <https://doi.org/10.1111/0023-8333.00143>
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Hove, England: Language Teaching Publications.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85–102.
- Li, Z., & Volkov, A. (2018). To whom it may concern: A study on the use of lexical bundles in email writing tasks in an English proficiency test. *TESL Canada Journal*. <https://doi.org/10.18806/tesl.v34i3.1268>
- Macalister, J., & Nation, P. I. S. (2020). *Language curriculum design*. New York and London: Routledge.
- Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111–129. <https://doi.org/10.1017/S0267190512000128>
- Moon, R. (1998). *Fixed expressions and imprints in English*. Oxford: Oxford University Press.
- Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom Foreign Language learning. *Language Learning*, 48(3), 323–364. <https://doi.org/10.1111/0023-8333.00045>
- Nattinger, J. R., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nekrasova, T. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning*, 59(3), 647–686. <https://doi.org/10.1111/j.1467-9922.2009.00520.x>
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: CUP.
- Palmer, H. E. (1933). *Second interim report on English collocations*. Submitted to the tenth annual conference of English teachers. Tokyo: Kaitakusha.
- Pearson, W. S. (2019). Critical perspectives on the IELTS test. *ELT Journal*, 73(2), 197–206. <https://doi.org/10.1093/elt/ccz006>
- Reppen, R. (2018). Teaching lexical bundles. In E. Hinkel (Ed.), *Teaching essential units of language*. Routledge.
- Saadatara, A. (2020). *Lexical bundles across band scores of the writing component of IELTS mock exams*. Unpublished MA Thesis: Tarbiat Modares University.
- Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching*. Amsterdam: John Benjamins Publishing Company.
- Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action - an introduction. In N. Schmitt (Ed.), *Formulaic sequences acquisition, processing, and use* (pp. 1–22). Amsterdam: Philadelphia: John Benjamins Pub.
- Shirazizadeh, M., & Amirfazlian, R. (2021). Lexical bundles in theses, articles and textbooks of applied linguistics: Investigating intradisciplinary uniformity and variation. *Journal of English for Academic Purposes*, 49, 1–3. <https://doi.org/10.1016/j.jeap.2020.100946>
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487–512. <https://doi.org/10.1093/applin/amp058>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Spötl, C., & McCarthy, M. (2003). Formulaic utterances in the multi-lingual context. In Hufeisen (Ed.), *The multilingual lexicon, jasone cenoz, ulrike jessner & britta* (pp. 133–151). Dordrecht: Kluwer.
- Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, 12(3), 214–225. <https://doi.org/10.1016/j.jeap.2013.05.002>
- Strunkyte, G., & Jurkūnaite, E. (2008). *Written academic discourse: Lexical bundles in humanities and natural sciences*. Vilnius: Vilnius University. Unpublished bachelor's thesis.
- Stubbs, M., & Barth, I. (2003). Using recurrent phrases as text-type discriminators: A quantitative method and some findings. *Functions of Language*, 10(1), 65–108. <https://doi.org/10.1075/fol.10.1.04stu>
- Swales, J. (2004). *Research genres: Exploration and applications*. Cambridge: Cambridge University Press.
- Syder, F. H., & Pawley, A. (2014). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In *Language and communication* (pp. 203–239). Routledge.
- Talebzadeh, H., & Khazraie, M. (2021). 'Ignoring the elephant in the room': (Under-)representation of impoliteness phenomenon in popular ELT textbooks. *Language Teaching Research*. <https://doi.org/10.1177/13621688211029028>, 0(0).
- Tremblay, A., Derwing, B., & Libben, G. (2007). Are lexical bundles stored and processed as single units?. In *Proceedings of 23rd northwest linguistics conference, victoria BC CDA, (feb. 17-19, 2007)* (pp. 258–279).
- Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, 61(2), 569–613. <https://doi.org/10.1111/j.1467-9922.2010.00622.x>
- Vo, S. (2019). Use of lexical features in non-native academic writing. *Journal of Second Language Writing*, 44, 1–12.
- Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Boston, MA: Harvard Business School Press.
- Wood, D. C., & Appel, R. (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes*, 15, 1–13.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, 20, 1–28.

Ali Saadatara is an MA graduate of TEFL from Tarbiat Modares University, Tehran, Iran. His main areas of interest are corpus linguistics and IELTS writing.

Gholam Reza Kiany is a Professor of TEFL at Tarbiat Modares University, Tehran, Iran. He has got his PhD in TEFL from the University of Essex in England in 1996. He has published many books and articles and has presented in a lot of international conferences. His areas of interest are program evaluation, language assessment and research.

Hossein Talebzadeh is an Assistant Professor of Applied Linguistics at Kharazmi University. He obtained his PhD from Tarbiat Modares University. His current research interests lie predominantly in LSP, Academic Writing and (Multimodal) Genre Analysis, Materials Development/Evaluation, and Applied Sociolinguistics and Socopragmatics.