

RESEARCH

Open Access



# Phraseological competence in IELTS academic writing task 2: phraseological units and test-takers' perceptions and use

Masoomah Estaji\*  and Mina Hashemi

\*Correspondence:  
mestaji74@gmail.com

Faculty of Persian Literature  
and Foreign Languages,  
Department of English Language  
and Literature, Allameh  
Tabataba'i University, South  
Allameh St, Modiriat Fly-over,  
Chamran Highway, Tehran, Iran

## Abstract

This study intended to explore the different types of phraseological units in IELTS academic writing task 2 and probe into the IELTS candidates' perceptions of phraseological competence. To this end, a corpus entailing 100 essays (26,423 words) written for IELTS writing task 2 was scrutinized, through which phraseological units were extracted and their types were identified based on Moon's (Phraseology: Theory, analysis, and application, 1998) typology. In addition, semi-structured interviews were conducted with nine IELTS candidates. The results indicated that simple formulae were the most commonly used phraseological units in IELTS writing task 2. The interview data, using inductive thematic and summative content analyses, also demonstrated that IELTS candidates have had varying perceptions of phraseological competence, wherein higher band score candidates appeared to have a deeper and more accurate understanding of the concept. As for the candidates' perceived phraseological competence, they were found to under- or overestimate their own competence at using phraseology. The findings of the study can be employed in making learner dictionaries based on the phraseologies used in IELTS.

**Keywords:** IELTS academic writing, Phraseological competence, Phraseological units, Test-takers' perception

## Introduction

In recent years, there has been an increasing interest in phraseology and, especially, the phraseological competence of foreign language learners. Phraseology is the “study of word combinations” (Howarth, 1998a, p. 24), and a language user's phraseological competence includes everything he or she needs to know to be able to use phraseology and phraseological units in communication. According to Cowie (1998), phraseological competence can be defined as exhibiting a native-like proficiency that “depends crucially on a stock of prefabricated units—or ‘prefabs’—varying in complexity and internal stability” (p. 1).

Phraseology and its use as a technical concept has been associated with many scholars, among whom Howarth (1996) and Cowie (1998) are prominent. It also indebted much to the Firthian linguistics and, more specifically, John Rupert Firth. Later, his students,

called “neo-Firthians,” including John Sinclair and Michael Halliday contributed much to the discipline. Among the vast body of literature on phraseology, the corpus-based studies of Howarth (1998a), Cowie (1998), Moon (1998), and Mel'čuk (1998) are also outstanding. The use and study of phraseology and phraseological competence have gained much prominence in foreign language learning and testing. EFL learners struggle with understanding and using phraseological units, which are probably very different from the phraseological units in their own mother tongue (Poulsen, 2005). Even with no such difference in meaning and use between phraseological units of a second language and those of the first language, foreign language learners have shown to be still at a disadvantage (Vyatkina, 2016).

Phraseological competence of a language learner is assumed to be represented in all four skills and is a contributing factor to the overall proficiency of a second or foreign language. One of the significant skills that constitutes the proficiency of a foreign language learner is writing. Many high-stake standardized tests have been developed to assess the proficiency of language learners, among which is IELTS. In IELTS, two tasks of writing are to be completed by the candidates, and the second task, in particular, requires the candidates to compose an essay by either arguing in favor or against a view, providing a solution to a problem, or expressing personal opinions on a topic. They are, therefore, scored based on four criteria of task achievement, coherence and cohesion, lexical resource, and grammatical resource and accuracy. The way candidates approach writing task 2 of IELTS is essentially in congruence with the value they place on each assessment criterion. Furthermore, this value comes directly from their perceived understanding of the importance of each criterion.

While there have been longitudinal studies on the development of phraseological competence (Candarli, 2021; Garner & Crossley, 2018; Siyanova-Chanturia & Spina, 2020), not many studies have investigated this concept in language testing settings. To identify the types of phraseological units used by IELTS test-takers and examine their perceptions and difficulties they encounter in the use of phraseology, the study at hand set out to explore, firstly, the representations of phraseology and the types of phraseological units in IELTS test-takers' academic writing task 2, and secondly, the IELTS candidates' perceptions of phraseology and phraseological competence.

### Literature review

While interest in vocabulary knowledge only began to gain grounds in the 1970s, the study of phraseology and phraseological units has its roots in the work of Harold Edward Palmer, who gave the first compilation of collocations specifically designed for foreign language learners. Phraseological competence is the ability to use phraseology and phraseological units. It is part of a language learner's communicative, linguistic, and cultural competence. The concept is interrelated with the concepts of idioms, word clusters, *n*-grams, collocations, colligations, collostructions, constructions, patterns, fixed expressions, phrasemes, phrasal lexemes, and phraseologisms. Closely related to the concept of phraseological competence is collocational competence, on which many research studies have been done. While these two concepts are intricately connected, this interconnectedness does not mean that they are one and the same. Phraseological competence is defined by Cowie (1998) as exhibiting a native-like proficiency that

“depends crucially on a stock of prefabricated units—or ‘prefabs’—varying in complexity and internal stability” (p. 1).

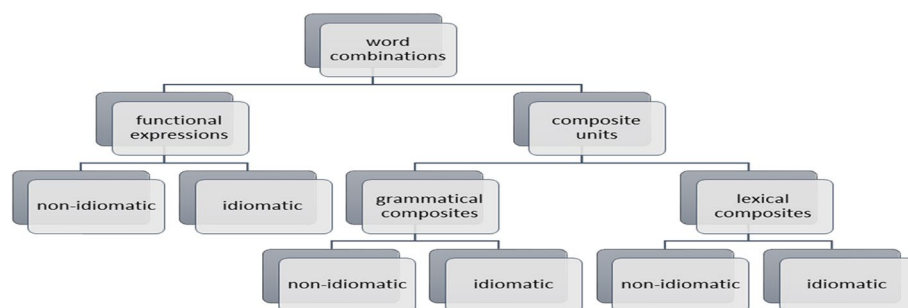
Despite the existence of various terms referring to the same concept, the term phraseological unit has been increasingly used in research regarding phraseology. According to Gläser (1998), a phraseological unit is “a lexicalized, reproducible billexemic or polylexemic word group in common use, which has relative syntactic and semantic stability, may be idiomatized, may carry connotations, and may have an emphatic or intensifying function in a text” (p. 125). According to Kunin (as cited in Naciscione, 2010), “a phraseological unit is a stable combination of words with a fully or partially figurative meaning” (p. 210). Any phraseological unit has a “base form” which “indicate[s] the form of the phraseological unit to which other forms of the PU can be related and with which they can be compared” (Naciscione, 2010, p. 31).

Altogether, phraseological units make up the “phrasicon” of a language, a term referring to the whole set of phrases, “both word-like and sentence-like set expressions” (Gläser, 1998, p. 126). Word-like phraseological units are “nominations”—the traditional parts of speech which are designated to a phenomenon, an object, a process, or a state in the outside world. Sentence-like phraseological units, however, are “propositions” which consist of a nomination and a prediction. They include proverbs, commonplaces, routine formulae, slogans, commandments and maxims, and quotations and winged words (Gläser, 1998).

One of the more widely known models of the study of phraseology is the model proposed by Howarth (1998a). According to Fig. 1, word combinations can either be functional expressions or composite units. These functional expressions are identified by their role in discourse, may be complete utterances in themselves, such as proverbs, catchphrases, and slogans (Alexander, 1984). Composite units, on the other hand, have a syntactic function and are the realizations of phrase structures (Howarth, 1998a).

The division between idiomatic and non-idiomatic is not so straightforward, rather it is a continuum “derived from the application of such criteria as restricted collocability, semantic specialization, and idiomaticity, each of which is gradable” (Howarth, 1998a, p. 28). This collocation continuum is illustrated in Table 1.

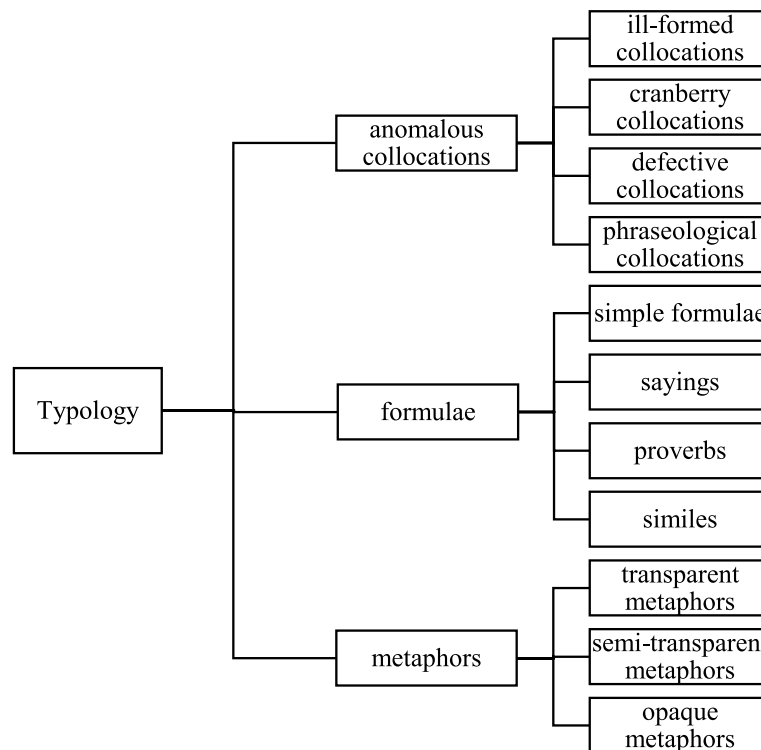
One another significant model is the typology Moon (1998) used in her study of phraseology. This typology is based on the notion of phrasal lexemes which are “the whole range of fixed and semi-fixed complex items,” commonly referred to as idioms or phrases (Moon, 1998, p. 79). A schematic representation of Moon’s typology can



**Fig. 1** Howarth's (1998a) phraseological categories

**Table 1** Collocational continuum

	Free combinations	Restricted collocations	Figurative idioms	Pure idioms
Lexical composite Verb + noun	<i>Blow a trumpet</i>	<i>Blow a fuse</i>	<i>Blow your own trumpet</i>	<i>Blow the gaff</i>
Grammatical composites Preposition + noun	<i>Under the table</i>	<i>Under attack</i>	<i>Under the microscope</i>	<i>Under the weather</i>

**Fig. 2** Moon's (1998) typology

be seen in Fig. 2. In her typology, Moon clarifies ill-formed collocation by giving the examples of *of course*, *by and large*. Cranberry collocations are “those which contain lexical items which are unique to the combination,” such as *put the kibosh on*, *to and fro* (Makkai, as cited in Moon, 1998). Defective collocations “contain items used with meanings which are unique to the combination, or which contain delexicalized or semantically depleted component items,” such as *be the question*, *in time* (Moon, 1998, p. 85).

The final category of anomalous collocations, phraseological collocations, are those “where there is a limited paradigm operating at one (or more) of the word slots (*on show/display*), but the realizations are restricted to a small set that is not fully productive (cf. the unacceptable *\*on demonstration*),” such as *in action* (Moon, 1998, p. 85). Examples of other categories include *in this day and age*, *you know* as simple formulae; *an eye for an eye* as a saying; *you can't have your cake and eat it*

as a proverb; *as nice as pie* as a simile; *alarm bells ring* as a transparent metaphor; *grasp the nettle* as a semi-transparent metaphor; and *kick the bucket* as an opaque metaphor (Moon, 1998).

Many studies have been done on the phraseology of the EFL learners as represented in their writing performance. A large number of these studies have used a native-speaker corpus as a benchmark to test the learners' phraseological competence. Although they have made use of different, incomparable sets of criteria, one finding remains the same: L2 writers use only a much more limited repertoire of phraseological units or word combinations, i.e., merely the ones they have mastered, than the native speakers (Bestgen & Granger, 2014; Wray, 2012); they often use the combinations appropriate to spoken language in their formal academic writing (Bestgen & Granger, 2014). English phraseology can be a major source of problem in high-stake standardized tests like IELTS, affecting the test takers' writing performance with various band scores. With the help of the band scores assigned, more can be learned "about key features of writing ability across different proficiency levels and within different domains" (Taylor, 2004, p. 2).

Howarth (1998b) studied the use of phraseology in the learners' academic writing. Basically, with his study, he aimed to identify and analyze non-standard phraseology in non-native academic writing. He classified these collocations into four categories of free combinations, restricted collocations, figurative idioms, and pure idioms and claimed that they are on a continuum of collocability and restrictedness. Based on the nature of phraseology and corpus linguistics, scholars have also adopted different methodologies to identify phraseologisms. Some studies (e.g., Stubbs, 2001, 2002) have taken the approach to produce *n*-grams, i.e., forms of maximum five adjacent elements. Some others (e.g., Hunston & Francis, 2000; Sinclair, 1991) have worked to generate concordances. Another group of researchers (e.g., Renouf & Sinclair, 1991) has taken an integrative approach of the two, studying "collocational frameworks" or "colligates."

Different types of collocations have been the focus of investigation throughout the years: Some researchers have focused on lexical collocations with specific combinations, for example, verb + noun (e.g., Bahns & Eldaw, 1993; Barfield, 2003; Bonk, 2001; Chan & Liou, 2005; Eyckmans, 2009; Howarth, 1996; Laufer & Waldman, 2011; Peters, 2009) or adjective + noun (e.g., Li & Schmitt, 2010; Siyanova & Schmitt, 2008) while others have focused on grammatical collocations, such as noun + preposition. While studying the collocational competence research, a recurring theme is the influence of L1 on the acquisition of collocations and the development of collocational competence (Bahns & Eldaw, 1993; Nesselhauf, 2003). Likewise, Bonk (2001), Gyllstad (2005), Meara (1996), Nizonkiza (2011), and Zareva et al. (2005) found that language proficiency can be predicated via lexical competence. Collocational competence was also found to be a reliable predictor of lexical competence. This correlation was, however, stronger in the lower levels of proficiency (Chen & Baker, 2016). It has also been found that even highly proficient learners may face difficulties when using or developing L2 knowledge of collocations (e.g., Arnaud & Savignon, 1997; Nesselhauf, 2005; Revier & Henrikson, 2006).

Studies of collocational competence have also taken advantage of different approaches: Firstly, a frequency-based approach might be employed. This approach observes the number of times two or three words have co-occurred. Secondly, a phraseological view, which identifies collocations based on such criteria as degree of opacity and syntactic

structure, might be used. Some other studies have used both computational approaches which, together, make it possible to identify collocations of varying frequencies (Henriksen, 2013). Boers et al. (2006) have differentiated between motivated and arbitrary collocations, and based on multiple experiments, have concluded that the differences between the two groups of collocations determine—more specifically, influence—the learnability of different collocations and teaching approaches.

Another closely related concept of syntactic complexity was studied by Li et al. (2022) in which they examined the development of syntactic complexity in Chinese freshmen who had an intermediate or upper-intermediate level of English proficiency. It was found that syntactic complexity was developed in a non-linear way, where nominal complexity was higher and clausal complexity was lower. However, the students with a higher level of English also showed a higher level of progress over time. To distinguish the writing proficiency of beginner, intermediate, and advanced Chinese English learners, Ouyang et al. (2022) examined a syntactically annotated corpus of 400 narrative compositions based on traditional syntactic complexity and dependency distance measures. The results revealed that among dependency distance measures, the overall mean dependency distance was found the best measure to differentiate the writing proficiency of learners in various proficiency levels. However, none of the traditional measures of syntactic complexity could significantly discriminate the writing proficiency levels of beginner, intermediate, and advanced learners. Hence, the mean dependency distance measures better represent the results of the traditional syntactic complexity measures from language processing perspectives.

More recent studies have employed other methods to examine the development of phraseological competence. In one longitudinal study by Chen and Zhang (2022), the effect of *xu*-argument based continuation task was tested in Chinese students to observe the development of their use of English bi-grams and tri-grams. They also used MI and Delta P scores, and the results revealed that Chinese students used more high-frequency *n*-grams with lower MI scores after the course of 5 months. Another longitudinal study by Candarli (2021) investigated the use of multi-word constructions in L2 learners' academic writing by using a linear mixed-effects modeling and finding the relationship between the reading input and the frequency of combinations in their essays. Although the frequencies of different multi-word constructions varied over time, it was found that the reading input had a direct impact on the frequency and dispersion of these combinations in the essays.

In his article, Li (2017) investigated the direct use of corpus-based instruction in Chinese postgraduate learners' collocational competence—a concept closely associated with phraseological competence—specifically the use of verb + preposition collocations in academic writing by conducting an experiment where one experimental group was instructed with a corpus-based approach and the other control group was taught a traditional essay writing approach. Both experimental and control groups passed a 15-week course on linguistics, including a range of topics—words and meanings, phraseology, collocations, and genre analysis—before being tested at the end of the course and once more after 2 months. Collocational competence of the learners was evaluated based on three factors: The frequency of their use of verb + preposition collocations, the variety of collocations, and the frequency of misused collocations. Overall, the results of the



experiment revealed that while both groups demonstrated improvement in their writing, the direct use of corpus had a more significant effect on the accurate and natural use of academic collocations and fixed phraseological forms (Li, 2017).

Corpus linguistics, however, is not limited to the context of L2 learning and development. Giczela-Pastwa (2021) reviewed how corpus can be used in L2 translation to make the final translated work more native-like, as it is now inevitable. This is the point where phraseology causes problems with the readability of the translated texts. To this end, a mining technique was proposed to be used by Polish novice translators of legal texts, which suggests that translators build a monolingual corpus compiled from electronic sources as this allows them to look for the most natural phraseological combination from non-translated texts. This, over time, would imply that translators get more and more familiar with the better collocations used in the context rather than opting for unnatural and calqued combinations.

Most studies reviewed in this section represent the relationship between phraseological competence, writing performance, and language proficiency. The common finding of these studies is that, first of all, English phraseology is a major source of problems for foreign language learners and test takers, and, secondly, lower proficiency level learners employ a small number of phraseological sequences repeatedly in their writings, whereas higher proficiency learners take advantage of a range of formulaic phrases which are more appropriate to academic writing. However, the results cannot be easily or accurately compared as there have been variations in the use of phraseology, formulaic language, collocations, and lexical bundles in different studies and that more empirical evidence is required for the test takers' use of phraseology and phraseological competence mainly in high-stake standardized tests like IELTS.

As the test-takers' perceptions and uses of phraseology matter mainly in high-stake tests, this study investigated the representations of phraseological competence in IELTS writing task 2 and sought to determine IELTS candidates' perceptions of phraseology and phraseological competence. In short, the study attempted to answer the following research questions.

1. What types of phraseological units are used in IELTS test-takers' task 2 writings with various band scores?
2. What is the IELTS test-takers' perceptions of phraseology and their phraseological competence?

## **Methodology**

### **Corpus**

To answer the first research question, a corpus entailing 100 essays written for the IELTS task 2 writing of mock tests by 100 test-takers were selected, used, and analyzed. The corpus consisted of 26,423 words in total. The IELTS mock tests were selected from an IELTS center in Tehran, Iran, upon receiving consent from the authorities of the center. The sample consisted of 100 essays, selected randomly. The mock exams were held in March 2019. Each essay was evaluated, and phraseological units were extracted partly based on Moon's (1998) typology. The details of the corpus can be seen in Table 2.

**Table 2** Details of the corpus

		Frequency	Percentage
Word count	— 200	16	16%
	200–300	66	66%
	300+	18	18%
Total		100	100%

**Table 3** Demographic information of the participants in the interviews

		Frequency	Percentage
Age	24–29	5	55%
	30–35	2	22%
	36–40	1	11%
	40+	1	11%
Gender	Female	6	67%
	Male	3	33%
Years of studying English	— 1	1	11%
	1–7	2	22%
	8–15	4	44%
	15+	2	22%
Years of preparing for IELTS	— 1	5	55%
	1–2	2	22%
	3+	2	22%
Writing band score	— 3	1	11%
	3–4	2	22%
	4.5–5.5	3	33%
	6+	3	33%
Total		9	100%

Another important reference corpus that was used in this study was the Corpus of Contemporary American English (COCA) which contains more than 560 million words of text, equally divided among spoken, newspaper, fiction, popular magazines, and academic texts. This corpus was used mainly for two reasons: Firstly, it is extremely large. And secondly, it is the most representative of the corpus under study out of all the other available corpora.

### Participants

The quantitative phase of the study was done through the analysis of 100 essays written by 100 IELTS test-takers. To answer the second research question, nine IELTS candidates were selected through purposive sampling, based on their writing band score and their agreement to participate in the study. Purposive sampling is a non-probability and non-random sampling technique which is used to select those cases, based on specific criteria, which would yield the richest information (Mackey & Gass, 2005; Patton, 2002). Table 3 demonstrates the characteristics of the participants who were involved in the interview.



As it can be seen in Table 3, all of the participants (with 67% female and 33% male) were Iranian test-takers in the age range of 24 to 45, who had been studying English for a much longer time than they had been preparing for the IELTS exam.

### Instruments

In this study, two instruments were employed: IELTS writing task 2 and interviews. Writing task 2, in both modules of the IELTS test—the General Training and Academic Module—involves the composition of an essay, in which the test-takers are given a topic and asked to write an essay entailing 250 words in 40 min. The types of the given prompts can be agree/disagree, advantages/disadvantages, cause/effect, cause/solution, problem/solution, and discuss both views. The academic writing tasks used in the corpus of the study were all of agree/disagree question type and were on topics related to technology. This task was scored based on four criteria: task response, coherence and cohesion, lexical resource, and grammatical range and accuracy, varying from 1 to 9 band scores. The test takers' scores in IELTS academic writing task 2 mock tests have been provided in Table 4.

Face-to-face interviews with 9 participants were also conducted to assess the test-takers' perceptions of phraseology and their own phraseological competence. The length of the semi-structured interviews, whose questions were developed by the researchers, was approximately 10 to 15 min, depending on the level of linguistic and communicative ability of the participants. The 11 interview questions were about the candidates' understanding of IELTS writing scoring criteria, the concept of phraseological competence, its significance in IELTS writing, as well as their own abilities in using the phraseological units. The rationale behind opting for semi-structured interviews was that “the researchers apply a written list of questions as their guide, but they are still free to digress and look for more information” (Mackey & Gass, 2005, p. 173). The selection criteria for the participants in individual interviews were their writing band score and their agreement to participate in the study.

### Procedure

This study was comprised of three main phases. Primarily, consent was obtained from an IELTS institute in Tehran for their cooperation and provision of their IELTS mock test-takers' overall band scores, writing task 2 scores, and their essays. The second phase involved compiling a corpus of 100 essays, containing 26,423 words, written for mock IELTS academic writing task 2. The researchers evaluated each essay individually. Then, phraseological units consisting of maximum 5 words were extracted manually by the

**Table 4** Details of the IELTS writing task 2 mock tests

		Frequency	Percentage
Band score	— 3	3	3%
	3–4	15	15%
	4.5–5.5	50	50%
	6+	32	32%
Total		100	100%

researchers out of each essay and coded considering Moon's (1998) typology. The rationale for using Moon's taxonomy is that, from among the existing models, it has been the most frequently used and comprehensive one.

The extracted units included nominal phrases (e.g., "profound impact"), verbal phrases (e.g., "go hand in hand"), adjectival phrases (e.g., "as good as gold"), adverbial phrases (e.g., "in the long run"), conjunctive phrases (e.g., "on the other hand", "in comparison with"), and prepositional phrases (e.g., "in the course of"). Other types of phraseological units extracted included units that functioned as nouns (i.e., noun + noun), units that functioned like verbs (i.e., verb + noun; verb + preposition; verb + sth + noun), and units functioning as adverbs (i.e., preposition + noun).

The units were inserted in Excel worksheet for further data analysis—frequency counts. The same procedure was conducted separately for each essay written for IELTS writing task 2. The units were, then, coded and named, and the use of each type of phraseological unit was determined across band scores in IELTS academic writing task 2. To analyze the frequency of each type of phraseological unit, the data were summarized, categorized, and analyzed using frequency count, percentages, and descriptive statistics. Likewise, to ensure the reliability of the coded data, the intercoder consistency was calculated using the Kappa agreement test. Based on the results, the index of  $k=0.88$  was obtained, resting assured that the data obtained from the 3 coders would have perfect consistency.

The third phase of the study consisted of nine individual semi-structured interviews that were conducted with IELTS candidates of a range of band scores to explore their perceptions of phraseology and phraseological competence in light of IELTS writing task 2. The interview questions were developed by the researcher and examined by two official IELTS examiners and one expert in the field of applied linguistics to establish the content validity of the interview questions. Each interview lasted 10 to 15 min, all of which were audio recorded using a Digital Voice Recorder (DVR). As for the analysis of the data compiled during the interviews, the data were transcribed manually for content and thematic analysis (Research question 2). This study used inductive thematic analysis at the semantic level, and summative content analysis for the interpretation and analysis of interview data.

## Results

### Types of phraseological units used in IELTS test-takers' academic writings task 2

The first research question explored the types of phraseological units that IELTS test-takers used in writing task 2. Looking back at Moon's (1998) typology, it immediately becomes clear that all types of phraseological units were used in task 2, each producing a different effect on the text. Overall, from all 100 essays, 1046 units were extracted, out of which only 684 units fit the typology provided by Moon (1998). It was noticed that no similes were used and when the only example of proverbs was used—by a candidate with a low band score (band 4)—it had an inverse effect on the band score assigned to the task, i.e., the test-taker lost band score. The only example of a saying was also used by a candidate of a rather higher band score (band 6.5).

However, the use of metaphors—be it transparent, semi-transparent, or opaque—seemed to be used by those candidates with higher band scores (above band 5), and

**Table 5** Phraseological units used in IELTS task 2 based on Moon's (1998) typology

Types	Unit	Example	Frequency	Percentage
Anomalous collocations	Ill-formed collocations	As well	86	12.57%
	Cranberry collocations	Pros and cons	9	1.31%
	Defective collocations	By the time	29	4.23%
	Phraseological collocations	In fact	182	26.6%
Formulae	Simple formulae	On the other hand	329	48.09%
	Sayings	Kill two birds with one stone	1	0.14%
	Proverbs	A picture is worth a thousand words	1	0.14%
	Similes	-	-	-
Metaphors	Transparent metaphors	In touch with	32	4.67%
	Semi-transparent metaphors	Lose track of time	4	0.58%
	Opaque metaphors	Silver lining	11	1.6%
Total			684	100%

**Table 6** Restricted collocations used in IELTS writing task 2

Restricted collocations	Composition	Example	Frequency	Percentage
Lexical collocations	Noun + noun	Mass media	302	84.78%
	Verb + noun	Spark debate		
	Adjective + noun	Controversial issue		
	Adverb + verb	Negatively affect		
	Adverb + adjective	Fully aware		
Grammatical collocations	Verb + preposition	Associate with	54	15.21%
	Noun + preposition	A variety of		
	Adjective + preposition	Aware of		
Total			356	100%

lower-score test-takers never used metaphors. This is while simple formulae were used by all candidates across band score and even overused by lower-score candidates (below band 5). The use of anomalous collocations varied across band scores. However, the sub-categories used did not produce a consistent pattern according to band scores received. Phraseological collocations were used most of all, followed in number by ill-formed collocation and defective collocations, leaving cranberry collocations to be the least used category. Table 5 demonstrated these findings in more detail.

In addition to the phraseological units proposed by Moon (1998), another group of units used in the IELTS task 2 was extracted. This group could not fit any of the categories in Moon's (1998) typology, so another group needed to be added. The researchers labeled this new group of phraseological units "restricted collocations" to conform to the terminology used in the literature, specifically in Howarth's (1998a) terms. These standard collocations accounted for approximately 34% of the extracted units and were used by all candidates across band scores, but were used more in number and more effectively by candidates of higher band scores (above band 5). Table 6 represents the findings more coherently.

As it can be seen in Table 6, the restricted collocations used in task 2 writing were divided into two separate categories of lexical and grammatical collocations. These

categories were formed based on the composition of the collocation. Simply put, a lexical collocation consists of the combination of two content words—for example, a noun followed by another noun. A grammatical collocation, on the other hand, is the combination of a content word and function word—for example, a verb followed by a preposition.

The following excerpt has been extracted from a band 6 essay, in which instances of simple formulae (“that is to say”), phraseological collocations (“due to”), and both kinds of restricted collocations can be seen:

*Another negative effect of the widespread use of modern communication technologies is that a huge number of people have access to these facilities; that is to say, they are unable to make new relationships due to lack of necessary equipment.*

Another extract from a band 5 essay is inserted here:

*According to new researches, [...] there are many different shapes of though which some of them are optimistic and others are pessimistic. First of all, it should be mentioned that any new changes may have some positive and negative influence on our life. For instance, new communication technology such as social networks which absolutely has pros and cons for human's life.*

In this excerpt, there are examples of simple formulae (“first of all,” “for instance,” and “such as”), phraseological collocations (“according to”), cranberry collocations (“pros and cons”), and both types of restricted collocations—lexical collocations (“positive influence,” “negative influence,” “communication technology,” and “social networks”) and grammatical collocations (“influence on”).

There were six phraseological units that were extracted from the essays, but did not fit any of the categories of Moon's (1998) typology or restricted collocations, nor did they seem to fit under the same category. These six units were labelled “free combinations.” Example of these combinations are “lack thereof,” which is a combination of a noun followed by an adverb, and “the public,” which is a combination of a determiner and a noun. Table 7 and Fig. 3 illustrate the overall findings for the first qualitative research question.

The following pie charts display each type of phraseological units determined in each band score (Figs. 4, 5, 6, 7, and 8).

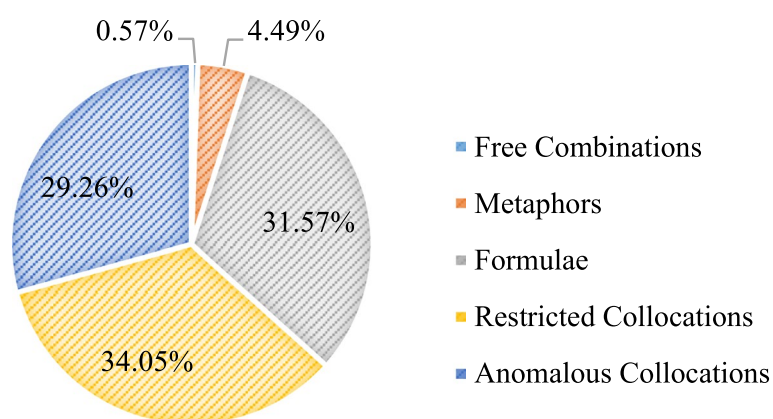
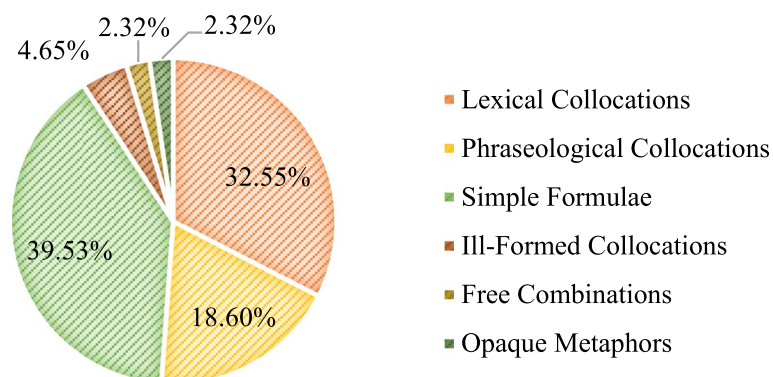
The essays were, then, divided into 10 groups based on the received band score. The number of units used in each band score essay was counted and the mean number of phraseological units in each essay was computed. Table 8 summarizes the information.

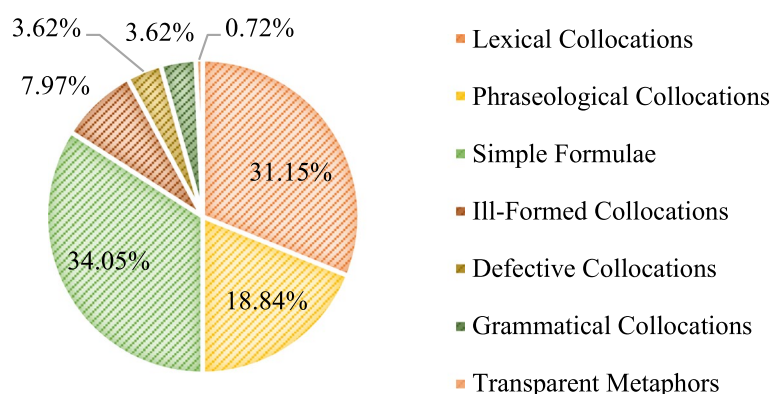
Overall, a rising trend can be seen in the mean number of phraseological units as the writing band score gets higher. The difference between the mean numbers is smaller in lower band scores, and it quickly rises when the band score gets to its highest—i.e., band score 7. The only exception to this rise is for band score 4.5, in which the mean number of phraseological units is lower than band score 4. Besides the knowledge of phraseology which has affected writing performance, another reason that might explain this increase to some extent was assumed to be the length of essay, as higher writing band scores came with longer essays.

Based on the findings for the first research question, a new typology for the use of phraseological units in IELTS task 2 writing can be proposed (Fig. 9). This typology

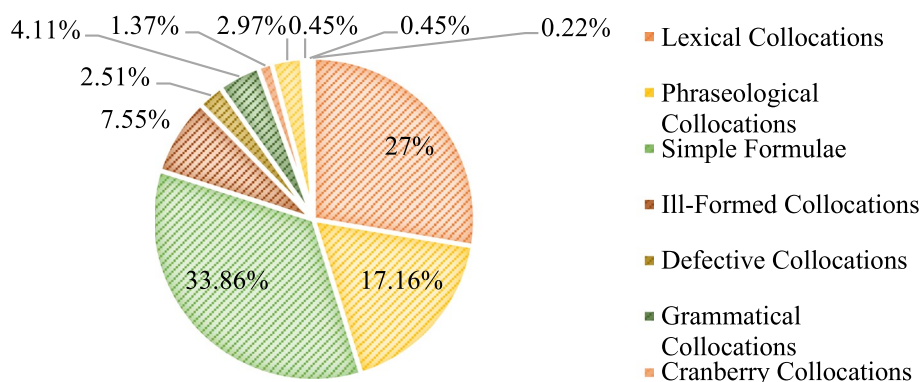
**Table 7** Overall phraseological units extracted from IELTS task 2 writing essays

		Frequency	Percentage
Anomalous collocation	Ill-formed collocations	86	8.22%
	Cranberry collocations	9	0.86%
	Defective collocations	29	2.77%
	Phraseological collocations	182	17.41%
Restricted collocations	Lexical collocations	302	28.89%
	Grammatical collocations	54	5.16%
Formulae	Simple formulae	329	31.48%
	Sayings	1	0.09%
	Proverbs	1	0.09%
	Similes	-	-
Metaphors	Transparent metaphors	32	3.06%
	Semi-transparent metaphors	4	0.38%
	Opaque metaphors	11	1.05%
Free combinations		6	0.57%
Total		1,046	100%

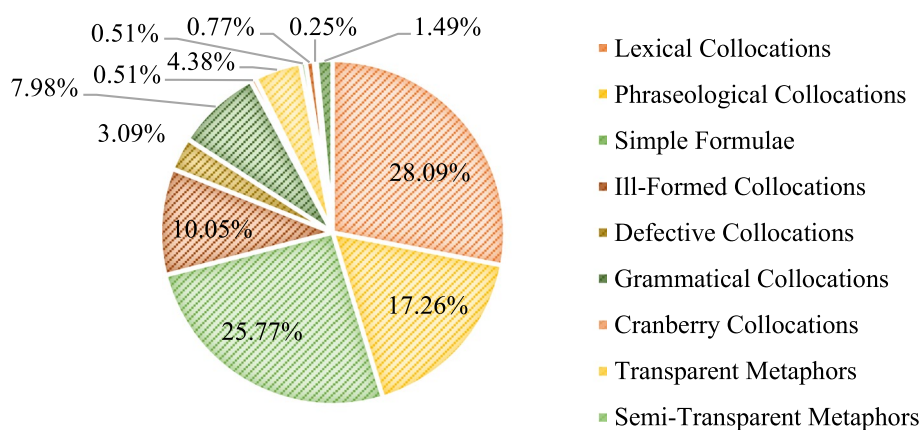
**Fig. 3** Percentages of each phraseological unit used in IELTS task 2 writing**Fig. 4** Percentages of each phraseological unit type used in band scores 2.5–3.5



**Fig. 5** Percentages of each phraseological unit type used in band scores 4–4.5



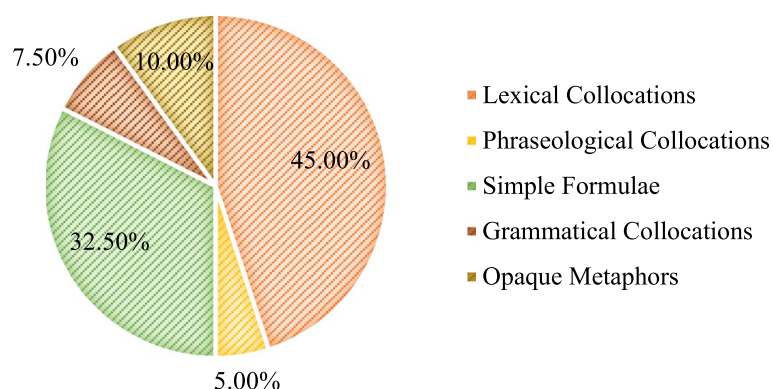
**Fig. 6** Percentages of each phraseological unit type used in band scores 5–5.5



**Fig. 7** Percentages of each phraseological unit type used in band scores 6–6.5

has omitted 3 categories from the original typology suggested by Moon (1998). These categories include sayings, proverbs, and similes. This is because these types of units were not used or used in the minimum number (proverbs were used only once in a low band score essay) in the essays written for IELTS task 2. Three new categories,





**Fig. 8** Percentages of each phraseological unit type used in band score 7

**Table 8** Frequency counts and mean number of the units used in each band score

Writing band score	Frequency of essays (%)	Frequency of phraseological units (%)	Mean number of units in each essay
2.5	3 (3%)	7 (0.66%)	2.8
3	-	-	-
3.5	6 (6%)	36 (3.54%)	6
4	9 (9%)	75 (6.79%)	8.33
4.5	9 (9%)	63 (6.28%)	7
5	14 (14%)	141 (13.68%)	10.07
5.5	27 (27%)	296 (28.7%)	10.96
6	25 (25%)	321 (30.33%)	12.84
6.5	5 (5%)	67 (6.41%)	13.4
7	2 (2%)	40 (3.82%)	20
Total	100 (100%)	1046 (100%)	

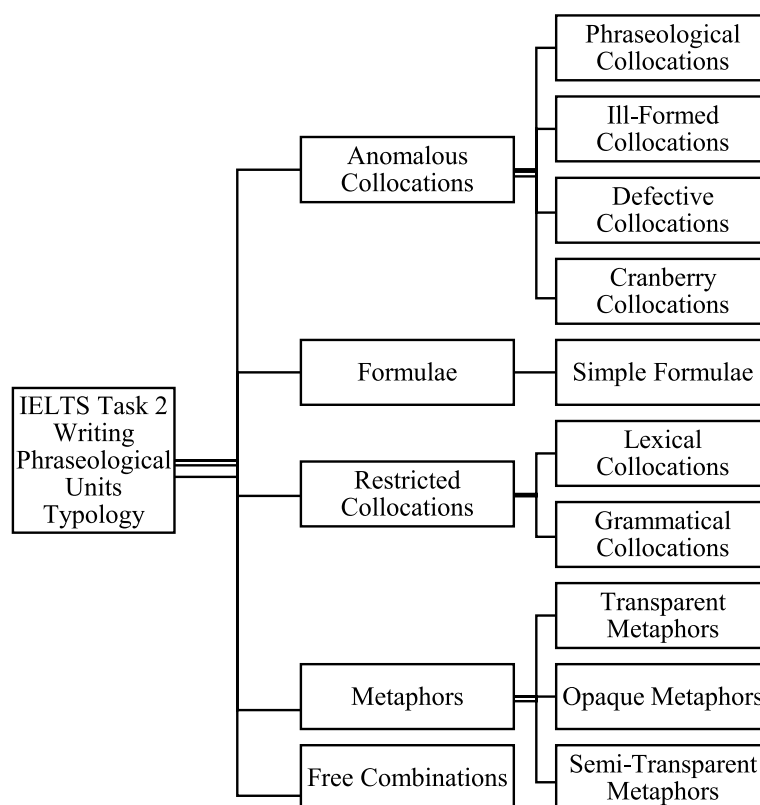
however, were added to this typology, namely lexical collocations, grammatical collocations, and free combinations.

The results of the first research question revealed that different types of phraseological units are used by candidates who have received different band scores. Overall, some types of phraseological units have been overused by lower band score candidates, namely simple formulae, while others were favored by higher band score candidates, namely lexical collocations and metaphors. In the meantime, other types of phraseological units were used more or less by all candidates of all band scores, and the use of each type has had different effects on the band score assigned to the writing task.

#### IELTS test-takers' perceptions of phraseology and their phraseological competence

The second research question explored the test-takers' perception of phraseology and phraseological competence. To answer this research question, individual interviews were conducted. The interview data indicated that IELTS candidates had 3 groups of definitions for the term "phraseological competence." Based on the results of the interviews, candidates with higher band scores appeared to have a more profound definition





**Fig. 9** Proposed typology for the use of phraseological units in IELTS task 2 writing

**Table 9** Phraseological competence definitions as reported by IELTS candidates

Phraseological competence definitions	Writing band score 5–		Writing band score 5+		Total	
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
Knowledge of phrase and fixed forms	-	-	4	44%	4	44%
Knowledge of phrasal verbs	1	11%	1	11%	2	22%
No definition was provided	3	33%	-	-	3	33%

of phraseological competence. These definitions provided by the interviewees are summarized in Table 9 and are explained through excerpts taken from the interviewees.

As Table 9 demonstrates, candidates with band scores of higher than band 5 seemed to have a more accurate and comprehensive understanding of phraseological competence. According to Table 9, the majority of candidates with higher writing band scores (80%) defined phraseological competence as the knowledge of phrases and fixed forms, whereas only one high score candidate (20%) defined phraseological competence as the knowledge of phrasal verbs. Similarly, the majority of lower-score candidates (75%) did not understand the question at all and could not reply properly. Only one low score candidate (25%) defined phraseological competence as the knowledge of phrasal verbs. Two excerpts on phraseological competence definition are provided here:

*I think it is the ability of using idioms, proverbs, phrasal verbs, or any other combi-*

*nation of words into a fixed expression. It seems that it has a major role in fluency. (Participant 2, band score 6)*

*Phraseological competence, in my opinion, is the quality use of set or fixed expressions, idioms, and other types of multi-word lexical units in the writing. (Participant 8, band score 7)*

Regarding the importance of phraseological competence—collocations, phrasal verbs, and formulae—the analysis of the interview data indicated that the candidates' band score did not influence their perceptions of the importance of phraseology in IELTS writing. However, interviewees with higher band scores were more willing to delve into the issue and put more effort into figuring out what phraseological competence means. Therefore, most answers were received from higher proficiency candidates. The findings about the participants' perceptions of the importance of phraseology in writing are given in Table 10.

As Table 10 displays, the majority of candidates (55%) believed that phraseological competence is an element of language proficiency. All candidates, regardless of their perceptions of the importance of phraseology, agreed that phraseology contributes to the overall writing band score only if word combinations are used naturally, appropriately, and accurately. One participant, in particular, stated:

*[...] Lexis is as important as the other three criteria and using collocations, idioms, phrasal verbs, and even proverbs is a way of showing one's range of vocabulary. However, one should be careful with the register of idioms and proverbs and use them if and only if they are confident how to use them naturally. (Participant 9, band score 7)*

Another excerpt is given from the interview with a higher band score participant:

*I think idioms and collocations about the subject are the simplest way to expand the writing body. You can explain your [ideas] in a short way which will be clear for*

**Table 10** Importance of phraseology in IELTS writing as reported by IELTS candidates

Importance of phraseology in IELTS writing	Writing band score 5 –		Writing band score 5 +		Total	
	f	%	f	%	f	%
Is not as important as other factors	1	11%	2	22%	3	33%
Shows overall linguistic ability	1	11%	4	44%	5	55%
Helps expand ideas in writing	-	-	2	22%	2	22%
Has psychological impact on examiner	-	-	1	11%	1	11%
Makes writing unique	2	22%	-	-	2	22%

*readers. Also in body of the writing, there is [extreme] need to use correct combination of words which [are] related to the issue. This leads to importance of collocations. If you have knowledge about collocations, you will apply complex words in [a] prompt way which [reduces] time [...]. (Participant 5, band score 6.5)*

The participants were also required to judge their own phraseological competence and how good they perceived themselves to be at using phraseological units. On the whole, participants could be divided into two groups based on their perceptions, regardless of their band score. One half of the participants were confident in their ability to use phraseology in writing, and the other half believed they had limited knowledge of phraseology and the use of phraseological units. The details can be seen in Table 11.

As Table 11 demonstrates, the participants' responses were irrespective of their writing band score, as there was one participant who had a low writing band score; however, she considered herself to be quite good at utilizing phraseological units. This participant stated that she can use 7 phraseological units out of the 10 units that she has learned. There were also two higher score participants who did not find themselves competent at phraseology. They also claimed that not having broad knowledge of phraseological units was their own choice and because they lacked the necessary time to prepare themselves for the official IELTS test, they would rather spend their time improving other aspects of their writing, for example, grammar. Hereupon, one interviewee stated:

*Honestly speaking, my knowledge of [phraseology is] absolutely limited and rarely do I use [phraseological units]. If I knew more, I would use them in my essays and speaking. Considering limitations of my time [...] for studying and learning, I prefer learning more grammar. (Participant 2, band score 6)*

One of the high-band score participants, who believed to be competent at phraseology, claimed that she actually double checked her own writing specifically to check her use of phraseological units.

*To be good at using collocations, idioms, etc., which I think are the treasuries of a language, [one] needs to be somehow rich... linguistically. [...] In this case, when I write an essay and look back and review my writing to check my weaknesses and strengths, I find myself good at using [these] fixed forms and usually get satisfied about that. (Participant 8, band score 7)*

One major theme surfaced from the analysis of interview data: *Phraseology contributes to the writing band score only if and when word combinations are used naturally, appropriately, and accurately.* On the whole, the data from the interviews concerning

**Table 11** Perception of own phraseological competence as reported by IELTS candidates

	Writing band score 5 –		Writing band score 5 +		Total	
	f	%	f	%	f	%
Is good at utilizing phraseology	1	11%	3	33%	4	44%
Has limited knowledge of phraseology	3	33%	2	22%	5	55%

candidates' general perception of phraseological competence indicated that the interviewees with higher writing band scores and overall proficiency scores have a deeper and more appropriate understanding of phraseological competence and its importance. In other words, interviewees with higher band scores and lower band scores reported different perceptions of phraseological competence, revealing that language proficiency and writing performance play important roles. It was also found that a candidate's understanding of their own phraseological competence is not influenced solely by their writing or overall band score.

## Discussion

The present study was a bid to investigate the types of phraseological units used by candidates in IELTS academic writing task 2. It was also an attempt to explore the test-takers' perceptions of phraseology and phraseological competence. Based on the results, it was found that all candidates, regardless of their band score, showed a tendency to use simple formulae at least once in their essays. The second most widely used phraseological unit was found to be lexical collocations—which was defined as any combination of content words with strong tendency to co-occur. It was also found out that as writing band score of the candidates got higher, the range and number of phraseological units employed increased, and the gap seemed to widen especially in higher band scores. The findings appear not to be very different from the studies in the extant literature.

Regarding the increase in the number of phraseological units used, one assumption is that higher band score candidates had written more words in their essays; therefore, the number of units they have used increased as well. Most of the analyzed essays that were underlength—which were also penalized—had lower band scores, with the exception of a few, which are assumed to be due to lack of concentration or fatigue on the part of the raters. As for the range of phraseological units, most L2 writers have previously been found to show a tendency to overuse a limited repertoire of phraseological units (Bestgen & Granger, 2014; Wray, 2012). This is similar to what was discovered in the current study: lower band score IELTS candidates did, in fact, use similar phraseological units repeatedly throughout their essays (which were not counted in the final frequency count for the respective research question). This means that especially lower proficiency test-takers have access to a very limited set of word combinations (Ädel & Erman, 2012; Chen & Baker, 2010; Staples et al., 2013). In addition, the most commonly used phraseological units were simple formulae which, according to the typology, are the easiest to acquire as they are more fixed and unvaried than other types and do not have idiomatic meanings.

Phraseology in academic writing, in general, has two characteristics: (1) There is limited use of traditional types of phraseological units like idioms (Biber et al., as cited in Ackerman & Chen, 2013), and (2) there is wide use of academic collocations and phrases (Ackerman & Chen, 2013; Simpson-Vlach & Ellis, 2010). This, in the present research study, can be explained in terms of the types of phraseological units that were found in lower and higher band scores. While lower band score candidates repeatedly used a fixed set of simple formulae—which were counted only once in the present study—higher band scores used a wide range of lexical collocations (Chen & Baker, 2014). This may also be indicative of the fact that higher band scores show more native-like

linguistic ability as they used a higher number of phraseological units. However, lower band score candidates claimed that memorization and learning specific types of phraseological units have been challenging for them.

What is interesting is that the increase in the number of phraseological units did not follow the same pace as the increase in band score, that is the gap in the number of used phraseological units is bigger as band scores increase. This is supported by Laufer and Waldman (2011) who compared the use of collocations in learner corpus and native speaker corpus and found that learners at various proficiency levels used far fewer collocations than native speakers, and the number of collocations significantly increased in advanced levels. Similarly, Li et al. (2022) found that the students with a higher level of English proficiency evinced a higher level of progress in the syntactic and phraseological complexity of their texts. In a recent longitudinal study, Siyanova-Chanturia and Spina (2020) reported that second language learners would use fewer frequent and strongly associated combinations with more exposure to L2 and over time as they learn to use language with more creativity and generativity. Estaji and Montazeri (2022) also found that this is the case in IELTS speaking, too, as the use and production of collocations increased in parallel with their band score increasing.

Regarding the conception of phraseological competence, two definitions were suggested by the interviewed participants, while most lower band score candidates did not—and probably could not—have any idea at all. Some participants believed that phraseological competence is the knowledge of using phrases, while the rest believed it refers to the knowledge of phrasal verbs. Not knowing what phraseological competence means on the part of the lower band score candidates may explain their lack of understanding of the assessment criteria, which can, in turn, explain their lower band score. However, as it was mentioned, there were high scoring candidates who did not provide an appropriate definition of the term, which can be another evidence that CEFR, as well as IELTS scoring rubrics, do not pay heed, at least explicitly, to the construct of phraseological competence (Paquot, 2018).

To be phraseologically competent in IELTS writing, the majority of candidates ascertained that phraseological competence is only a part of overall linguistic ability and language proficiency. The interview data also revealed that the candidates viewed phraseological competence as a way to expand their ideas in writing and make their essay unique, and a factor that influences the examiner psychologically. Some interviewees claimed that phraseological competence is of less significance in IELTS writing, which can be due to the fact that there is no explicit and direct mention of the construct in the band descriptors. It was also assumed that the reason why many participants claimed grammatical accuracy has been of a much more significance in writing band score was the focus IELTS instructors put on grammar and lack thereof on phraseology.

Regarding the candidates' judgement of their own phraseological competence, the data revealed that the participants either were confident in their use of phraseology or perceived themselves as lacking necessary and adequate knowledge of phraseology. Surprisingly, the band score of the participants did not affect how they viewed their own phraseological competence. This could possibly be because of the participants' improper understanding of the concept of phraseological competence, their self-concept, personality traits, self-appraisal regarding their own phraseological competence, overestimating

and underestimating their abilities, and self-flattery syndrome, where some participants have tried to make themselves look better, however deceitfully.

In band scores under 4.5, many of the fixed simple formulae had been used incorrectly, which were not counted at all. For example, many instances were found where the candidates had used “in the other hand” instead of the correct form “on the other hand.” Hence, the differences in IELTS band scores can be explained by the types of phraseological units they use, a feature that is missing in the IELTS writing rubric. The phraseological dimension of language should be considered in writing assessment (to determine a text’s macrostructure and functions, the level of formality and interactivity, and intended register) as it can shed light on the social aspects and uses of language (Paquot, 2018). In this study, there were instances of the correct use of a phrase in an inappropriate context. These units were counted in the present study; however, it remains ambiguous to what extent the knowledge of register should be taken into consideration when it comes to the study of phraseological competence.

There were also units that were used so few in number (for example, only once throughout all 100 essays) that the effect they might possibly have on writing band score was not clear. The inaccurate and inappropriate use of phraseological units can be due to factors outside of the domain of the present study, such as the effects of first language (Koya, 2005; Laufer & Waldman, 2011). The results of this study and those found in the extant literature (Gablasova et al., 2017; Gries, 2010) highlight this vision that frequency alone should not be the only metric for determining the use and exposure and that dispersion should also be coupled with frequency information in learner corpus studies. All in all, to better differentiate the test-takers based on their band scores, language tests should rather concentrate on the use of academic phraseology and the types of phraseological units the test-takers employ. By placing phraseology in their rating scale descriptors, language testers can create rubrics by which they can better determine the test-takers’ language ability and test scores and recognize those linguistic features which are challenging for them at any proficiency level and with any test score.

### Conclusion and implications

The study at hand was an attempt to explore the different types of phraseological units used in IELTS academic writing task 2 across band scores. It also explored IELTS test-takers’ perceptions of phraseology and phraseological competence. The findings revealed that most IELTS candidates tended to use simple formulae and lexical collocations. However, simple formulae were overused by lower band score candidates, while the range and number of phraseological units increased as the band scores rose. This increase was more noticeable in higher band scores. Generally, metaphors were used more by higher band score candidates. The participants considered phraseological competence to be the knowledge of phrases or phrasal verbs. Concerning how they judged their own phraseological competence, the participants viewed themselves as either phraseologically competent or incompetent, irrespective of the band score they had received. Hence, it can be concluded that the test-takers’ phraseological competence, as shown by their use of a wide range and great number of phraseological units, is representative of their writing performance enabling them to employ a broader range and number of units in their essays, which, in turn, increases their band score.

Even having a deep understanding of the concept of phraseological competence does not mean that the candidates are able to accurately judge their own phraseological competence and others'. This is particularly true when there are other variables, namely affective, cognitive, and instructional variables, such as the level of confidence, practices, and learning styles at work which might overshadow the impact of writing band score on the participants' judgement of their own phraseological competence. Based on the results, a new typology is proposed which highlights the various types of phraseological units that are employed in IELTS task 2 writing. This typology, however, fails to give information on how these phraseological units are prioritized by IELTS test-takers and examiners. The proposed typology consists of five main categories: restricted collocations, anomalous collocations, formulae, metaphors, and free combinations. This typology adds two more categories to an already-existing typology provided by Moon (1998), while eliminating 3 subcategories of formulae—proverbs, sayings, and similes—for the typology to be matched with the requirements of IELTS academic writing task 2.

The findings of the study would assist IELTS candidates in developing their awareness regarding the concept of phraseological competence and the important role it plays in their writing and overall band score. Furthermore, teachers would raise their awareness of phraseology and the challenges and problems that IELTS candidates encounter to promote their phraseological competence. The findings of the study are also helpful for examiners in that they can be made aware of the impacts phraseological competence has on their overall ratings of the candidates' writings. This could also be considered in the training of the examiners. Another important implication that the findings of the present study can have for the materials developers would be dictionary making (Howarth, 2013), as "reference to the dictionary" was identified to be the main strategy in promoting the IELTS test-takers' phraseological competence, and learner dictionaries seem to be neglecting this focal area of language (Siepmann, 2008).

The most important limitation of the study was analyzing mock tests instead of the official IELTS tests. This is limiting because while taking the actual IELTS test, many factors, such as anxiety, are at work that may influence how the test-takers perform. Another reason for opting for mock tests instead of official IELTS tests was the lack of access on the part of the researchers. Building upon the limitation and delimitation of this study, several suggestions have been provided for future and further research. For example, future studies can explore phraseology and phraseological competence in relation to other skills, such as speaking. Future researchers may intend to explore IELTS examiners' perceptions of the extent to which phraseological complexity may influence their overall judgement and rating of the essays. Future studies could make it the focus of attention to count, identify, and analyze the erroneous phraseological combinations, rather than the accurately employed ones. This study focused on the range and types of phraseological units. Future studies may intend to investigate their functions or degree of fixedness.

#### Abbreviations

CEFR	Common European Framework of Reference
COCA	Corpus of Contemporary American English
DVR	Digital voice recorder
EFL	English as a foreign language
IELTS	International English Language Testing System
L2	Second language



### Acknowledgements

The authors highly appreciate the insightful comments suggested by the editor and anonymous reviewers.

### Authors' contributions

All authors have materially participated in the research and article preparation. Additionally, all authors have approved the final article. Masoomeh Estaji: Conceptualization, formal analysis, methodology, supervision, writing—original draft, and writing—review and editing. Mina Hashemi: Conceptualization, data collection, corpus analysis, investigation, methodology, resources, writing—original draft, and writing—review and editing. Both authors read and approved the final manuscript.

### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Availability of data and materials

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

### Declarations

#### Competing interests

The authors declare that there are no competing interests.

Received: 10 December 2021 Accepted: 12 August 2022

Published online: 08 September 2022

### References

- Ackerman, K., & Chen, Y.-H. (2013). Developing the Academic Collocation List (ACL) – a corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, 12(4), 235–247. <https://doi.org/10.1016/j.jeap.2013.08.002>
- Ädel, 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*, 27(1), 24–43. <https://doi.org/10.1016/j.esp.2011.08.004>
- Alexander, R. J. (1984). *Primary teaching*. Cassell.
- Arnaud, P. J. L., & Savignon, S. J. (1997). Rare words, complex lexical units and the advanced learner. In J. Coady & T. Huchin (Eds.), *Second language vocabulary acquisition* (pp. 157–173). Cambridge University Press.
- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101–114. [https://doi.org/10.1016/0346-251x\(93\)90010-e](https://doi.org/10.1016/0346-251x(93)90010-e)
- Barfield, A. (2003). *Collocation recognition and production: Research insights*. Chuo University.
- Bestgen, Y., & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing*, 26, 28–41. <https://doi.org/10.1016/j.jslw.2014.09.004>
- Boers, F., Eyckmans, J., & Stengers, H. (2006). Motivating multiword units. Rationale, mnemonic benefits and cognitive style variables. *EUROSLA Yearbook*, 6(1), 169–190. <https://doi.org/10.1075/eurosla.6.11boe>
- Bonk, W. J. (2001). Testing ESL learners' knowledge of collocations. In T. Hudson & D. Brown (Eds.), *A focus on language test development: Expanding the language proficiency construct across a variety of tests* (pp. 113–142). University of Hawaii.
- Candarli, D. (2021). A longitudinal study of multi-word constructions in L2 academic writing: The effects of frequency and dispersion. *Reading and Writing*, 34(5), 1191–1223. <https://doi.org/10.1007/s11145-020-10108-3>
- Chan, T. P., & Liou, H. C. (2005). Effects of web-based concordancing instruction of EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231–250. <https://doi.org/10.1080/09588220500185769>
- Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14(2), 30–49.
- Chen, Y., & Baker, P. (2014). 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>
- 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>
- Chen, J., & Zhang, X. (2022). L2 development of phraseological knowledge via a xu-argument based continuation task: A latent curve modeling approach. *System*, 106, 102767. <https://doi.org/10.1016/j.system.2022.102767>
- Cowie, A. P. (1998). *Phraseology: theory, analysis, and applications*. Oxford University Press.
- Estaji, M., & Montazeri, M. R. (2022). The representation of collocational patterns and their differentiating power in the speaking performance of Iranian IELTS test-takers. *Frontiers in Education*, 7, 827927. <https://doi.org/10.3389/feduc.2022.827927>
- Eyckmans, J. (2009). Towards an assessment of learners' receptive and productive syntagmatic knowledge. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 139–152). Palgrave Macmillan.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Exploring learner language through corpora: comparing and interpreting corpus frequency information. *Language Learning*, 67(S1), 130–154. <https://doi.org/10.1111/lang.12226>
- Garner, J., & Crossley, S. (2018). A latent curve model approach to studying L2 n-gram development. *Modern Language Journal*, 102, 494–511. <https://doi.org/10.1111/modl.12494>

- Giczela-Pastwa, J. (2021). Developing phraseological competence in L2 legal translator trainees: A proposal of a data mining technique applied in translation from an LLD into ELF. *The Interpreter and Translator Trainer*, 15(2), 187–204. <https://doi.org/10.1080/1750399x.2020.1868177>
- Gläser, R. (1998). The stylistic potential of phraseological units in the light of genre analysis. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 125–143). Oxford University Press.
- Gries, S. T. (2010). Dispersions and adjusted frequencies in corpora: further explorations. In S. T. Gries, S. Wulff, & M. Davies (Eds.), *Corpus linguistic applications: Current studies, new directions* (pp. 197–212). Rodopi.
- Gyllstad, H. (2005). Words that go together well: Developing test formats for measuring learner knowledge of English collocations. In F. Heinat & A. Kligvall (Eds.), *The department of English in Lund: Working papers in linguistics*, 5 (pp. 1–31). Lund University. Available online at [http://www.sol.lu.se/engelska/wp.html?expand\\_menu=14](http://www.sol.lu.se/engelska/wp.html?expand_menu=14).
- Henriksen, B. (2013). Research on L2 learners' collocational competence and development: A progress report. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 29–56). Eurosla.
- Howarth, P. A. (1996). *Phraseology in English academic writing: some implications for language learning and dictionary making*. Max Niemeyer.
- Howarth, P. A. (1998a). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), 24–44. <https://doi.org/10.1093/applin/19.1.24>
- Howarth, P. A. (1998b). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161–186). Oxford University Press.
- Howarth, P. A. (2013). *Phraseology in English academic writing: Some implications for language learning and dictionary making* (Vol. 75). Walter de Gruyter.
- Hunston, S., & Francis, G. (2000). Pattern grammar: A corpus-driven approach to the lexical grammar of English. *Journal of Linguistics*, 37(2), 451–462. <https://doi.org/10.1017/S0022226701001001>
- Koya, T. (2005). *The acquisition of basic collocations by Japanese learners of English* (Unpublished doctoral dissertation). Waseda University.
- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second-language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647–672. <https://doi.org/10.1111/j.1467-9922.2010.00621.x>
- Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology*, 21(3), 153–171.
- Li, J., & Schmitt, N. (2010). The development of collocations use in academic texts by advanced L2 learners: A multiple case study approach. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 23–46). Continuum.
- Li, Y., Nikitina, L., & Riget, P. N. (2022). Development of syntactic complexity in Chinese university students' L2 argumentative writing. *Journal of English for Academic Purposes*, 56, 101099. <https://doi.org/10.1016/j.jeap.2022.101099>
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Competence and performance in language learning* (pp. 35–53). Cambridge University Press.
- Mel'čuk, I. (1998). Collocations and lexical functions. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and application* (pp. 23–53). Oxford University Press.
- Moon, R. (1998). Frequencies and forms of phrasal lexemes in English. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and application* (pp. 79–100). Oxford University Press.
- Naciscione, A. (2010). *Stylistic use of phraseological units in discourse*. John Benjamins Publishing Company.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242. <https://doi.org/10.1093/applin/24.2.223>
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. John Benjamins Publishing Company.
- Nizonkiza, D. (2011). The relationship between lexical competence, collocational competence, and second language proficiency. *English Text Construction*, 4(1), 113–146. <https://doi.org/10.1075/etc.4.1.06niz>
- Ouyang, J., Jiang, J., & Liu, H. (2022). Dependency distance measures in assessing L2 writing proficiency. *Assessing Writing*, 51, 100603. <https://doi.org/10.1016/j.asw.2021.100603>
- Paquot, M. (2018). Phraseological competence: A missing component in university entrance language tests? Insights from a study of EFL learners' use of statistical collocations. *Language Assessment Quarterly*, 15(1), 29–43. <https://doi.org/10.1080/15434303.2017.1405421>
- Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work*, 1(3), 261–283. <https://doi.org/10.1177/1473325002001003636>
- Peters, E. (2009). Learning collocations through attention-drawing techniques: A qualitative and quantitative analysis. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 194–207). Palgrave Macmillan.
- Poulsen, S. (2005). *Collocation as a language resource: a functional and cognitive study in English phraseology* (Unpublished doctoral dissertation). University of Southern Denmark.
- Renouf, A., & Sinclair, J. (1991). Collocational frameworks in English. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honor of Jan Svartvik* (pp. 128–143). Longman.
- Revier, R. L., & Henriksen, B. (2006). Teaching collocations: Pedagogical implications based on a cross-sectional study of Danish EFL. In M. Brendsten, M. Björklund, C. Fant, & L. Forsman (Eds.), *Språk, lärande och utbildning i sikt [Language, learning, and education]* (pp. 191–206). Åbo Akademi Vasa.
- Siepmann, D. (2008). Phraseology in learners' dictionaries: What, where and how? In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 185–202). John Benjamins Publishing Company.
- 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 University Press.
- Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review*, 64(3), 429–458. <https://doi.org/10.3138/cmlr.64.3.429>

- Siyanova-Chanturia, A., & Spina, S. (2020). Multi-word expressions in second language writing: A large-scale longitudinal learner corpus study. *Language Learning*, 70(2), 420–463. <https://doi.org/10.1111/lang.12383>
- 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>
- Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics*, 7(2), 215–244. <https://doi.org/10.1075/ijcl.7.2.04stu>
- Stubbs, M. (2001). *Words and phrases: corpus studies of lexical semantics*. Blackwell Publishers.
- Taylor, L. (2004). Second language writing assessment: Cambridge ESOL's ongoing research agenda. *Research Notes*, 16, 2–3.
- Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology*, 20(3), 159–179.
- Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*, 32(1), 231–254. <https://doi.org/10.1017/s026719051200013x>
- Zareva, A., Schwanenflugel, P., & Nikolova, Y. (2005). Relationship between lexical competence and language proficiency. *Studies in Second Language*, 27(4), 567–595. <https://doi.org/10.1017/s0272263105050254>

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---