frontiers in Education

ORIGINAL RESEARCH
published: 22 March 2022
doi: 10.3389/feduc.2022.827927

The Representation of Collocational Patterns and Their Differentiating Power in the Speaking Performance of Iranian IELTS Test-Takers

Masoomeh Estaji and Mohammad Reza Montazeri*

Department of English Language and Literature, Allameh Tabataba'i University, Tehran, Iran

Corpus studies have highlighted the role of multiword units in naturally occurring language. Speech theories, too, have underlined the linkage between such formulaic sequences- collocations in particular- and speech production. Few studies, however, have focused their attention on examining collocations in speaking assessment, especially in high-stakes tests. This study investigated the most frequently used collocational patterns and their discriminating power across three groups of participants with band scores 6, 7, and 8 respectively. To collect data, a corpus entailing 60 IELTS speaking samples, 20 samples from each band score, and approximately 110,000 words was gleaned. The results revealed that L1 (adjective + noun) and L7 (verb + noun) were the most frequently used types of lexical collocations, and G8 (verb + preposition) was the most frequently used grammatical collocation. The study also found that L1(adjective + verb), L5 (noun of noun), L8 (phrasal verb and adverb), L9 (noun and phrasal verb), and L10 (phrasal verb and noun) were the five types of lexical collocations with the most discriminating power across the band scores. Given the grammatical collocations, G4 (preposition + noun) and G5 (adjective + preposition) had the power to differentiate across the three band scores.

Keywords: corpus analysis, grammatical collocations, IELTS test-takers, lexical collocations, speaking assessment

OPEN ACCESS

Edited by:
George Waddell,
Royal College of Music,
United Kingdom

Reviewed by:
Yangyu Xiao,
The Chinese University of Hong Kong, China
Musa Nushi,
Shahid Beheshti University, Iran

****Correspondence:***
Masoomeh Estaji
mestaji74@gmail.com

Specialty section:
This article was submitted to Assessment, Testing and Applied Measurement, a section of the journal Frontiers in Education

Received: *02 December 2021*
Accepted: *02 February 2022*
Published: *22 March 2022*

Citation:
Estaji M and Montazeri MR (2022) The Representation of Collocational Patterns and Their Differentiating Power in the Speaking Performance of Iranian IELTS Test-Takers. Front. Educ. 7:827927. doi: 10.3389/feduc.2022.827927

INTRODUCTION

Vocabulary knowledge is an indispensable feature in organizational competence as a subcategory of language competence (Bachman and Palmer, 1996). When it comes to vocabulary learning, not only is the lexical size considered, but importance is also placed on the lexical depth, which is germane to the words relationship (Caro and Mendinueta, 2017). Over the past few years, the topic of formulaic sets has experienced a mushrooming interest, being spotlighted particularly by corpus linguists (e.g., Sinclair, 1991; Biber et al., 1999; Hyland, 2008; Paquot and Granger, 2012; Macis and Schmitt, 2016; Paquot, 2018). Studies pertinent to formulaic sequences have been in vogue for the past four decades. This highlights the importance of such pre-constructed units available to language users of which language is mostly composed (Pawley and Syder, 1983; Sinclair, 1991). Formulaic sequences, i.e., chunks of words with different lengths (Xu, 2018) entail some categories such as phraseological units (Gläser, 1986), collocations, and idioms. As an essential part of lexical cohesion in a language, collocation consists of lexical items which tend to go-together (Halliday and Hasan, 1976).

In L2 production, the speakers' correct use of collocations improves the collocational competence both in writing and speaking (Xu, 2018). This is provided that they use collocations appropriately as there exist certain types of collocations—or as Granger (1998) calls them "significant collocations"—which could differentiate between lower-level and higher-level language learners. That is, lower-level students tend to underuse native-like types of collocations caused by "an underdeveloped sense of salience and of what constitutes a significant collocation" (Granger, 1998, p. 6). Another example of the differentiating power of collocations among learners with different levels is the greater use of highly restricted collocations (i.e., collocations with limited number of substitutions for the headword such as "lunar calendar") on the part of higher-level students (Xu, 2015, 2018). Collocations with phrasal verbs are more challenging for elementary English learners (Xu, 2015). However, most of these studies have mainly scrutinized collocations including one or some types of syntactic patterns, for example, verb + noun collocations (e.g., Bahns and Eldaw, 1993; Laufer and Waldman, 2011) or premodifier-noun collocations (e.g., Durrant and Schmitt, 2009; Granger and Bestgen, 2014), failing to consider all types of collocations altogether so as to better differentiate collocations among language learners or test-takers with different levels of proficiency.

The apropos use of collocations is an emblem of native-like communicative competence (Keshavarz and Salimi, 2007) as it is seen in the native-speakers' propensity to use such chunks in lieu of other words in many situations (Wray, 2000). Thus, achieving this native-like competence and possessing the mastery of a new language would not be feasible without using formulaicity, prefabricated patterns, and collocations in particular (Wray, 2000; Nesselhauf, 2005; Nizonkiza, 2011). Moreover, formulaicity is conducive for faster language processing, leading to the production of language with minimized cognitive load (Bygate, 1987). This, consequently, leads to more improved oral proficiency (e.g., Hsu and Chiu, 2008). Not only have the teaching of collocations and raising collocational competence been accentuated in EFL and ESL classes, but also its benefits accrue to the test-takers. The utilization of collocations goes hand in hand with all of the four skills, i.e., listening, speaking, reading, and writing. Thus, collocation use augments the general level of language proficiency (e.g., Bonk, 2000; Hsu, 2007) which can be assessed through tests. Tests, particularly high-stakes ones, are largely influenced by the use of such native-like sequences.

To illustrate, the International English Language Testing System (IELTS), as one of the most universally recognized high-stakes tests, stresses the importance of lexical resources, in which collocation use is mandatory for obtaining higher scores. Among the four skills tested in IELTS, speaking is one of the most challenging one. Generally, an oft-raised issue in making hindrance for language test-takers has been the speaking portion of the proficiency tests. This is because speaking is certainly an important and perhaps the most puzzling skill among these four skills (Lazaraton, 2014). Due to the dearth of exposure to the target language, language learners have poor speaking ability "especially regarding fluency, control of idiomatic expressions, and understanding of cultural pragmatics" (Shumin, 2002, p. 204). As an efficacious way to ameliorate the test-taker's speaking proficiency, the use of collocations has proved to augment their oral proficiency scores in high-stakes tests (Hsu and Chiu, 2008). In IELTS speaking test, apropos use of collocations eventuates in band scores 7 or above according to the speaking band descriptors (public version).

Using an "evidence-based approach," a corpus-driven study offers invaluable and authentic insights regarding the data the researchers are seeking to analyze (Hyland, 2006). To this end, corpus linguists collect a corpus of representative samples of naturally occurring collocations in spoken or written texts. Previous research has mostly focused on collocation in EFL or ESL contexts or general English corpus, yet few studies have examined collocations in the testing domain, particularly when it comes to the use of grammatical and lexical collocations in important high-stakes tests such as IELTS.

LITERATURE REVIEW

Although many scholars have introduced different definitions of collocations, never have they arrived at a consensus (Mel'cuk, 1998; Wray, 2002; Nesselhauf, 2005; Xu, 2018). However, one important definition has been that of Benson et al. (2010), who referred to collocations as "fixed, identifiable, non-idiomatic" combinations which are used repetitively in a language (p. 19). Despite the multifaceted essence of collocation and formulaic language, researchers agree that collocations are divided into two chief categories, namely grammatical and lexical collocations. Grammatical collocations are collocations between one open class (noun, adjective, and verb) and one close class (a preposition or a grammatical structure) such as "depend on" (Benson et al., 2010). Lexical collocations, on the other hand, contain two content words, both of which are open classes. Lexical collocations include nouns, adjectives, verbs, and adverbs.

Regarding the two types of collocations, grammatical collocations are of eight types while there exist seven types of lexical collocations (Benson et al., 2010). Having added three more collocation categories, Xu (2015) maintained that collocations with phrasal verbs are more complex types of collocations, which call for more language competence. According to Benson et al. (2010), the syntactic patterns of grammatical collocations include: noun + preposition, noun + to + infinitive, noun + that-clause, preposition + noun, adjective + preposition, predicative adjective + to + infinitive, adjective + that-clause, and collocational verb patterns.[1] Based on Xu's (2015) syntactic patterns of lexical collocations, these types of collocations include the following structures: adjective + noun, adverb + adjective, adverb + verb, noun + noun, noun + of + noun, noun + verb, verb + noun, phrasal verb + adverb, noun + phrasal verb, and phrasal verb + noun.

Although there exist different types of grammatical and lexical collocations, previous studies proved that not all types are used

[1]According to Benson et al. (2010), there exist 19 English verb patterns in this category, whose discussion would be irrelevant and time-consuming in this study. However, among these verb patterns, verb + preposition category was chosen in the current study for the analysis of collocations used by IELTS test-takers.

with the same frequency due largely to the difficulty certain types of collocations could pose for language learners (Palmer, 1933; Bahns and Eldaw, 1993; Granger, 1998; Gitsaki, 1999; Nesselhauf, 2005). Additionally, what might impinge on language learners' collocational competence is their underuse, overuse, or misuse of certain types of collocations. This, consequently, would differentiate lower-level L2 learners with more advanced ones or native-speakers. Regarding the underuse of certain types of collocations, Laufer and Waldman (2011), say, found that L2 learners made use of less number of verb + noun collocations (5.9%) compared to native-speakers (10%). Similarly, Granger (1998) concluded that non-native speakers' use of collocations with intensifying adverbs was less than native speakers'. On the other hand, non-native speakers tend to overuse certain types of collocations in situations where more precise meaning is required (e.g., Shih, 2000; Durrant and Schmitt, 2009). Finally, the production of deviant collocations is another problem, differentiating between lower-level and higher-level L2 learners. It has been proven that the chief reason of such misuses is the negative L1 transfer based on the learners' reliance on their first language to produce L2 collocations, resulting in semantically inaccurate collocations (Kormos, 2006; Namvar, 2012; Xu, 2015).

Another proved belief about collocations is their position in the continuum of idiomaticity. The position of collocations is between free combinations such as "under the table" and fixed, figurative combinations such as "under the weather" (Howarth, 1998; Xu, 2018). Therefore, collocations are semifixed structures that contain possibly interchangeable lexical items, but with particular limitations (Xu, 2018). For instance, "make an effort" could only be substituted by "put forth an effort", and not with other elements.

Collocational competence, or the lack thereof, plays an important role in both language production and comprehension. The significance of collocational competence as one of the pivotal components in the four skills and language proficiency, in general, has been highlighted by a plethora of research (e.g., Bonk, 2000; Hsu, 2007; El-Dakhs, 2015). Collocational studies have been on the march since collocation is a widespread phenomenon in languages (El-Dakhs, 2015), helping learners improve efficiency in both comprehension and production by reducing the cognitive load throughout language production (Bygate, 1987).

As authentic language is interwoven with the utilization of collocations, in language testing, particularly spoken evaluations of the test-takers too, the use of collocations has proved to be of considerable use (e.g., Hsu and Chiu, 2008; Attar and Allami, 2013; Keshavarz and Taherian, 2018). The significance of collocations in spoken evaluations of language stems from speech-processing theories in spontaneous speaking tests. These theories substantiate the logical nexus between collocational ability and the construct of L2 oral proficiency, highlighting the role of formulaicity in speech production (Bygate, 1987; Levelt, 1999; Kormos, 2006; Xu, 2015, 2018). To illustrate, Kormos (2006) stated that formulaic expressions are crucially necessary in spoken language, and most of our utterances are replete with such expressions. What advanced L2 speakers mostly resort to is the store of the lexicon rather than L2 declarative rules because their declarative knowledge has changed to procedural knowledge (Kormos, 2006). That is why the memorization of formulaic expressions can pave the way for such a transformation as it creates limited attentional resources.

When it comes to language tests, the role and measurement of collocations has proved to be of high value (Xu, 2015). Drawing on speech production theories, Xu (2015) formulated the new construct of spoken collocational competence. Xu's (2015) study on 60 adults Chinese L2 learners' speech output in an oral English test made it clear that generally the production of accurate, complex, and fluent collocations was the part and parcel of more advanced test-takers' speech. Based on this construct and the results of his study, Xu (2015, 2018) maintained that language assessors need to pay closer attention to the measurement of collocations in language tests.

In recent years, there has been an increasing number of studies on formulaicity and collocations. Collocations have been studied on the four language skills to determine the impact of collocational knowledge on language proficiency. What is common among most of these studies is the fact that proficient L2 language learners have a better performance than lower-level learners in collocational tests (e.g., Bahns and Eldaw, 1993; Zughoul and Abdul-Fattah, 2003; Keshavarz and Salimi, 2007; Namvar, 2012). For instance, having studied 62 Taiwanese EFL students at a university of science and technology to explore the students' utilization of lexical collocations in online writing, Hsu (2007) found positive correlations between the learners' frequency of lexical collocations and their online writing scores. He also found that the variety of lexical collocations was positively correlated with their writing scores.

As to the role of collocations in speaking, Boers et al. (2006) studied the impact of formulaic sequences in general (such as idioms and collocations) on L2 oral proficiency. They studied 32 Belgian college students, 11 of whom were in the experimental group and were made cognizant of formulaic sequences, and the rest 15 were taught in the control group, with the traditional method of instruction. The findings of Boers et al. (2006) demonstrated the better performance of the experimental group in oral proficiency. Similarly, Sarvari et al. (2016) researched 60 EFL learners who were divided into the experimental and control groups. They found that the experimental group, who were taught collocations, were more fluent than the control group, who were taught solely single lexical items, in the IELTS speaking test.

However, there are obvious differences in the research results on the relationship between collocational knowledge and language proficiency. For one thing, the term "collocation" has been defined differently by different researchers, so when a researcher defines collocation in a phraseological sense (e.g., Nesselhauf, 2003), the essence of such a study would differ from another researcher who defines it in a frequency-based sense (e.g., Cowie, 1994). In addition, various researchers have based their collocational studies on a specific collocational pattern. For instance, Bahns and Eldaw (1993) focused on the lexical collocational patterns for investigating the importance of teaching collocations in EFL classes. They studied 58 advanced German learners, 34 of whom were asked to complete a translation task and the rest were asked to complete a cloze task. However, Alsulayyi (2015) attempted to study the grammatical

collocational patterns to ascertain the cause of collocational errors among Arab undergraduate students. He compared the use of grammatical collocations among Arab students majoring in English. The results indicated that noun + preposition and adjective + preposition patterns were the most erroneously used collocations, the crux of which lay in L1 transfer.

None of these studies focused on the test-takers' use of collocations and the frequency of different collocational patterns they represent to analyze their test scores in high-stakes oral proficiency tests such as IELTS. Most of the previous studies, however, researched collocations in non-testing contexts solely based on a limited number of syntactic patterns, disregarding other types of collocations particularly the grammatical ones. To bridge the mentioned gap, this study inspected the most typical grammatical and lexical collocational patterns and their differentiating power in the speaking section of IELTS. To this end, the present study sought to answer the following questions.

1. Which collocational patterns are represented in IELTS test-takers' speaking with various band scores?
2. To what extent do the collocational patterns differentiate among test-takers across various band scores in IELTS speaking?

METHODOLOGY

The Corpus

In the current study, a corpus consisting of 60 recordings of the IELTS speaking mock tests was gleaned. The speech samples were collected from different IELTS centers in Tehran, which administer IELTS mock tests regularly each year. The 60 recordings, approximately 13 h, were selected randomly from among 90 speaking samples, and then through purposive sampling, samples with band scores 6, 7, and 8 were chosen. Purposive sampling was also chosen to select the most reliable IELTS institutes in Tehran through consultation with IELTS experts. As is seen in **Table 1**, each band score involved 20 recordings to be analyzed, and the corpus contained approximately 110,000 words. The mock exams were held in 2020.

Instrumentation

In the current study, different instruments were used to collect data, including (a) IELTS mock tests and (b) IELTS speaking tests.

IELTS Mock Tests

IELTS mock tests held in 2020 were used for the quantitative phase of the study. An IELTS mock test includes four components: Listening, speaking, reading, and writing. The listening and speaking tests are the same for all of the candidates, who take the mock test. However, the reading and writing tests of the academic module are different from the reading and writing tests of the general training module. The reading test consists of three texts with a total of 40 questions. There exist different types of questions, including Multiple choice, identifying information, identifying the writer's views, matching type questions (e.g., matching headings), completion type questions (e.g., flow-chart completion), and short answer questions.

The listening test includes four sections and a total of 40 questions. The first two sections are germane to social contexts, while sections 3 and 4 are concerned with the academic contexts. Half of the sections are set in dialogues between 2 or more people, but the other two sections are set in monologues. The speaking test entails three main parts. The first part questions revolve around general and familiar topics. In the second part, the test-takers are given a cue card with a topic and some suggestions about which they are to talk for two minutes. Finally, in part 3, the examiner asks more detailed questions regarding the topic raised in part 2. The writing test involves two tasks. In task 1 of academic writing, the test-takers are asked to describe a piece of visual information, be it a graph, a table, or a chart. Task 1 of the general training module asks the test-takers to write a letter as a response to a specific situation. Task 2 of both modules requires the test-takers to respond to a problem, provide a solution, compare and contrast different ideas, and evaluate and challenge arguments. The centers administering the mock tests either use the retired or expired versions of the test, whose reliability and validity have already been confirmed, or establish the reliability and validity of the compiled ones through several pilot tests and content analysis of Iranian certified examiners.

TABLE 1 | Details of the Samples.

Speaking band scores	Number of words	Length of recording (h)
6	35,800	4.2
7	36,680	4.3
8	37,820	4.4

IELTS Speaking Tests

IELTS speaking test is an encounter between an examiner and a candidate that is designed to take between 11 to 14 min. It entails three main sections, each of which has a specific function. Part 1 which is called the introduction takes 4 to 5 min, revolving around questions concerning familiar topics such as hobbies, interests, and jobs. The second part, which is also called individual long turn, takes 3–4 min. This part entails a verbal prompt on a card about which the candidates are asked to talk for 2 min. They have one minute to prepare before talking for 2 min. Part 3 is called a two-way discussion in which the examiner asks more abstract concepts related to the topic of part 2. It takes 4–5 min. Scoring of these tasks is reported based on four criteria: Fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation, each varying from 1 to 9 IELTS bands.

Data Collection and Analysis Procedure

Before the instigation of the study, to ensure the psychometric quality of the study and richness of the data (i.e., the richness of the utilization of collocations in each sample), a speech sample containing 12 IELTS interviews between an examiner and a candidate was scrutinized. Besides, the issue of confidentiality was considered. That is, gaining the consent of the institutes

which provides the researchers with the speech samples was the primary concern of the study. Afterward, a corpus entailing 60 IELTS speaking tests recordings, which were chosen based on the test-takers' speaking section band scores, was gleaned. That is, the IELTS mock test-takers' band scores of the speaking sections were scrutinized, and those samples with the band scores 6, 7, and 8 were selected. Then around 13 h of recordings, which contained around 110,000 words, were transcribed for further analysis utilizing the application Nuance Dragon Professional Individual. Subsequently, the samples were proofread by the researchers. Finally, the examiner's speeches were italicized. A sample of the IELTS speaking test is shown in **Appendix A**.

Having adopted a corpus-driven approach, the frequency of collocations was examined after the manual extraction. That is, the frequency of 10 lexical collocations and four grammatical collocations was analyzed. The lexical and grammatical collocations were coded using Xu's (2015) (**Appendix B**) and Benson et al.'s (2010) (**Appendix C**) coding scheme, respectively. To measure the collocational strength in the English language and judge the acceptability of collocations, the use of a reference native corpus was mandatory (Paquot and Granger, 2012). Such reference corpora are quite large in size, including widely representative samples of the language used in the real and naturally occurring speech (Paquot and Granger, 2012). To do so, the present study made use of the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). Therefore, the frequency of each collocation was determined with regard to the BNC and COCA reference corpora. Moreover, having analyzed the frequency of each collocation, chi-square analysis was run to determine the differentiating power of the collocations across the three band scores.

RESULTS

Reliability Analyses of Collocation Extraction

To ensure the accuracy of the collocation extraction, the precision and recall equations were taken into account. These equations, which are used to evaluate the NLP applications (Futagi et al., 2008), help the researchers identify and include the most accurate collocations for further analysis. The following equations show the formulae for calculating precision and recall.

$$precision = \frac{\left|Collocation_{extracted} \cap Collocation_{true}\right|}{\left|Collocation_{extracted}\right|}$$

$$Recall = \frac{\left|Collocation_{extracted} \cap Collocation_{true}\right|}{\left|Collocation_{true}\right|}$$

Through analyzing the pilot study sample, the first coder identified 572 collocations. The second coder analyzed this sample and identified 580 collocations. Then the precision and recall equations were utilized to check the accuracy of collocations. The coders also discussed the differences and resolved the discrepancies by referring to the reference corpora (i.e., BNC and COCA). Having computed the precision and recall, the researchers reached the high precision of 0.98 and the recall of 0.97 which substantiate the accuracy of collocation extraction. This means that 98% of the collocations that the first coder had detected were true collocations, and solely 3% of the true collocations embedded in the speaking transcriptions were missed by the first coder.

To ensure that the coded data are reliable, the coding of the different types of collocations was done by two coders. The inter-coder consistency was calculated for 20% of the sample (five cases from each band, $N = 15$) before proceeding to the rest of the coding, using the Kappa formula (**Table 2**).

The result of the Kappa agreement test ($\kappa = 0.89$, SE = 0.017, $p = 0.000 < 0.05$) shows almost perfect agreement (values between 0.81 and 1.00 are considered almost perfect), ensuring that the data obtained from the two coders would have the acceptable consistency.

Results for the Representation of Collocational Patterns in IELTS Test-Takers' Speaking

The first question of the present study was germane to the frequency of different types of lexical and grammatical collocations across the three band scores. Based on the results, it was found that 2,178 collocations were of lexical types, while 1,074 collocations were grammatical ones. Overall, 3,252 collocations were found in the corpus. The sub-corpus of band score 6 included 570 collocations, while more collocations were utilized by the test-takers of band scores 7 and 8 (1,087 and 1,595 collocations, respectively). To analyze the frequency of each type, a frequency analysis of the types of lexical and grammatical collocations found in the analyzed corpus was done. To do so, the number of collocations used by the participants within each band score was counted. **Table 3** below presents the descriptive statistics of the lexical collocations used in each band score.

As reported in **Table 3**, the two most frequently used types of lexical collocations were L1 (adjective + noun) and L7 (verb + noun), followed by L4 (noun + noun), L2 (adverb + adjective), L3 (adverb + verb), L5 (noun of noun), and L10 (phrasal verb + noun). Moreover, L8 (phrasal verb + adverb), L9 (noun + phrasal verb), and L6 (noun + verb) had the overall lowest frequency of uses. To get a better picture of the pattern, a multiple boxplot of the frequencies in each band score was generated (**Figure 1**).

As illustrated in **Figure 1**, in all types of lexical collocations, the frequency of use increased with the increase in the band score. Moreover, the pattern seems still throughout the band sores, with L1 (adjective + noun) and L7 (verb + noun) being the most frequent ones. The following excerpts show the use of these two types of collocations.

Excerpt 1: They should have eh different eh characteristics for start their own business, they should be eh very knowledgeable, and they should have **interpersonal skills** and also manage. **(L1 (adjective + noun)/ band score 7 test-taker)**

Excerpt 2: And eh when eh weather is rainy eh and eh especially in accidents or eh rainy day eh it leads eh to **increase**

TABLE 2 | Kappa agreement: inter-coder consistency.

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	0.89	0.01	54.52	0.00
No. of valid cases		355			

[a]*Not assuming the null hypothesis.*
[b]*Using the asymptotic standard error assuming the null hypothesis.*

eh eh **traffic congestion**. **(L7 (verb + noun)/ band score 6 test-taker)**

After L1 (adjective + noun) and L7 (verb + noun), L4 (noun + noun), L2 (adverb + adjective), L3 (adverb + verb), L5 (noun of noun), and L10 (phrasal verb + noun) were the most frequently used types, respectively. The following excerpts show the use of these types.

TABLE 3 | Descriptive statistics for the frequency of lexical collocations by band scores.

	Band	N	Minimum	Maximum	Mean	SD
L1	6	20	0	12	4.50	3.08
	7	20	6	28	13.75	6.65
	8	20	10	37	21.20	6.37
	Total	60	0	37	13.15	8.82
L2	6	20	0	5	1.20	1.64
	7	20	0	12	2.25	3.52
	8	20	0	10	3.30	2.25
	Total	60	0	12	2.25	2.69
L3	6	20	0	2	0.40	0.68
	7	20	0	5	1.65	1.75
	8	20	0	7	2.10	2.15
	Total	60	0	7	1.38	1.77
L4	6	20	0	7	2.10	1.65
	7	20	0	10	4.95	2.91
	8	20	2	19	7.35	4.20
	Total	60	0	19	4.80	3.74
L5	6	20	0	1	0.25	0.44
	7	20	0	3	1.05	0.82
	8	20	0	6	2.15	1.42
	Total	60	0	6	1.15	1.24
L6	6	20	0	1	0.15	0.36
	7	20	0	2	0.50	0.76
	8	20	0	2	0.55	0.68
	Total	60	0	2	0.40	0.64
L7	6	20	3	12	7.30	3.06
	7	20	4	24	12.60	6.21
	8	20	9	29	16.00	5.28
	Total	60	3	29	11.97	6.12
L8	6	20	0	0	0.00	0.00
	7	20	0	0	0.00	0.00
	8	20	0	2	0.30	0.57
	Total	60	0	2	0.10	0.35
L9	6	20	0	0	0.00	0.00
	7	20	0	1	0.15	0.36
	8	20	0	1	0.40	0.50
	Total	60	0	1	0.18	0.39
L10	6	20	0	2	0.15	0.48
	7	20	0	4	0.85	1.26
	8	20	0	6	2.15	1.84
	Total	60	0	6	1.05	1.54

Excerpt 3: Uh but some of them are really eh motivating, for example, some uh **talent shows** are really motivating for eh young children and eh **(L4 (noun + noun)/band score 6 test-taker)**

Excerpt 4: Well, I had this eh **excruciatingly painful** burden of my parents on me because I was as good at engineering as I was good at mathematics. **(L2 (adverb + adjective)/band score 8 test-taker)**

Excerpt 5: First of all, I have to say that computers have **drastically changed** our lives. **(L3 (adverb + verb)/band score 8 test-taker)**

Excerpt 6: Of course. Uh, I have a lot of **circle of friends** and they I mean they have- I definitely put the- put them on X mm due to their behavior but also my life. **(L5 (noun of noun)/band score 7 test-taker)**

Excerpt 7: You shouldn't uh **bottle up your feelings** and which eh gradually eventuate in some disaster. **(L10 (phrasal verb + noun) band score 7 test-taker)**

Finally, L8 (phrasal verb + adverb), L9 (noun + phrasal verb), and L6 (noun + verb) were the least frequents. The following excerpts show the use of these types of lexical collocations produced by the test-takers.

Excerpt 8: I try to just eh keep eh normal eh sleeping pattern-sleeping pattern and uh not **staying up** much **late**. **(L8 (phrasal verb + adverb)/band score 8 test-taker)**

Excerpt 9: I play the guitar, and I think my whole **life revolves around** the music. **(L9 (noun + phrasal verb)/band score 8 test-taker)**

Excerpt 10: Well, probably those kinda uh cool **titles** that **hook** you. They really can captivate your attention. **(L6 (noun + verb)/band score 8 test-taker)**

Concerning grammatical collocations, only four types were inspected. The rationale for the selection of these four types has been discussed in the delimitations of the study. **Table 4** presents the descriptive statistics for the frequency of use in this type of collocation.

Table 4 shows that the most frequently used type of grammatical collocations was G8 (verb + preposition) and the least G1 (noun + preposition). **Figure 2** illustrates the pattern in a multiple boxplot.

Figure 2 also shows similar patterns throughout the band scores considering the frequency of use, G8 (verb + preposition) having the highest, followed by G4 (preposition + noun) and G5 (adjective + preposition). Moreover, G1 (noun + preposition) had the lowest frequency in all the three band scores. The pattern also exists in the other two types of grammatical collocations. The following examples illustrate the utilization of grammatical collocations by the test-takers.

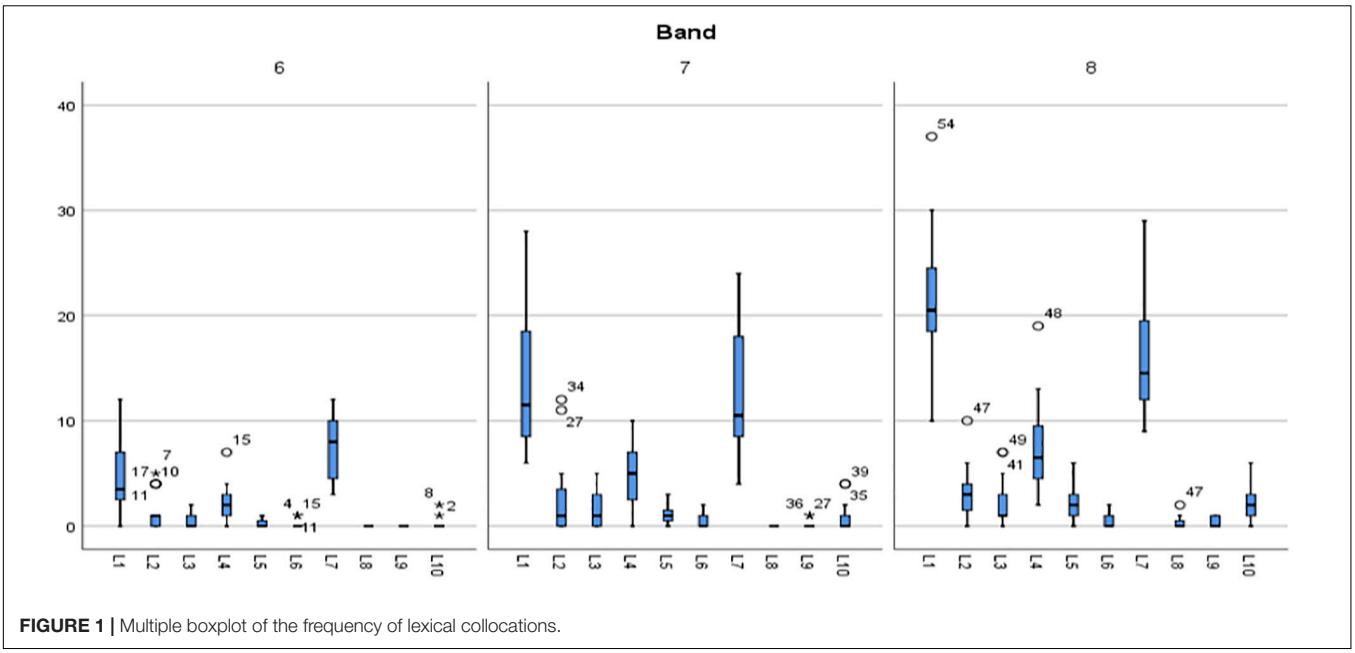

FIGURE 1 | Multiple boxplot of the frequency of lexical collocations.

Excerpt 11: Yes, em . . . to my mind, Iranians are very welcoming and they are hospitable, and I think eh as- our country has a **reputation for** its natural attraction. **(G1 (noun and preposition) band score 7 test-taker)**

Excerpt 12: I guess, um reliable information cannot necessarily be found **on the internet**. **(G4 (preposition + noun) band score 8 test-taker)**

Excerpt 13: Eh mmm as I know, in Iran, eh people are eh... You are so **friendly to** foreigner eh they invite them maybe to their home. **(G5 (adjective + preposition) band score 6 test-taker)**

Excerpt 14: Mmm I think eh, it **depends on** the situation but eh mostly the things that you need eh very strong concentration, for example, doing some... kind of eh drawing. **(G8 (verb + preposition) band score 6 test-taker)**

Results for the Collocational Patterns Differentiating Among IELTS Test-Takers Across Various Band Scores

To capture the differentiating power of the collocations in various band scores, series of chi-square tests were run on each type of collocations. **Table 5** shows the results for lexical collocations and **Table 6** for the grammatical ones.

Based on the results of the chi-square tests, it was found that there did exist certain types which can differentiate test-takers across band scores 6, 7, and 8. As it is evident from **Table 5**, half of the lexical collocations' types can differentiate among the three band scores. They include L1 (adjective + verb) ($\chi^2_{(52)} = 72.8$, $p = 0.03 < 0.05$), L5 (noun of noun) ($\chi^2_{(10)} = 33.4$, $p = 0.000 < 0.05$), L8 (phrasal verb and adverb) ($\chi^2_{(4)} = 10.91$, $p = 0.028 < 0.05$), L9 (noun and phrasal verb) ($\chi^2_{(2)} = 10.91$, $p = 0.004 < 0.05$), and L10 (phrasal verb and noun) ($\chi^2_{(12)} = 32.76$, $p = 0.001 < 0.05$). The following excerpts are indicative of the use of these five collocations among the three band scores. With regard to the use of L1 (adjective + noun), the following excerpts are presented.

Excerpt 15: I actually rarely eh use **public transportations** like bus, subway and uh. **(band score 6 test-taker)**

Excerpt 16: The majority of the time uh we are in faced with eh scarcity and- and water shortage, but I hope in the not- in the **foreseeable future** eh take action to tackle the problem. **(band score 7 test-taker)**

Excerpt 17: Uh, not to mention, I would say it is also famous for uh- for the **heavy traffic** that we face every day here, which makes it actually eh to some extent difficult to < get > around. **(band score 8 test-taker)**

TABLE 4 | Descriptive statistics for the frequency of grammatical collocations by band scores.

	Band	*N*	Minimum	Maximum	Mean	SD
G1	6	20	0	4	.85	1.26
	7	20	0	6	1.45	1.60
	8	20	0	5	1.50	1.53
	Total	60	0	6	1.27	1.48
G4	6	20	1	8	3.55	4.89
	7	20	0	11	4.70	8.32
	8	20	3	12	7.45	4.99
	Total	60	0	12	5.23	2.93
G5	6	20	0	5	1.80	2.37
	7	20	1	7	2.65	3.08
	8	20	1	10	5.00	5.36
	Total	60	0	10	3.15	2.31
G8	6	20	2	14	6.35	9.60
	7	20	2	15	7.65	8.55
	8	20	3	16	10.25	11.67
	Total	60	2	16	8.08	3.50

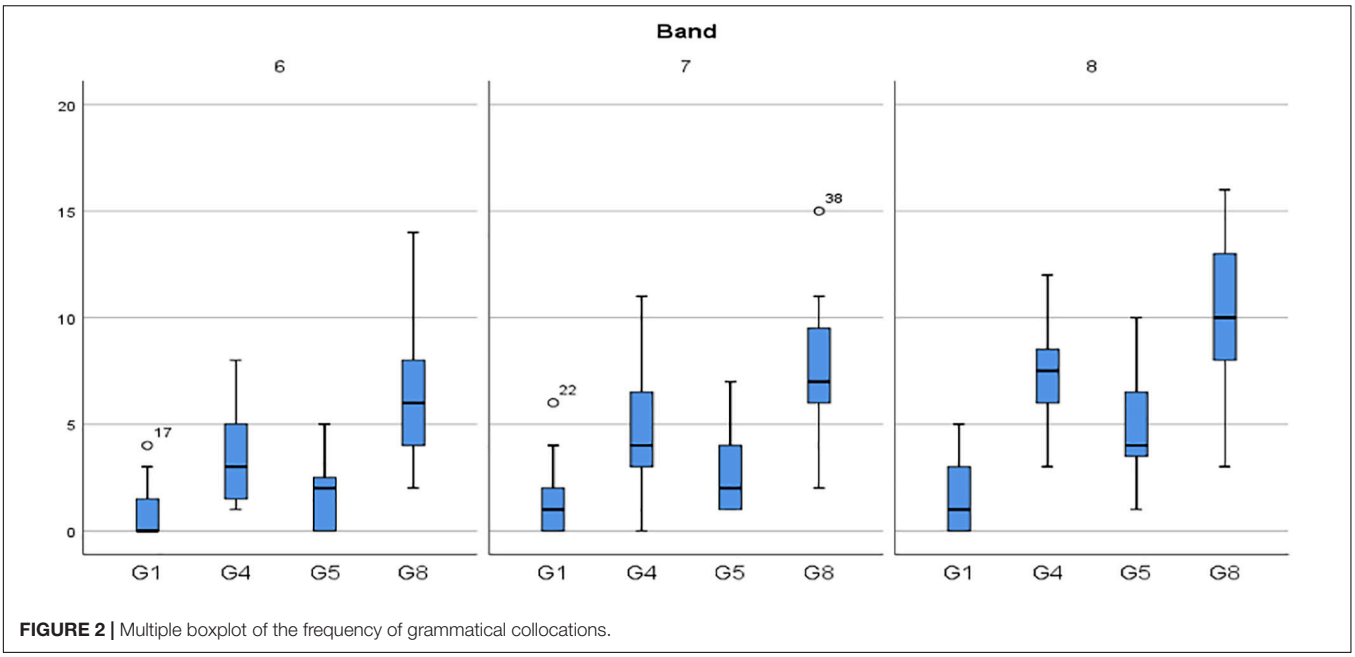

FIGURE 2 | Multiple boxplot of the frequency of grammatical collocations.

The following examples show the use of L5 (noun of noun) across the three groups.

Excerpt 18: It's very eh . . . useful because it makes eh . . . it- it makes eh income for me. And eh . . . it makes income for me and eh it eh eh eh sour- I think it's- it's good eh **source of revenue** for our lives. **(band score 6 test-taker)**

Excerpt 19: Maybe they don't have enough information, or even if they have prepared enough em **piece of information**, it might be nervous to stand in front of so many eyes and they are just staring at you. **(band score 7 test-taker)**

Excerpt 20: And when I exercise, I can feel it in every inch of myself and in every **ounce of blood** that I can think better. **(band score 8 test-taker)**

The use of L8 (phrasal verb and adverb) was solely detected in band score 8, as the excerpt 21 illustrates:

Excerpt 21: Uh, but not everybody. I think uh possesses the eh like ability and capability to **get along well** with English.

The use of L9 (noun and phrasal verb) was only found in band scores 7 and 8. The following examples show the use of this type.

Excerpt 22: He helped us eh about subjects because he is eh very eh resourceful, and whenever we co- come up eh eh to- **a problem comes up**, he helps us a lot. **(band score 7 test-taker)**

Excerpt 23: Their eh potential benefits and eh effect eh might not be eh clear or touchable now for the people but as **time goes by** actually through investing, these can actually uh mm later surprise people. **(band score 8 test-taker)**

Finally, excerpts 24, 25, and 26 show the use of L10 (phrasal verb and noun) across the three band scores.

Excerpt 24: When I was a little girl eh loved eh special **shoes** that eh as a little girl I- I hadn't the uh- the permission to- to **put** them(on). **(band score 6 test-taker)**

Excerpt 25: To this- so this uh you have to be able to **deal with** different kinds of **problems** or **issues** that you will face in the future of your business. **(band score 7 test-taker)**

Excerpt 26: When you cannot make ends meet, you have to eh **look for** a secondary **job** eh that eh- that partially helps you eh- your life to run sm- more smoothly. **(band score 8 test-taker)**

Considering the grammatical collocations, as it is evident from **Table 6**, two out of four types of the collocations could differentiate among the three band scores. They include G4 (preposition + noun) ($\chi^2_{(24)} = 38.49$, $p = 0.031 < 0.05$) and G5 (adjective + preposition) ($\chi^2_{(20)} = 46.6$, $p = 0.001 < 0.05$). The

TABLE 5 | Chi-square test on the type of lexical collocations differentiating the band scores.

	χ^2	df	Asymptotic significance (2-sided)
L1	72.80	52	0.03
L2	26.25	18	0.09
L3	17.48	12	0.13
L4	34.56	24	0.07
L5	33.40	10	0.00
L6	6.02	4	0.19
L7	45.00	40	0.27
L8	10.90	4	0.02
L9	10.90	2	0.00
L10	32.75	12	0.00

TABLE 6 | Chi-square test on the type of grammatical collocations differentiating the band scores.

	χ^2	Df	Asymptotic significance (2-sided)
G1	8.53	12	0.74
G4	38.47	24	0.03
G5	46.60	20	0.00
G8	34.19	28	0.19

following excerpts illustrate the use of G4 (preposition + noun) across the three levels:

Excerpt 27: Uh, this is a difficult question because I usually can't remember my dreams **at night,** but eh eh sometimes eh I usually eh have some dreams about immigration to Australia. **(band score 6 test-taker)**

Excerpt 28: Uh the most popular one I think is the uh subway because it super X uh the other transport systems uh in our town uh and also it's eh so fast and **on time**. **(band score 7 test-taker)**

Excerpt 29: But parents have to educate themselves. Is is- it is not something that you do that just **by nature**. **(band score 8 test-taker)**

Overall, the results indicated that there existed such differences among the test-takers concerning their use of certain types of collocations. Hence, the null hypothesis, which claimed that the collocational patterns do not differentiate among test takers across various band scores in IELTS speaking, was rejected.

DISCUSSION

This study aspired to analyze the representation of collocational patterns and their differentiating power across three band scores in the IELTS speaking test. The first question of the present study was pertinent to the frequency of lexical and grammatical collocations. According to the research results, the use of both types of collocations with all of the syntactic types was observed in each band score. In particular, the production of collocations—both the lexical and grammatical types—increased with the increase of the test-takers' speaking band scores. This highlights that the more advanced the students became, the more collocations they used in their speech.

This finding shows that lower-level test-takers still lack adequate collocational competence to convey their messages through varied types of collocations. This finding is in tune with the literature in that the number of collocation use tends to increase with the increase of the learners' proficiency level (Bahns and Eldaw, 1993; Laufer and Waldman, 2011). This finding must be interpreted with care, though, since the literature has shown that even advanced students tend to make (sometimes more) mistakes. By way of example, some studies (e.g., Obukadeta, 2014; Men, 2018) suggested that "collocation lag" (Men, 2018, p. 2) does exist as the English language learners' proficiency level increases and they tend to encounter more collocational challenges.

The findings indicated that out of the ten patterns of the lexical collocations, the most common patterns were L1 (adjective + noun), L7 (verb + noun), and L4 (noun + noun) among all the three groups. This finding was quite predictable since the literature has substantiated that the majority of collocations produced by English learners include L7 (verb + noun), L1 (adjective + noun), and L4 (noun + noun) (e.g., Xu and Xi, 2010; Xu, 2015). These findings are because such collocations carry the most important information while producing language (Xu, 2015; Men, 2018). In line with the literature (e.g., Mei, 1999; Xu, 2015), the greater frequency of these three syntactic patterns show the role of syntactic transfer in which Iranian IELTS test-takers applied "L1 rules for encoding an L2 phrase" (Kormos, 2006, p. 175). That is, by using such structures, they could lighten their cognitive load, thereby facilitating their L2 oral language production (Kormos, 2006). Although these three types are the commonest types of collocations, it does not mean that error-free production of these combinations is easy. However, research proved that these types carry the most frequent sources of challenge for English L2 learners. To illustrate, regarding L7 (verb + noun), Gitsaki (1999) concluded that verb+noun combinations were the most difficult types of collocations which were acquired at later stages of learning. Moreover, based on the Chinese Learner English Corpus, L1 (adjective + noun) and L4 (noun + noun) were the second and third most deviant collocation types (Gui and Yang, 2003).

The present study also found that L1 (adjective + noun) outnumbered L7 (verb + noun) in general (save for band score 6), although the difference was not significant. This is in contrast with some other studies such as Namvar's (2012) and Xu's (2015). This could be justified by what Johansson and Hofland (1989) stated. They held that L1 (adjective + noun) and L4 (noun + noun) collocations are the most frequently used types by native English speakers. Gitsaki's (1999) study also demonstrated that L1 (adjective + noun) was the easiest combination, being acquired at an early stage of collocational knowledge development.

The paucity of the use of other types of collocations such as L2 (adverb + adjective) and L3 (adverb + verb), especially on the part of band score 6 test takers illustrates their lack of knowledge concerning these collocations. The use of such syntactic patterns increased with the increase of band scores; however, the test-takers still tended not to use these patterns much. This highlights that adverbial collocations are not as common as L1 (adjective + noun), L7 (verb + noun), and L4 (noun + noun), which carry the most important information in oral communication (Xu, 2015). This finding aligns with the literature in that non-native speakers tend to underuse collocations with intensifying adverbs (Granger, 1998). As Xu (2015) maintained, English learners tend to replace such intensifying collocations (such as adverb + adjective) with single words. Collocations with phrasal verbs were also proved to be less utilized in all the three band scores. This is because phrasal verb collocations are more difficult to learn than collocations with single verbs due to their idiomatic nature and syntactic patterns (Xu, 2015; Maeen and Chilukuri, 2019). That is why, in lieu of such advanced syntactic patterns, English learners tend to cling on, as Granger (1998) put it, their "safe bets" or certain types of fixed expressions about which they are more confident (p. 147).

Of the four types of 1,074 grammatical collocations, the most frequent pattern was G8 (verb + preposition), followed by G4 (preposition + noun), G5 (adjective + preposition), and G1 (noun + preposition). This could mean that G8 (verb + preposition) was the easiest for the learners to learn, and they probably had greater exposure to verb + preposition combinations since verb patterns are quite common in English (Benson et al., 2010). On the other hand, the lack of knowledge in the use of G1 (noun + preposition) was pronounced among test-takers, especially among band score 6 test-takers. Lack of exposure to certain types of collocations could result in the test-takers' difficulty of using such structures (Bortfeld and Brennan, 1997; Hsu and Hsu, 2007). Another problem for using

adjective/noun+preposition patterns would be due to the effect of negative transfer as in Persian, the most frequent preposition is "az" (from) (Maeen and Chilukuri, 2019) which makes hindrance for Persian L2 learners to use a wide range of prepositions in using grammatical collocations in English. To illustrate, in English, the adjective "surprised" is followed by "at", the adjective "afraid" by "of", and the adjective "bored" by "with", whereas in Persian all of them are followed by "az" (from). This difficulty for the use of G1 (noun + preposition) was also confirmed by Hatami (2015), who found that this type of collocation was more difficult than the production of G4 (preposition + noun).

The second question of the study examined whether or not the collocational patterns differentiated among IELTS test-takers across various band scores. To answer this question, the study made use of chi-square tests on each type of collocations. It was found that five lexical collocations, i.e., L1 (adjective + verb), L5 (noun of noun), L8 (phrasal verb and adverb), L9 (noun and phrasal verb), and L10 (phrasal verb and noun), were able to differentiate test-takers based on their band scores. The existence of certain types of collocations, which could distinguish higher-level L2 learners from lower-level ones has been supported in literature (e.g., Durrant and Schmitt, 2009; Granger and Bestgen, 2014). In their study, Granger and Bestgen (2014) found that intermediate L2 learners tended to underuse lower-frequency collocations, using a large proportion of high-frequency combinations. They also noticed that adjective+noun types had the differentiating power between intermediate and advanced learners.

The results of the current study were also quite expected since lower-level test-takers tended to avoid complex types of collocations, such as those with phrasal verbs (i.e., L8, L9, and L10), due to their lack of knowledge of these difficult syntactic patterns (Xu, 2015; Maeen and Chilukuri, 2019). This conservative strategy, as is seen in the literature (e.g., Xu and Xi, 2010; Xu, 2015), can eventuate in lower-level test-takers' dearth of utilizing such syntactic patterns and their considerable reliance on easier patterns. Therefore, the apropos use of such complex collocations was proved to be of help to ameliorate the test-takers' speaking band-scores. This differentiating power can help to determine the specific group that the test-takers may belong to as to their use of these collocations. To illustrate, band-score 6 and 7 test-takers did not make use of L8 (phrasal verb and adverb), highlighting the fact that not only is learning phrasal verbs difficult but also more difficulty is created when phrasal verbs collocate with adverbs. This is because collocations with adverbs are not frequent among non-native speakers (Granger, 1998; Xu, 2015).

The two types of grammatical collocations with differentiation power were G4 (preposition + noun) and G5 (adjective + preposition). This finding indicates that learning the right prepositions, which collocate with nouns and adjectives (i.e., before a noun and after an adjective), could considerably change the test-takers' level of proficiency; if applied appropriately. This finding highlights the significance of prepositions in the English language as learning prepositions has proved to be the most challenging part of English learning (Takahaski, 1969). More importantly, the English language contains more diverse prepositions than the Persian language (Ghasemi et al., 2014), meaning that learning them could be an issue particularly for Iranian L2 learners due to the negative transfer of L1. Many adjectives, for instance, are followed by prepositions which cannot be produced by the literal translation of them, such as the most famous preposition in Persian which is "az" (Maeen and Chilukuri, 2019) as used in "عصبانی از" or "کسل از", which are translated into "angry at" and "bored with". In another study, Hatami (2015) found that learning G4 (preposition + noun) was easier for Iranian learners than G1 (noun + preposition). This is in contrast with the current study's results in which G4 (preposition + noun) was proved to have the differentiation power across the test-takers. Therefore, it revealed that with the improvement of their language proficiency and band score, test-takers made more effective use of G4 (preposition + noun) and G5 (adjective + preposition) in their speaking tests.

CONCLUSION AND IMPLICATIONS

In this study, the frequency of collocational use was examined in the IELTS speaking test. It was also an attempt to analyze the differentiating power of collocations across three band scores of 6, 7, and 8. Findings of the current study suggested that collocation use was a quite common phenomenon in the three corpora (i.e., band score 6,7, 8 speaking samples). In particular, the results of the study evinced that both the lexical and grammatical collocations were quite frequent across the three groups. It was also revealed that with the increase in the band scores, the number of collocations uses increased as well. That is, while 570 collocations were used in band score 6 sub-corpus, 1,087 and 1,595 were found in band score 7 and 8 sub-corpora, respectively.

According to the results, L1 (adjective + noun) and L7 (verb + noun) were the most frequently used types of lexical collocations. After these types, L4 (noun+ noun), L2 (adverb + adjective), L3 (adverb + verb), L5 (noun of noun), and L10 (phrasal verb+ noun) were detected to be the most common types. However, L8 (phrasal verb + adverb), L9 (noun + phrasal verb), and L6 (noun + verb) had the lowest frequency of uses across the band scores. In addition, G8 (verb + preposition) was the most frequently used grammatical collocation, followed by the pattern G4 (preposition + noun) and G5 (adjective + preposition), while the least used syntactic pattern was G1 (noun + preposition).

The study also sought to determine the collocational patterns, which differentiated among IELTS test-takers across band scores 6, 7, and 8. Having made use of chi-square tests, the study found that among the lexical collocation types, L1 (adjective + verb), L5 (noun of noun), L8 (phrasal verb and adverb), L9 (noun and phrasal verb), and L10 (phrasal verb and noun) had the most differentiating power. Moreover, G4 (preposition + noun) and G5 (adjective + preposition) possessed the most differentiating power among the grammatical collocation types.

According to the results of the current study, it can be concluded that the more proficient the test-takers became, the more collocations they produced in the speaking test. This means that the apt use of collocations would result in more advancement in language proficiency and communicative

competence (Crossley et al., 2014; Xu, 2015). Moreover, as the differentiating power of the collocations indicated, the more advanced test-takers tended to use more complex syntactic patterns both in terms of lexical and grammatical collocations. As an illustration, more advanced test-takers tended to use more complex patterns of collocational phrasal verbs.

The findings of this study have significant implications for IELTS trainers, IELTS materials developers, and syllabus designers. IELTS trainers should have an encyclopedic knowledge of collocations and be keenly aware of the importance of formulaicity. They, thus, can highlight the merits of formulaic language so that the test-takers would pay heed to them to ameliorate their collocational competence. Moreover, materials developers are required to pay special attention to embedding useful collocations that can contribute to the improvement of test-takers' oral proficiency. They can also use the findings of the study as to the differentiating power of collocations across the three band scores. Stated in another way, they can raise the learners' awareness of the more challenging types of collocations, which would result in the improvement of their band score.

Using lexis-based syllabus as a supplementary syllabus to the core syllabus, syllabus designers can highlight the importance of collocations in IELTS courses. They can also explicitly draw the students' attention to the differences of collocations in L1 and the English language to pre-empt the potential negative-transfer-related collocational errors. Therefore, the importance of formulaic language should be underlined in designing a course, particularly those courses preparing the students for high-stakes tests.

One of the limitations of the study included not having access to a sufficient number of speaking band score 9 since few test-takers may achieve the highest IELTS band score. Thus, the collocations were collected from band scores 6, 7, and 8. In addition, the study's delimitation relates to the choice of those grammatical types of collocations with prepositions for further analysis. According to Takahaski (1969), the most serious problem an L2 learner encounters while learning English is certainly the correct use of specific prepositions. Thus, these specific types would be the most representative of the grammatical competency of an L2 English learner.

As this study did not enjoy an enormous wealth of band scores 9, future researchers can entertain the idea of analyzing the collocations produced by this group of test-takers. Moreover, when it comes to the analysis of different types of collocations, future researchers may wish to merely work on either the lexical or grammatical ones. They can also analyze the other types of grammatical collocations, which were not considered in the present study. IELTS researchers are also recommended to conduct qualitative studies to descry the reasons behind producing miscollocations such as apprehension or anxiety-related problems. Finally, future researchers can carry out greater numbers of corpus-based studies with larger samples to throw more light on the significance of collocations in the speaking parts of the high-stakes tests.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Both authors have materially participated in the research and manuscript preparation and approved the final manuscript.

REFERENCES

Alsulayyi, M. N. (2015). The use of grammatical collocations by advanced Saudi EFL learners in the UK and KSA. *Int. J. Engl. Linguist.* 5, 32–43. doi: 10.5539/ijel.v5n1p32

Attar, E. M., and Allami, H. (2013). The effects of teaching lexical collocations on speaking ability of Iranian EFL learners. *Theory Pract. Lang. Stud.* 3, 1070–1079. doi: 10.4304/tpls.3.6.1070-1079

Bachman, L. F., and Palmer, A. S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

Bahns, J., and Eldaw, M. (1993). Should we teach EFL students collocations? *System* 2, 101–114. doi: 10.1016/0346-251X(93)90010-E

Benson, M., Benson, E., and Ilson, R. (2010). *The BBI Combinatory Dictionary of English*, 3rd Edn. Amsterdam: John Benjamins. doi: 10.1075/z.bbi

Biber, D., Johansson, J., Leech, G., Conrad, S., and Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Ltd.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., and Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: putting a lexical approach to the test. *Lang. Teach. Res.* 10, 245–261. doi: 10.1191/1362168806lr195oa

Bonk, W. J. (2000). *Testing ESL Learners' Knowledge of Collocations (ERIC Document Reproduction Service No. ED 442 309)*. Available online at: http://files.eric.ed.gov/fulltext/ED442309.pdf (accessed April 16, 2020).

Bortfeld, H., and Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Process.* 23, 119–148. doi: 10.1080/01638537709544986

Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.

Caro, K., and Mendinueta, N. R. (2017). Lexis, lexical competence and lexical knowledge: a review. *J. Lang. Teach. Res.* 8, 205–213. doi: 10.17507/jltr.0802.01

Cowie, A. P. (1994). "Phraseology," in *The Encyclopedia of Language and Linguistics*, ed. R. E. Asher (Oxford: Pergamon), 3168–3171.

Crossley, S. A., Salsbury, T., and McNamara, D. S. (2014). Assessing lexical proficiency using analytic ratings: a case for collocation accuracy. *Appl. Linguist.* 36, 1–22. doi: 10.1093/applin/amt056

Durrant, P., and Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL Int. Rev. Appl. Linguist. Lang. Teach.* 47, 157–177. doi: 10.1515/iral.2009.007

El-Dakhs, D. (2015). Collocational competence in English language teaching: an overview. *Arab World Engl. J.* 6, 68–82. doi: 10.24093/awej/vol6no1.5

Futagi, Y., Deane, P., Chodorow, M., and Tetreault, J. (2008). A computational approach to detecting collocation errors in the writing of non-native speakers of English. *Comput. Assist. Lang. Learn.* 21, 353–367. doi: 10.1080/09588220802343561

Ghasemi, F., Janfaza, A., and Soori, A. (2014). A contrastive analysis of the prepositions "of" and "from". *Int. J. Educ. Lit. Stud.* 2, 17–21. doi: 10.7575/aiac.ijels.v.2n.3p.17

Gitsaki, C. (1999). *Second Language Lexical Acquisition: A Study of the Development of Collocational Knowledge*. San Francisco, CA: International Scholars Publications.

Gläser, R. (1986). *Phraseologie der Englischen Sparche.* [Phraseology of English]. Leipzig: VEB Verlag Enzyklopädie. doi: 10.1515/9783111562827

Granger, S. (1998). "Prefabricated patterns in advanced EFL writing: collocations and formulae," in *Phraseology: Theory, Analysis, and Applications*, ed. A. P. Cowie (Oxford: Oxford University Press), 145–160.

Granger, S., and Bestgen, Y. (2014). The use of collocations by intermediate vs. advanced non-native writers: a bigram-based study. *Int. Rev. Appl. Linguist. Lang. Teach.* 52, 229–252. doi: 10.1515/iral-2014-0011

Gui, S. C., and Yang, Z. H. (2003). *Chinese Learner English Corpus*. Shanghai: Shanghai Foreign Language Education Press.

Halliday, M. A. K., and Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hatami, S. (2015). *Collocations in Farsi L2 Learners of English. The Role of Proficiency and L1 Language Transfer* (Unpublished Master's thesis). Tromsø: The Arctic University of Norway.

Howarth, P. A. (1998). Phraseology and second language proficiency. *Appl. Linguist.* 19, 24–44. doi: 10.1093/applin/19.1.24

Hsu, J. Y., and Hsu, L. C. (2007). Teaching lexical collocations to enhance listening comprehension of English majors in a technological university of Taiwan. *Soochow J. Foreign Lang. Cult.* 24, 1–33.

Hsu, J.-y. (2007). Lexical collocations and their relation to the online writing of Taiwanese college English majors and non-English majors. *Electron. J. Foreign Lang. Teach.* 4, 192–209.

Hsu, J.-y., and Chiu, C. y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *Asian EFL J.* 10, 181–204.

Hyland, K. (2006). *English for Academic Purposes an Advanced Resource Book*. London: Routledge. doi: 10.4324/9780203006603

Hyland, K. (2008). As can be seen: lexical bundles and disciplinary variation. *Engl. Specif. Purp.* 27, 4–21. doi: 10.1016/j.esp.2007.06.001

Johansson, S., and Hofland, K. (1989). *Frequency Analysis of English Vocabulary and Grammar*. Oxford: Oxford University Press.

Keshavarz, M. H., and Salimi, H. (2007). Collocational competence and cloze test performance: a study of Iranian EFL learners. *Int. J. Appl. Linguist.* 17, 81–92. doi: 10.1111/j.1473-4192.2007.00134.x

Keshavarz, M. H., and Taherian, P. (2018). The Effect of explicit instruction of collocations on EFL learners' language proficiency. *Hacettepe Univ. J. Educ.* 33, 987–1001. doi: 10.16986/HUJE.2018038632

Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Laufer, B., and Waldman, T. (2011). Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Lang. Learn.* 61, 647–672. doi: 10.1111/j.1467-9922.2010.00621.x

Lazaraton, A. (2014). "Second language speaking," in *Teaching English as a Second or Foreign Language*, eds M. Celce-Murcia, D. M. Brinton, and M. A. Snow (Boston, MA: National Geographic Learning/Heinle Cengage Learning), 106–120.

Levelt, W. J. M. (1999). "Producing spoken language: a blueprint of the speaker," in *The Neurocognition of Language*, eds C. M. Brown and P. Hagoort (Oxford: Oxford University Press), 83–122. doi: 10.1093/acprof:oso/9780198507932.003.0004

Macis, M., and Schmitt, N. (2016). The figurative and polysemous nature of collocations and their place in ELT. *ELT J.* 71, 50–59. doi: 10.1093/elt/ccw044

Maeen, N., and Chilukuri, B. A. (2019). Comparative study of phrasal verbs in English and Persian. *Spec. J. Lang. Stud. Lit.* 3, 16–31.

Mei, J. J. (1999). *Dictionary of Modern Chinese Collocations.* Shanghai: Hanyu Dictionary Press.

Mel'cuk, I. (1998). "Collocations and lexical functions," in *Phraseology: Theory, Analysis, and Application*, ed. A. P. Cowie (New York, NY: Oxford University Press), 23–53.

Men, H. (2018). *Vocabulary Increase and Collocation Learning.* Singapore: Springer. doi: 10.1007/978-981-10-5822-6

Namvar, F. (2012). The relationship between language proficiency and use of collocation by Iranian EFL students. *3L Southeast Asian J. Engl. Lang. Stud.* 18, 41–52.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Appl. Linguist.* 24, 223–242. doi: 10.1093/applin/24.2.223

Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam: John Benjamins. doi: 10.1075/scl.14

Nizonkiza, D. (2011). The relationship between lexical competence, collocational competence, and second language proficiency. *Engl. Text Constr.* 4, 113–145. doi: 10.1075/etc.4.1.06niz

Obukadeta, P. (2014). "L2 collocations: a problematic linguistic phenomenon?," in *Proceedings of the Talk Given at the Birmingham English Language Postgraduate Conference*, (Birmingham: University of Birmingham).

Palmer, H. E. (1933). *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.

Paquot, M. (2018). Phraseological competence: a missing component in university entrance language tests? Insights from a study of EFL learners' use of statistical collocations. *Lang. Assess. Q.* 15, 29–43. doi: 10.1080/15434303.2017.1405421

Paquot, M., and Granger, S. (2012). Formulaic language in learner corpora. *Annu. Rev. Appl. Linguist.* 32, 130–149. doi: 10.1017/S0267190512000098

Pawley, A., and Syder, F. H. (1983). "Two puzzles for linguistic theory: nativelike selection and nativelike fluency," in *Language and Communication*, eds J. C. Richards and R. W. Schmidt (London: Longman), 191–226.

Sarvari, S., Gukani, A. H., and Khomami, H. Y. (2016). Teaching collocations: further developments in L2 speaking fluency. *J. Appl. Linguist. Lang. Res.* 3, 278–289.

Shih, R. H.-H. (2000). "Collocation deficiency in a learner corpus of English: from an overuse perspective," in *Paper Presented at the 38th Annual Meeting of the Association for Computational Linguistics*, Hong Kong.

Shumin, K. (2002). "Factors to consider: developing adult EFL student speaking abilities," in *Methodology in Language Teaching: An Anthology of Current Practice*, eds J. C. Richards and W. A. Renandya (Cambridge: Cambridge University Press), 204–211. doi: 10.1017/CBO9780511667190.028

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Takahaski, G. (1969). Perception of space and the function of certain English prepositions. *Lang. Learn.* 19, 217–234. doi: 10.1111/j.1467-1770.1969.tb00464.x

Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Appl. Linguist.* 21, 463–489. doi: 10.1093/applin/21.4.463

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: University Press. doi: 10.1017/CBO9780511519772

Xu, J. (2015). *Predicting ESL Learners' Oral Proficiency by Measuring the Collocations in Their Spontaneous Speech* (Unpublished doctoral dissertation). Ames, IA: Iowa State University.

Xu, J. (2018). "Spoken collocational competence" in communicative speaking assessment. *Lang. Assess. Q.* 15, 255–272. doi: 10.1080/15434303.2018.1482900

Xu, J., and Xi, X. (2010). "Comparing human and machine judgments of collocations and relating them to speaking proficiency," in *Paper Presented at the 32nd Language Testing* Research Colloquium (Cambridge: University of Cambridge).

Zughoul, M. R., and Abdul-Fattah, H. (2003). Translational collocational strategies of Arab learners of English: a study in lexical semantics. *Babel* 49, 59–81. doi: 10.1075/babel.49.1.05zug

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

Appendix A: A Sample Transcription of the IELTS Speaking Test Interview (With a Band-Score 8 Test-Taker):

Examiner: This is The X mock speaking test for the International English Language Testing System conducted on February 6'h, 2020 at the X. The candidate number is X. Good afternoon. Can I have your full name please?
Candidate: My name is XX.

E. In first part, I'm going to ask you some questions about yourself. Let's start by talking about what you do. Do you work or are you a student?
C: Both I should say. Actually, I'm a teacher and I'm about to finish my PhD program.
E. Ohoom. What work do you do exactly?
C: I am an English instructor. This is my main job. I also do some editing and some research if you call it a work.
E: Ok.
C: Yeah.
E. How popular is this job in your country?
C: Well I should say is quite popular eh especially in terms of prestige I think English teachers have a high prestige although the salaries are often very low and most people just look- uh look up to English teachers somehow. Uh well I think many people think of- it consider it as a kind of uh prosperous of I mean job.
E. Ok. Thank you.
C: You're welcome.
E. Let's talk a little bit about television. How much TV do you usually watch?
C: Well uh actually I watched TV every night but actually I don't plan to watch TV it's just on and I just want in front of it while doing my other stuff I just spend time watching those soap operas, too.
E. Do you have a favorite TV program?
C: Not actually I don't watch TV. I mean it's not my favorite hobby to watch TV. I just do it as a kind of every day activity.
E. How much did you watch TV when you were a child?
C: Actually, I watched TV a lot. I used to watch a lot of uh cartoons and animations actually I didn't miss anything on TV and uh well I stayed up late at night watching those soap operas with my family too. And uh I used to watch sports uh like different kinds of matches. It was somehow very interesting.
E. Thank you. Do you think television has changed in the past few decades?
C: Television has changed in the past! Uh I guess yes. The quality of the programs of course the-the quality of everything in-in the society has changed. I should say the quality has raised a little bit. Of course, we are not comparable to other countries but as far as we are talking about Iranian programs compared to the past, I should say yeah, the- the quality has changed, and the types of the programs have also changed because I remember when I was watching TV as a child, uh most of the cartoons we used to watch were just so sad and very
E: Ok.
C: Yeah.
E. Has television changed your life in any way?
C: My life! Uh I don't think so. No. I haven't been influenced by TV.
E. Thank you. Let's talk a little bit about the countryside. Uh would you like to live in the countryside in the future?
C: Uh you mean living permanently? Uh It's an option for me maybe after I get rid of course I should decide later but yeah, I just uh don't uh reject it was an al-alternative because I think we need- we all need all kinds of uh peace in our later lives so yeah maybe.
E. Ok. What are the benefits of living in rural areas?
The benefits in first part- the first one is peace and quiet. The second one is like the fresh air eh you just have to every day I mean it's kind of eh eh it's a very big oppor-opportunity for the people who are living in uh metropolitan cities like Tehran So I guess these two are enough to just convince a person to move to the countryside. The other stuff like are I think privacy If you just have good- good place for yourself
E. Thank you very much.
C: You're welcome.
E. Now I'm going to give you a topic and I'd like you to talk about 1 to 2 minutes. Before you talk you've got one minute to think. You can make some notes if you wish. Do you understand?
C: Sure.
E. So, here's some paper and sometimes
C: Thank you.
E. And here's your topic. I'd like you to describe a small business that you would like to own.
C: The small business! Ok.

[After two minutes]

C: Here you are.

E. You can have it to two minutes.

C: I can have it?

E. Yeah. So, remember you have one to two minutes to talk about this topic so don't worry if I stop you. I'll tell you when the time's up. Can you start please?

C: Sure. Well uh the best small business I would like to uh start uh actually to own by myself is a café. I-I always thought about having a café uh in just uh decorating it the way I wish and running it the way that I would like. And uh before I start this business I think I have to educate myself in uh some related area such as how to run a café first of all and how to manage just uh these kinda business then I have to know about the different kinds of desserts and the food. Uh even I should know about like uh how to deal with the customers I think and I would uh run this business of course run this cafe using some uh qualified maybe eh eh baristas and I don't know waiters and I have to just uh find intricate rules some of those eh best ones may be those who are just uh made for this job and experts so that I can run the away I wish so uh this is my first choice because I've al-I've always enjoyed being in cafés. I-I also uh have good memories of our friends X spending time and I think most-most of the people who just go to café our younger ones and I also like spending time with young generation so I think the energy in a café < > properly somehow very fresh and into me so that's why I just choose this one after that I would just go for other businesses. my own my- husband is just into it so that's another reason I would just uh do it in the future. Yeah

E. Have you ever own your own business?

C: No, I haven't.

E. Would you like to have a family business?

C: Uh if-if my partner is uh considered yes but the other family members I wouldn't say yes because we have different ideas but-but my husband is actually on my side so I think we're just on the same track if we just wanna run a business together so I think he's a good option to be a partner.

E. Ok. thank you very much can I have the card and the notes please? C: Here you go.

E: Thank you very much. We've been talking about a small business you would like to on now I'm going to ask you some more general questions relate to this. Let's consider first of all small businesses. Uh what types of small businesses are most popular in your country?

C: These days I see a lot of start-ups being- working like eh eh being run- run by the young-young people, eh university students and I think mostly in IT and to l-like to start applications to-to pro-program- app- computer programs to design applications I think now this is very popular Iran.

E: Why? Why do you think these businesses are popular?

C: Well it's quite clear you know because internet is just the main uh uh- the main forum- or the let's say the ma- the mainstream uh medium for communication for running different businesses so many people are just- tend to- they just tend to uh get f-familiar with eh how to work with Internet so I think it's the first reason the most important reason is actually the eh of the commonality of Internet-based businesses.

E: Ohoom. What challenges and difficulties do people face when they try to start a small business?

C: A small business in general, you mean?

E: Yeah.

C: Uh first of all you- I mean a person who just starts a business has to have some qualities I mean there are some features that a person has to possess like self-discipline and let's say uh leadership or like the ability to work on his or hor- her own. So, the first thing eh comes to my mind regarding the question uh you just asked is eh having these personal qualities. This-this is a challenge for many people because sometimes they're not used to that- that kind of work and the other stuff is like the support from the government and from the society, so they need just you uh uh work very hard. They have to just uh eh spend a lot of time and energy and sometimes money to attract uh customers and also to attract support in part of their work.

E: Ok thank you. How can small businesses benefit local people?

C: Uh small businesses! uh I think eh there is something eh which has to be considered when we are talking about marketing and stuff that- that is trust. So, for the local people I think it's much easier to trust eh eh to trust eh one's eh whom X uh. Eh I myself I would just trust the person whom I-I just know for a long time. I think is the benefit for the local people around that eh small business. For bigger companies, usually we don't have this kind of uh knowledge and familiarity about them, so it's really hard to just f- trust them.

E: Ok. Let's talk a bit about business owners right now. Why do some people start their own businesses?

C: First of all, it's because of the many of the problems that exists. When you are- when you work for others- when you're employed by someone such- such as lack of freedom or uh sometimes the uh hostile atmosphere some people experience in eh their workplaces. So, they just go for running their own business I mean standing on their own feet somehow. And the other part I guess s- running a business by their own is actually more beneficial financially so people just uh hope to earn more if they run their own businesses, so I guess these are enough reasons to just to uh business of course if they- when this is kind of eh risk-taking job, so they take a risk to uh with the hope of earning some benefit in the future.

E: Ok. Thank you very much. This is the end of the speaking.

Appendix B: Syntactic Patterns of Lexical Collocations (Adapted From Xu, 2015, p. 80)

Type	Example from the corpus
1. Adjective and noun	pensive mood
2. Adverb and adjective	undeniably significant
3. Adverb and verb	sincerely thank
4. Noun and noun	job opportunities
5. Noun of noun	acts of terrorism
6. Noun and verb	the business flourishes
7. Verb and noun	make a decision
8. Phrasal verb and adverb	stay up late
9. Noun and phrasal verb	time goes by
10. Phrasal verb and noun	come across obstacles

Appendix C: Syntactic Patterns of Grammatical Collocations (Adapted From Benson et al., 2010, pp. 19–30)

Type	Example from the corpus
1. Noun and preposition	skill in
4. Preposition and noun	on TV
5. Adjective and preposition	aware of
8. Verb and preposition	depend on