

## Research Article

# Mapping Deviation from the Norm: A Corpus-Driven Analysis of Phrasal Verb Creativity in Iranian EFL Writing

Abdolvahed Zarifi <sup>1\*</sup>, Najrjes Shayeste <sup>2</sup>

1. Associate Professor, English Language Dept., Faculty of Humanities, Yasouj University, Yasouj, Iran, (Corresponding Author: [vzarifi@yu.ac.ir](mailto:vzarifi@yu.ac.ir))
2. M.A., English Language Dept., Faculty of Humanities, Yasouj University, Yasouj, Iran, [shaistenarjes@gmail.com](mailto:shaistenarjes@gmail.com)

### ARTICLE INFO

#### Submission History

Received: 2025-09-02  
Accepted: 2025-12-22

#### Keywords

Corpus Linguistics  
Creativity  
EFL Learners  
IELTS Test  
Phrasal Verbs  
Unnaturalness

### ABSTRACT

Recent developments in corpus analysis have inspired studies that examined both how often linguistic structures occur and their unique, learner-specific qualities in second language acquisition. This corpus-based study investigated creativity and (un)naturalness in phrasal verb usage in writings by Iranian EFL learners at IELTS and pre-IELTS levels. The corpus comprised 500 writing samples collected from candidates enrolled in IELTS preparation courses at two institutions in Shiraz. Wordsmith software (v.4.0) was used to locate all particle elements. Each instance preceded by a lexical verb was identified and, employing the Zar Test of Initialization, phrasal verbs were classified and quantified. Their combinations, particles, and contexts were then analyzed using the Oxford Dictionary of Phrasal Verbs as a reference. Using the British National Corpus (BNC) as a benchmark for natural usage, learner deviations were identified. Results revealed a significant under-use of phrasal verbs alongside prevalent grammatical deviations in their application.

## Introduction

The advent of learner corpora in the 1990s marked a new phase in the study of learner English. In addition, subsequent developments in corpus-linguistic methodology and software have enabled deeper investigations of learner language. Researchers can now move beyond numerical features like frequency counts to examine idiosyncratic, learner-specific patterns. Furthermore, the parallel availability of first-language and second-language corpus data allows applied linguists to investigate phenomena such as avoidance, which were seldom addressed in traditional error analysis.

Within the field of second language learning research, corpus-based investigations into learner language have consistently highlighted the critical importance of vocabulary. Lexis serves as the essential foundation for achieving accurate, fluent, and impactful communication (Rudzka-Ostyn, 2008). The selection of words fundamentally shapes the meaning and communicative intent of any message. Consequently, as Waibel (2007) notes, the proficient use of vocabulary is key to achieving more natural-sounding learner language—an objective for many advanced students. Furthermore, vocabulary represents a uniquely complex component of linguistic knowledge. It is an open and dynamic system, continually incorporating new elements, which makes it resistant to simple, rule-based organization (Waibel, 2007).

A significant portion of natural language is also composed of multi-word units rather than isolated words. Moon (1998) emphasizes that language use is often idiomatic and formulaic, and multi-word units often convey meanings that are irreducible to their constituent parts. As such, research into formulaic language

forms a cornerstone of corpus linguistics and, more specifically, learner corpus research (Paquot & Granger, 2012). Given that items such as collocations, phrasal verbs, and fixed expressions are particularly challenging for learners to master (Garbatovič & Grigaliūnienė, 2020), they have constituted a major focus of investigation since the early stages of the discipline.

Phrasal verbs, a key type of multi-word units and a notoriously difficult part of the lexicon, present a well-documented challenge in foreign language teaching and learning. As Rudzka-Ostyn (2008) argues, this difficulty stems from their semantic path from concrete to abstract, which is often opaque to learners. However, with knowledge of the verb and particle, the idiomatic meaning can remain elusive, leading even advanced learners to struggle with comprehension and use.

EFL learners face considerable challenges in acquiring and using phrasal verbs, a difficulty attributable to several interconnected reasons. Firstly, phrasal verbs represent a defining characteristic of Germanic languages (Celce-Murcia et al., 1999). Consequently, they appear particularly foreign and unintuitive to speakers whose first language, such as Farsi, belongs to a different linguistic family (Siyanova & Schmitt, 2007). This inherent unfamiliarity establishes an initial barrier to comprehension. Secondly, these constructions are notoriously difficult due to their intrinsic semantic complexity, which is compounded by significant syntactic diversity. A single phrasal verb can operate under different grammatical rules, adding a layer of structural confusion to the puzzle of meaning.

A third major factor is the contextual dependence and polysemy of these multi-word units. Their meaning can shift dramatically based on the surrounding text,

making them ubiquitous yet highly ambiguous. This semantic slipperiness requires learners to grasp not just a definition, but a range of potential senses. Fourthly, the system of phrasal verbs is remarkably productive. Particles can combine freely with verbs from shared semantic domains, creating a dynamic and creative area of the lexicon (Armstrong, 2004; McArthur, 1989). This very productivity, while a testament to the linguistic economy of English, means the inventory of possible phrasal verbs is vast and ever-expanding, presenting a moving target for learners. As a result of this creative potential, phrasal verbs are considered a fertile ground for lexical innovation and are thus a rich area for linguistic investigation in EFL contexts. Scholars note that they form an essential and economical subsystem within English. The capacity for almost any lexical verb to merge with various particles underscores their generative nature (Alangari et al., 2020). However, this productivity intensifies the problems learners encounter. The challenges—whether syntactic, semantic, or both—are magnified by the endless possibilities for combination.

Paradoxically, the creative power of phrasal verbs, when coupled with their frequent idiomaticity, non-compositionality, and multiplicity of meanings, causes substantial learning hurdles (Armstrong, 2004). This combination makes it demanding for second language learners to internalize, understand, and deploy phrasal verbs both naturally and creatively. Despite their often complicated structure and sometimes opaque meanings, mastering phrasal verbs is crucial for EFL learners. Their relevance is high because they are pervasive not only in informal, colloquial speech but are also abundantly used in formal and academic registers (Alangari et al., 2020).

This widespread application signifies that learners must acquire phrasal verbs to achieve functional proficiency and authenticity in English. On one hand, their frequent use makes them indispensable. On the other hand, their layered semantic, syntactic, and pragmatic complexities create persistent difficulties in both learning and practical application, often causing avoidance and errors among even advanced students. This tension between necessity and difficulty lies at the core of the challenge phrasal verbs pose.

In addition to empirical studies focusing on phrasal verbs within general corpora, a number of investigations have tested these combinations in learner corpora. To begin with, (Schneider, 2004) examined the use of phrasal verbs across four sub-divisions of the International Corpus of English (ICE)—Indian, Singaporean, Philippine, and East African sub-corpora—and juxtaposed these multi-word forms with their counterparts in the British English ICE corpus. He further examined their occurrence, frequency of occurrence, structural patterns, and productive deployment across the sub-corpora relative to British English ICE. The analyses revealed that Singaporean learners employed phrasal verbs with greater frequency than other variety users of English and even more than native English speakers. Additionally, the propensity of non-Singaporean learners toward employing phrasal verbs appeared to decline with increasing formality in style. A key finding was that while Singaporean ESL learners, similar to native speakers, strongly favoured phrasal verbs in oral communication, other ESL communities tended to deploy these combinations in more formal contexts.

Among the broad-scale inquiries, Waible's (2007) study stands out for examining phrasal verb use in learner data. His analysis of phrasal

verb usage in the International Corpus of Learner English (ICLE), including the German (G-ICLE) and Italian (I-ICLE) sub-corpora, revealed that German learners employed these constructions more frequently than both their Italian counterparts and native speakers in the Louvain Corpus of Native English Essays (LOCNESS). This strong preference among German learners was attributed to their tendency to employ lexical verbs of Germanic origin. Conversely, Italian learners tended to avoid phrasal verbs, opting for verb forms of Latin origin.

In a separate investigation, Chen (2013) examined how phrasal verbs function within argumentative and academic writing genres produced by Chinese university students. He compared the argumentative and academic writing portions of the Chinese learner English corpus against corresponding segments in four native corpora representing American and British English varieties. Findings indicated that American students' usage of these multi-word units differed from that of British students both in total numbers and in the types of constructions. Although American students employed a higher quantity of phrasal verbs with broader variety than their British counterparts, in argumentative writing, American and British learners both employed phrasal verbs more often than they did in academic writing. Conversely, while the overall rates of phrasal verbs in the Chinese corpus seemed similar to those in the British variety, Chinese learners acted differently from their American peers in this respect.

Despite the extensive corpus-based work on the broad aspects of phrasal verbs within both general and learner corpora, investigations into the distinctive traits of these expressions remain scarce and scattered. For example, (Zarifi & Mukundan, 2014)

investigated creativity and unnaturalness in Malaysian ESL learners' phrasal verb use. Results showed that Malaysian ESL learners exhibited a pronounced tendency toward both overuse and coinage of novel phrasal verb combinations in their language output. This evidence allowed researchers to argue that although theoretical and empirical pointers often support under-using phrasal verbs by non-native speakers, ESL learning in the Malaysian setting appears to counteract the disparities between L1 and L2. While Malaysian ESL learners were nearly error-free with non-idiomatic pairings, such as 'push out', 'walk out', etc.; their attempts to construct new idiomatic forms were error-prone, for instance, 'use out', 'voice out', etc. Finally, in a recent investigation, Garbatovič and Grigaliūnienė (2020) contrasted the use of phrasal verbs in the Lithuanian and Polish sub-sections of the International Corpus of Learner English. They observed that both Lithuanian and Polish learner groups underused phrasal verbs relative to native English speakers. They attributed this underuse to learners' restricted repertoire of these idiomatic units. They added that, beyond the intrinsic complexity of phrasal verb constructions, part of learners' difficulties with these patterns can be explained by their absence in the linguistic inventory of EFL Lithuanian and Polish learners. To conclude, while these studies confirm a general trend toward avoidance and unnatural usage, there is a need for more nuanced analysis of the character of this unnaturalness, especially concerning idiosyncratic and creatively intended forms among under-researched learner populations, such as Iranian EFL candidates.

## Research Questions

The current study aimed at investigating the use of phrasal verbs by Iranian EFL learners, with a specific focus on identifying patterns of unnatural usage of these multi-word units in learners' writings at IELTS and pre-IELTS levels. Therefore, the research questions are twofold:

1. How are phrasal verbs distributed in the written corpus of Iranian IELTS and pre-IELTS level learners?
2. How natural is the usage of phrasal verbs in Iranian IELTS and pre-IELTS level candidates' writing?

## Methodology

### Corpus of the Study

Employing a corpus-based methodology, this research was designed to investigate the use of phrasal verbs in the written production of Iranian EFL learners. The corpus of the study was collected from the written texts produced by Pre-IELTS and IELTS level language learners who attended the IELTS preparation classes in Bahar and Kish Language Institutes in Shiraz, Iran. The corpus was composed of 500 pieces of written texts from the first and second writing tasks of academic as well as general modules of IELTS test. 274 counts of texts which consisted of 73916 words belonged to the IELTS level (172 texts from general and 102 texts from academic writings), and 226 texts containing 43784 words were composed by the pre-IELTS level (81 texts from pre-general and 145 texts from pre-academic writings).

### Procedure

#### Extraction and Identification of Phrasal Verbs

The identification of phrasal verbs involves several considerations, including the number of potential particle elements that can

contribute to combinations, the length of combinations, and the syntactic role of the particle element in each combination. Given the extensive inventory of lexical verbs in the English language and their tendency to combine with one or more particle elements to make up phrasal verbs, there exist a fairly limited number of particles in the language. Consequently, in the current investigation, we utilized the 22-item list of potential particle elements that is often recognized as the particle elements in making phrasal verbs. Table 1 shows this potential list of particles.

**Table 1**

#### *Potential particle elements*

About	After	Along	Away	By	For
In	On	Over	Round	Up	
Across	Ahead	Around	Back	Down	Forward
Off	Out	Part	Through	With	

### Classification of Phrasal Verb Types

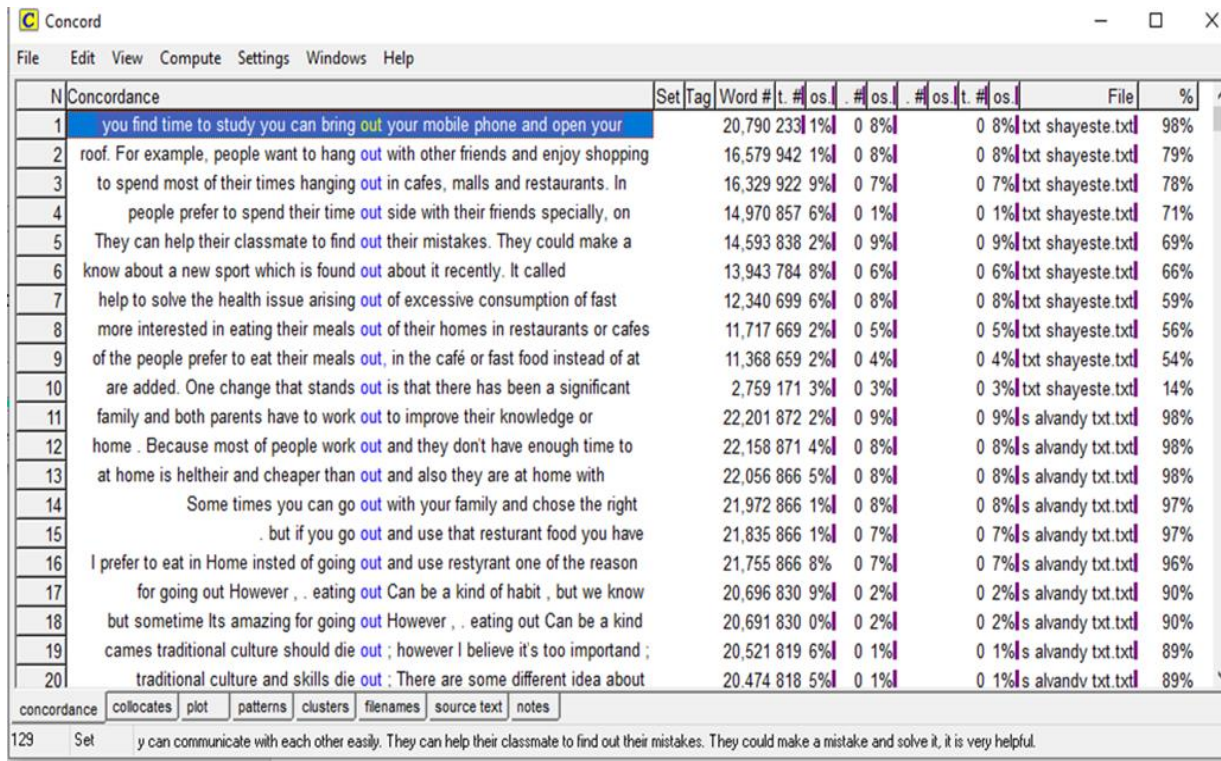
The identification of particle elements within the study corpus was conducted using the concordance function of Wordsmith Software (Version 4.0). The linear separation between the lexical verb and its particle governs the syntactic behavior of a phrasal verb combination. Consequently, for the purposes of this analysis, any co-occurrence of a lexical verb with an adjacent or nonadjacent particle was initially coded as a potential phrasal verb instance.

In the next stage, the software's concordance function was employed to isolate all particle elements appearing in the corpus. Subsequently, a second concordance analysis was run to locate every instance where a particle was preceded by a lexical verb to its left, thereby capturing both contiguous and discontinuous verb-particle constructions.

To illustrate, the sequence *bring out* (Line 1, Figure 1) was classified as a phrasal verb, consisting of the lexical verb *bring* and the

nearby adverbial particle *out*. In contrast, the expression *eat their meals out* (Line 9, Figure 1) was also recognized as a phrasal verb construction, even though its components are not adjacent. Here, the noun phrase *their*

*meals*, serving as the direct object, appears between the lexical verb *eat* and the particle *out*. A representative sample of the concordance lines generated for the particle *out* is shown in Figure 1.



N	Concordance	Set	Tag	Word #	t. #	os.	. #	os.	. #	os.	t. #	os.	File	%
1	you find time to study you can bring out your mobile phone and open your			20,790	233	1%	0	8%	0	8%	0	8%	txt shayeste.txt	98%
2	roof. For example, people want to hang out with other friends and enjoy shopping			16,579	942	1%	0	8%	0	8%	0	8%	txt shayeste.txt	79%
3	to spend most of their times hanging out in cafes, malls and restaurants. In			16,329	922	9%	0	7%	0	7%	0	7%	txt shayeste.txt	78%
4	people prefer to spend their time out side with their friends specially, on			14,970	857	6%	0	1%	0	1%	0	1%	txt shayeste.txt	71%
5	They can help their classmate to find out their mistakes. They could make a			14,593	838	2%	0	9%	0	9%	0	9%	txt shayeste.txt	69%
6	know about a new sport which is found out about it recently. It called			13,943	784	8%	0	6%	0	6%	0	6%	txt shayeste.txt	66%
7	help to solve the health issue arising out of excessive consumption of fast			12,340	699	6%	0	8%	0	8%	0	8%	txt shayeste.txt	59%
8	more interested in eating their meals out of their homes in restaurants or cafes			11,717	669	2%	0	5%	0	5%	0	5%	txt shayeste.txt	56%
9	of the people prefer to eat their meals out, in the café or fast food instead of at			11,368	659	2%	0	4%	0	4%	0	4%	txt shayeste.txt	54%
10	are added. One change that stands out is that there has been a significant			2,759	171	3%	0	3%	0	3%	0	3%	txt shayeste.txt	14%
11	family and both parents have to work out to improve their knowledge or			22,201	872	2%	0	9%	0	9%	0	9%	s alvandy txt.txt	98%
12	home . Because most of people work out and they don't have enough time to			22,158	871	4%	0	8%	0	8%	0	8%	s alvandy txt.txt	98%
13	at home is heltheir and cheaper than out and also they are at home with			22,056	866	5%	0	8%	0	8%	0	8%	s alvandy txt.txt	98%
14	Some times you can go out with your family and chose the right			21,972	866	1%	0	8%	0	8%	0	8%	s alvandy txt.txt	97%
15	. but if you go out and use that resturant food you have			21,835	866	1%	0	7%	0	7%	0	7%	s alvandy txt.txt	97%
16	I prefer to eat in Home insted of going out and use restyrant one of the reason			21,755	866	8%	0	7%	0	7%	0	7%	s alvandy txt.txt	96%
17	for going out However . . eating out Can be a kind of habit , but we know			20,696	830	9%	0	2%	0	2%	0	2%	s alvandy txt.txt	90%
18	but sometime Its amazing for going out However , . eating out Can be a kind			20,691	830	0%	0	2%	0	2%	0	2%	s alvandy txt.txt	90%
19	comes traditional culture should die out ; however I believe it's too important ;			20,521	819	6%	0	1%	0	1%	0	1%	s alvandy txt.txt	89%
20	traditional culture and skills die out ; There are some different idea about			20,474	818	5%	0	1%	0	1%	0	1%	s alvandv txt.txt	89%

Figure 1

*A snapshot of the Concordance lines of particle 'out' in corpus*

The category of phrasal verbs is traditionally delineated by three distinct syntactic configurations: verb + real particle (or “predicative particle”), verb + adverbial particle, and verb + prepositional particle. While each configuration comprises a lexical verb and a morphologically invariable particle, these structures are closely related, and a single particle may function variably as a real particle, adverb, or preposition alongside the same verb. This functional ambiguity poses a persistent challenge in reliably differentiating between phrasal verb types.

To address this analytical difficulty, several syntactic diagnostics (Biber et al., 2003; Bolinger, 1971; Darwin & Gray, 1999; Fraser, 1976) have been proposed in the literature.

However, existing tests exhibit notable limitations. First, their complexity—often involving seven to eleven distinct criteria—can lead to cumbersome application and interpretive complications. Second, they are frequently subject to exceptions and counterexamples, undermining their reliability. Furthermore, while some diagnostics can distinguish verb + real particle combinations from verb + prepositional particle ones, they often fail to adequately differentiate these from verb + adverbial particle constructions.

To overcome these shortcomings, AUTHOR (YEAR) introduced the “Zar-Test of Initialization.” This diagnostic employs a single syntactic criterion—initialization—

applied across three successive stages. Its theoretical foundation rests on the principle that a preposition forms a constituent with its subsequent noun phrase, a true particle forms a cohesive unit with its preceding verb, and an adverbial particle operates more freely, remaining syntactically independent of the lexical verb.

In the present research, the Zar-Test of Initialization was applied to identify and classify phrasal verb tokens into the three target combinations (verb + adverbial particles, verb + real particles and verb + prepositional particles). Following identification, all instances were lemmatized to aggregate different inflectional forms under a single lexical head (e.g., *take, takes, taking, took, and taken* were grouped under the lemma TAKE), ensuring that frequency counts reflected lexical types rather than surface forms.

To enhance the reliability of the classification and mitigate coder idiosyncrasy, a seasoned English Language teaching professional with two decades of experience independently analyzed a subset of concordance lines, applying the Zar-Test to classify each phrasal verb unit. Inter-rater reliability between the primary researcher and the expert coder was assessed using Cohen's Kappa. The resulting Kappa coefficient of 0.97 indicates an excellent level of coding consistency.

### **Analysis of Naturalness and Creativity**

To tackle the central question about whether IELTS writers use phrasal verbs in a natural or non-native way, we checked every verb-particle pair against the Oxford Dictionary of Phrasal Verbs to determine if an entry exists. If a combination lacked a dictionary entry, we then consulted the British National Corpus (BNC), which encompasses

a wide array of written and spoken samples of British English. Any verb-particle pair that appeared at least five times in the BNC was treated as a fresh, genuine usage even though it was not listed in the dictionary. Conversely, if the combination had no dictionary entry and showed fewer than five instances in the BNC, the learner-produced form was regarded as an unnatural construction. It should be noted that misapplication of particles, inappropriate contextual or semantic use of phrasal verbs, and either under-use or over-use of these forms in the learner corpus relative to the BNC were counted as additional indicators of unnatural phrasal verb usage by learners.

### **Scope of Analysis**

Although, in the identification and quantification of phrasal verb combinations, all the three forms of phrasal verbs in the whole corpus were considered and analyzed, the combinations of verb + prepositional particle were excluded from the later qualitative stages of analysis. We assumed this position, since verb + prepositional particle combinations are inherently of a different nature. Moreover, the verb + preposition combinations tend to be fixed units in the language; not all the lexical verbs can assume a preposition element; and the number of prepositions that might follow a verb is highly limited (mainly one preposition). For instance, lexical verbs such as 'demand', 'insist' and 'buy' can be used only with 'of', 'on' and 'for' as a preposition. In addition, the English language system does not allow for innovations in the verb-preposition combinations. A preposition always precedes the noun phrase that follows its related verb, and its position does not shift as it often does with verb plus adverbial particles. Consequently, learners' atypical employment of these forms is

interpreted as unacceptable and non-English rather than as a creative choice.

## Results

The analysis yielded a total of 193 phrasal verb combination lemmas with 494 tokens in the whole corpus. Out of these, 136 phrasal verb lemmas with 361 tokens appeared in the IELTS sub-corpus and 57 phrasal verb lemmas with 133 tokens appeared in the pre-IELTS sub-corpus. It is worth noting, however, that non-English combinations were also included for analysis. For instance, while ‘die out’ was identified as unacceptable and unnatural since it has neither an entry in the Oxford Dictionary of Phrasal Verbs nor does it have any record in the BNC, yet it was included in data analysis. Therefore, in the current study, ‘die out’ and some other non-English combinations like ‘gain up’ were counted as phrasal verbs.

Further data analysis revealed that in the IELTS level sub-corpus with a total of 73916 word-tokens, as Table 2 reveals, there occurred 73 lemmas of verb + real particle combination with 205 tokens, 13 lemmas of verb + adverbial particle with 69 tokens and 50 lemmas of verb + prepositional particle with 87 tokens.

**Table 2**

*Frequency of Verb-Particle combination in the two sub-corpora*

PV Type	Verb + Real particle token	Verb + Adverb token	Verb + Preposition token	Total
Learner Level				
IELTS	205	69	87	361
Pre-IELTS	75	34	24	133
Total	280	103	111	494

## Frequency and Underuse

Mirroring the trend in the IELTS sub-corpus, the pre-IELTS sub-corpus (43,784

words) contained fewer phrasal verbs, with 57 lemmas occurring 133 times. More particularly, in this sub-corpus, there were 32 lemmas of verb + real particle combination which appeared in 75 tokens, 10 lemmas of verb + adverbial particle combination which occurred with 34 tokens, and 15 lemmas of verb + prepositional particle combination which appeared in 24 tokens. The normalized frequency (see Table 3) was 0.73 phrasal verbs per 150 words in the IELTS sub-corpus and 0.45 in the pre-IELTS sub-corpus. This falls notably below the benchmark of approximately 1 phrasal verb per 150 words reported for native English in the BNC (Gardner & Davies, 2007, p. 347), indicating a clear pattern of underuse by both learner groups.

**Table 3**

*Phrasal verb lemmas and tokens per 150 words in IELTS and Pre-IELTS sub-corpora*

Learners' Level	Word tokens in the sub-corpus	PV tokens in the sub-corpus	Verb-particle combination per 150- word	Verb-particle lemma in the sub-corpus	Verb-particle lemma per 150- word
IELTS	73916	361	0.73	136	0.27
Pre-IELTS	43784	133	0.45	57	0.19

The most frequent phrasal verb lemma across both sub-corpora was *sum up* (a combination of a verb + real particle with 50 tokens in the IELTS and 12 tokens in the pre-IELTS sub-corpus). Other high-frequency items included *deal with*, *eat out*, *throw away*, and *get out*. Among particles, *up*, *out*, and *around* were combined with lexical verbs most frequently by learners at both levels.

It should be pointed out that a small set of verb-particle lemmas produced by the learners were not listed in the Oxford Dictionary of Phrasal Verbs. These fell into two distinct

categories, that is, Analogous Creations and Unnatural Deviations. Combinations such as *walk around* and *boost up*, while not dictionary-listed, were attested in the BNC. These represent productive, native-like extensions of phrasal verb patterns. However, combinations such as *die around*, *gain up*, and *watch after* had no entry in the reference dictionary and no attestation in the BNC. These are considered idiosyncratic, non-standard formations, likely resulting from overgeneralization or erroneous analogy.

## Discussion

This study revealed that while Iranian IELTS candidates demonstrated an awareness of both the functional importance and high-frequency occurrence of phrasal verbs across genres, their actual deployment of these multi-word units was significantly less frequent than that of native speaker benchmarks. Furthermore, their productive use was markedly constrained. Attempts to coin novel combinations were largely limited to structures employing lexical verbs with adverbial particles, yielding predominantly literal or aspectual meanings. Crucially, efforts to generate novel items with lexical verbs and true prepositional particles to convey idiomatic meaning were largely unsuccessful. This often resulted in deviant, non-conventional forms and, consequently, unnatural usage.

Research indicates that Iranian EFL learners exhibit unnatural usage of phrasal verbs, characterized primarily by under-use across proficiency levels (Gardner & Davies, 2007). This aligns with Granger's (2012) observation that learner language unnaturalness often stems more from the "over- and under-use of linguistic items or structures" than from outright errors (p. 41).

The present study found *sum up* to be the most frequent phrasal verb employed by learners. Notably, this combination does not rank among the top 100 phrasal verbs in the British National Corpus (BNC), suggesting its inflated frequency represents a case of unnatural use in learner writing. However, this prominence can be functionally explained: *sum up* serves as a conventional meta-discursive device for text structuring in academic essays, often signalling conclusion or recapitulation (Granger & Tyson, 1996). As Waibel (2007) notes, its prevalence may reflect pedagogical emphasis in IELTS preparation programs, where instructors highlight its role in marking the end of an argument and summarizing preceding points. Consequently, learners may deliberately position this construction at sentence-initial positions to enhance rhetorical clarity and guide the reader through the argumentative flow.

The particles *up*, *out*, and *around* were the most frequently employed by Iranian EFL learners at both proficiency levels in their production of phrasal verbs. This finding aligns with studies of other learner populations, where these same particles are identified as highly productive. For instance, AUTHOR and Waibel (2007) note that Malaysian, German, and Italian learners also over-rely on these three particles, often resulting in creative but unnatural phrasal verb constructions.

Focusing on the particle *around*, its semantic flexibility presents both an opportunity and a challenge for learners. Rudzka-Ostyn (2008) outlines its core spatial senses, such as movement "along a circular path," "in different directions within some bounded area," or "in no particular direction" (p. 188). Critically, as Lindstromberg (2010)

emphasizes, motion denoted by *around* is often non-teleological, implying “goalless motion” when no endpoint is specified (p. 135). Iranian learners in this study appeared to intuitively grasp this non-goal-directed concept. By analogizing with existing patterns like *run around*, they creatively and accurately produced the non-idiomatic phrasal verb *walk around* to describe aimless ambulation, as evidenced in both IELTS and pre-IELTS sub-corpora:

...and a couple are walking around. (pre-IELTS)

People often take their ability to walk around and get wherever... (IELTS)

Although *walk around* is not listed with this specific meaning in standard references like the *Oxford Dictionary of Phrasal Verbs*, its acceptability to native speakers is high. This can be attributed to the grammatical flexibility of adverbial particles, which can combine freely with a wide range of lexical verbs, and is further supported by its considerable frequency in native corpora such as the BNC. However, this very productivity can lead to error when learners overextend the pattern. An illustrative case from the pre-IELTS corpus is the unnatural construction *die around*:

... and people die around. We must...

While *around* combines productively with verbs of motion, the verb *die* does not semantically encode physical movement. The creation of *die around* is likely an interlingual transfer, potentially influenced by a religious-cultural conceptualization of death as a transition to another realm. This results in a formulation that, despite a plausible underlying logic, is pragmatically and collocationally unacceptable in standard English

The particles *up* and *out* were prominent in the corpus as loci of both creative extension and unnatural usage. Analysing these instances reveals how learners employ particle semantics, often through analogy, with varying success. The particle *up* frequently encodes completive aspect. As an adverbial element, ‘Up’ signals movement from a lower position to a higher one, while as a real particle, it evokes the sense of “completeness or thoroughness” (Lindstromberg, 2010, p. 24). Higher-proficiency (IELTS-level) learners applied this semantic constructively, generating the acceptable, though marginally conventional, combination *boost up* by analogy with verbs like *grow up* (e.g., “...can boost up our understanding”). In contrast, lower-proficiency (pre-IELTS) learners overextended this pattern, producing the unnatural form *gain up* (e.g., “...gain up a life style”), where the simplex verb *gain* alone is standard. This indicates a developmental overgeneralization of the completive *up*.

Another particle contributing to the unnatural use of phrasal verb combinations in the corpus was ‘out’. ‘Out’, generally conveys the idea of “a container and an object which moves out of the container” (Rudzka-Ostyn, 2003, p. 14). Despite the learners’ creation of few phrasal verbs by analogy with the existing patterns, it turned out that they would run into forging a large number of unacceptable forms. The pre-IELTS level learners, for instance, created the combination ‘work out’ based on some existing units like ‘eat out’ and ‘throw out’ in which the particle ‘out’ acts as the adverbial particle in its spatial meaning. This combination is an acceptable English form although it might be unlikely to be used in such a context by native speakers, and it can, therefore, be considered as a case of unnatural use of phrasal verb.

- Because most people work out and they don't have enough time to...

...it is good for them to work out, they say children can experience....

The particle *out*, prototypically denoting motion from a container (Rudzka-Ostyn, 2003, p. 14), was also subject to analogical creation. Learners correctly formed the literal phrasal verb *work out* (e.g., "most people work out"), likely modelled on patterns like *eat out*. However, its application in the given context—referring simply to working outside the home—constitutes an unnatural usage, as the intended meaning is not idiomatic for native speakers. This suggests that even structurally acceptable formations can be pragmatically infelicitous, highlighting a gap between syntactic analogy and contextual appropriateness. These patterns illustrate a key challenge for learners: moving beyond the productive generation of phrasal verb forms via semantic analogy to acquiring the lexical, collocational, and pragmatic constraints that govern their natural use.

Further evidence of non-native-like phrasal verb usage in the Iranian EFL corpus involves semantic extension through L1 transfer. For instance, one writer produced the string "... *it goes back to their government*," intending the meaning "is the responsibility of." While the phrasal verb '*go back*' is highly frequent in native corpora such as the BNC, its application here with an abstract noun phrase (*their government*) constitutes a marked collocational deviation from native norms. This usage appears to be a calque, directly translating a Persian expression into English, thereby generating an unnatural construction.

Beyond high-frequency particles, participants also produced unconventional combinations with other particles, exemplified

by the non-standard form '*watch after*' (intended meaning: "to take care of"). This innovation likely stems from analogical reasoning with the established pattern '*look after*,' indicating a strategy of overgeneralization when idiomatic knowledge is incomplete.

These deviations—both collocational (*go back*) and combinatorial (*watch after*)—primarily stem from contextual misapplication, where learners extend phrasal verbs to linguistic contexts atypical for native speakers. This phenomenon aligns with research attributing such errors to a limited repertoire of conventionalized combinations (Garbatovič & Grigaliūnienė, 2020) and a tendency toward erroneous semantic extension (Waibel, 2007).

The observed underuse and generation of unnatural phrasal verb combinations in the corpus can be partly attributed to learners' L1 backgrounds. Consistent with findings for speakers of other Romance languages (Sung, 2020), Iranian EFL learners—whose first language (Farsi) lacks an analogous phrasal verb structure—demonstrate a marked tendency to avoid these constructions in favour of one-word L1 equivalents when available. This avoidance results in less idiomatic and effective L2 written production. Conversely, learners from Germanic language backgrounds (e.g., German, Dutch, Swedish) typically exhibit more natural and frequent usage, facilitated by the presence of similar verb-particle constructions in their L1 (Waibel, 2007).

While the mother tongue is clearly not the sole factor influencing phraseological production (Garbatovič & Grigaliūnienė, 2020; Paquot & Granger, 2012), it remains the most discussed variable in learner corpus research. A growing body of empirical studies

substantiates L1 influence on patterns of over- and underuse, particularly reporting phrasal verb underuse by learners whose L1s lack this feature (e.g., Sung, 2017). This finding aligns with the wider claim that the lack of a particular element in the L1 can influence the L2 just as strongly as introduction of a different element would (Ellis, 1994).

The challenge for learners extends beyond avoidance to problematic creativity. As Prodromou (2003, p. 46) notes, while native speakers may creatively bend linguistic rules, learners' analogous attempts at phraseological innovation are often perceived as deviant. This is evident in the corpus: Iranian IELTS-level learners successfully employed some conventional phrasal verbs (e.g., '*walk around*', '*boost up*') but also produced unacceptable analogical combinations (e.g., '*gain up*', '*become up*', '*watch after*'). This suggests an awareness of the importance of phrasal verbs, coupled with an incomplete intuitive grasp of their constraints. Consequently, as proficiency increases, learners appear to adopt a more conservative approach, retreating to safer, attested combinations. These findings indicate that instruction must address not only the avoidance of phrasal verbs but also guide their productive use. Explicit instruction should clarify the boundaries of acceptable analogy to prevent the fossilization of deviant forms, thereby fostering both accuracy and idiomatic fluency.

## Conclusion

This corpus-based study investigated the naturalness of phrasal verb (PV) use by Iranian IELTS and pre-IELTS candidates. The findings present a dual challenge: a significant underuse of PVs compared to native-speaker norms, coupled

with unsuccessful attempts at creative formation that frequently result in unnatural, deviant combinations (e.g., '*gain up*', '*die around*'). While learners demonstrated a capacity for logical analogy in creating transparent, non-idiomatic forms (e.g., '*walk around*'), the generation of acceptable idiomatic PVs remained largely beyond their reach. This underscores that for these learners, unnaturalness manifests not only as avoidance but also as innovation without intuitive native-like constraints.

The study reinforces the significant role of L1 transfer in phraseological competence, highlighting how the absence of a parallel structure in Farsi predisposes learners to both avoidance and erroneous overgeneralization. Pedagogically, these findings suggest that instruction must move beyond merely presenting phrasal verbs as opaque lexical items. Textbook writers and instructors should instead organize instruction around systematic principles (Claridge, 2000). This involves teaching semantic clusters of particles (e.g., the completive *up*, the externalizing *out*), demonstrating how they predictably combine with specific verb classes (e.g., verbs of movement + *around*), and establishing clear boundaries to prevent overgeneralization by explaining, for instance, why *die around* fails while *walk around* succeeds. Such a systematic approach could mitigate the observed tendency toward avoidance by providing a logical framework for learners, while simultaneously reducing errors by clarifying the limits of productive analogy.

This study has several limitations that, in turn, point to fruitful avenues for future research. The corpus, while substantial, is drawn from a single educational context and genre (IELTS writing tasks). To build upon these findings, subsequent studies could

compare phrasal verb use across different genres, such as speaking versus writing or academic versus informal registers. Employing longitudinal designs would allow researchers to track the development of creative phrasal verb use alongside gains in general proficiency. Furthermore, experimental research could investigate the comparative effectiveness of pedagogy focused on semantic patterns versus traditional list-based methods. Finally, expanding the focus to explore phrasal verb processing and comprehension, not just production, in similar learner groups would yield a more comprehensive view of the challenges involved in the acquisition.

## References

Author.

- Alangari, M. A., Jaworska, S., & Laws, J. (2020). Who's afraid of phrasal verbs? The use of phrasal verbs in expert academic writing in the discipline of linguistics. *Journal of English for Academic Purposes*, 43, 100814.
- Armstrong, K. (2004). Sexing up the dossier: A semantic analysis of phrasal verbs for language teachers. *Language Awareness*, 13(4), 213-224.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (2003). *Longman grammar of spoken and written English*. MIT Press.
- Bolinger, D. (1971). *The phrasal verb in English*. Harvard University Press.
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1999). *The grammar book: An ESL/EFL teacher's course*. Newbury House Rowley, MA.
- Chen, M. (2013). Overuse or underuse: A corpus study of English phrasal verb use by Chinese, British and American university students. *International Journal of Corpus Linguistics*, 18(3), 418-442.
- Claridge, C. (2000). *Multi-word verbs in early Modern English: A corpus-based study* (2nd ed.). Rodopi Bv Editions.
- Darwin, C. M., & Gray, L. S. (1999). Going after the phrasal verb: An alternative approach to classification. *Tesol Quarterly*, 33(1), 65-83.
- Ellis, R. (1994). The study of second language acquisition. In. Oxford: Oxford University Press.
- Fraser, B. (1976). *The verb-particle combination in English*. Academic Press.
- Garbatovič, B., & Grigaliūnienė, J. (2020). Phrasal verbs in learner English: A corpus-based study of Lithuanian and Polish learners of English. *Baltic Journal of English Language, Literature and Culture*, 10, 36-53.
- Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *Tesol Quarterly*, 41(2), 339-359.
- Granger, S. (2012). Learner corpora. *The encyclopedia of applied linguistics*.
- Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27.
- Lindstromberg, S. (2010). *English prepositions explained*. John Benjamins Publishing Company.
- McArthur, T. (1989). The long-neglected phrasal verb. *English Today*, 5(2), 38-44.
- Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford University Press.
- Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130-149.
- Prodromou, L. (2003). Idiomaticity and the non-native speaker. *English Today*, 19(2), 42-48.
- Rudzka-Ostyn, B. (2003). *Word power: Phrasal verbs and compounds: A cognitive approach*. Walter de Gruyter.
- Rudzka-Ostyn, B. (2008). *Word Power: Phrasal Verbs and Compounds: a Cognitive Approach*. Mouton de Gruyter.
- Schneider, E. W. (2004). How to trace structural nativization: Particle verbs in world Englishes. *World Englishes*, 23(2), 227-249.
- Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs.
- Sung, M.-C. (2020). Underuse of English verb-particle constructions in an L2 learner corpus: Focus on structural patterns and one-word preference. *Corpus Linguistics and Linguistic Theory*, 16(1), 189-214.
- Waibel, B. (2007). *Phrasal verbs in learner English: A corpus-based study of German and Italian students* Freiburg (Breisgau), Univ., Diss., 2007].
- Zarifi, A., & Mukundan, J. (2014). Creativity and unnaturalness in the use of phrasal verbs in ESL learner language. *3L: Southeast Asian Journal of English Language Studies*, 20(3), 51-62.