LDK 2023

Unlocking the Complexity of English Phrasal Verbs and Polysemes: An Analysis of Semantic Relations Using A-Level Vocabulary Items

Kohei Takebayashi
Tokyo University of Foreign Studies
takebayashi.kohei.v0@tufs.ac.jp

Yukio Tono
Tokyo University of Foreign Studies
y.tono@tufs.ac.jp

Abstract

This two-part study aims to explore semantic transformations exhibited by English phrasal verbs (PVs) and polysemous verbs. Despite the prevalence of PVs in English communication, L2 learners of English have a noticeable tendency to avoid PVs in favour of their one-word equivalents. In order to overcome this avoidance, this research argues that PVs may serve as significant building blocks for developing learners' vocabulary knowledge. To this end, this study explores the possibility of utilising PVs as a bridge between semantic representations of A-level verbs and those of B/C-level verbs as defined by the CEFR. To ascertain the vocabulary levels of verbs found in the most common PVs, a corpus of PV textbooks (size = 3.5 million tokens) was compiled, and frequency data of word pairs composed of verbs and particles were extracted. Also, pairs of PVs and their single-word verb equivalents (SVs) were retrieved from a thesaurus. After producing a list of [PV – SV] pairs, the vocabulary levels of the SVs found on the list were identified in accordance with the English Vocabulary Profile in order to investigate the extent to which PVs can replace their SV counterparts. In addition, both PVs' semantic transparency and the degree of semantic transformation between PVs and their SV equivalents were examined. This study will demonstrate how PVs have the potential to serve as a bridge between A-level and B/C-level verbs, and a selected group of PVs will make a significant impact on the expansion of the range of meaning related to verb semantics. Furthermore, a similar methodology was applied to the investigation of sense relations and semantic transparency exhibited by polysemous A-level verbs in relation to their synonymous SVs. The findings of this study show that verb semantics display a cline of transparency, and learners' deconstruction effort of various senses displayed by polysemes may be facilitated by the semantic precision provided by higher-level SVs.

1 Introduction

The phrasal verb structure is a unique aspect of the Germanic languages (Dagut & Laufer, 1985; Darwin & Gray, 1999), playing an essential role in everyday English communication. However, phrasal verbs' ambiguity in semantic transparency, as well as their irregular syntactic features continue to confuse learners of English worldwide, leading them to choose single-word verbs instead in their production (Siyanova & Schmitt, 2007). Despite the challenges posed by phrasal verbs (PVs) and learners' tendency to avoid PVs in their output in favour of single-word verbs, the current study aims to validate the claim that learners' familiarity with the constituent verbs found in most PVs and their inclination for choosing single-word verbs are two attributes that, together, could offer a more effective means of developing vocabulary. Approaches taken for this study was threefold. First, a corpus of approximately 3.5 million tokens consisting of text data derived from 25 English phrasal verbs textbooks was compiled in order to identify what forms of PVs have been perceived to be most common and essential for learners by teachers and materials developers. Second, by assigning the vocabulary levels indexed in the English Vocabulary Profile presented by the English Profile Programme (Saville & Hawkey, 2010) (hereafter the EVP) to the constituent verbs of the PVs extracted from the corpus, the proportion of basic-level verbs that were present in the form of PVs was identified. And lastly, the EVP's vocabulary levels were assigned to single-word synonyms which were found to correspond to PVs

in order to quantitatively comprehend the relationship between PVs and their single-word verb equivalents in relation to vocabulary level development and its utility in formulating vocabulary learning strategies that would be of help to learners. The study aims to discern the leverage of A-level lexical verbs and the limitations thereof in an effort to obtain quantitative findings which may lend support to a realisation of more efficient or effective vocabulary building strategies for beginner-level learners, and ultimately to the encouragement of vocabulary acquisition among foreign language learners as a whole.

2 Review of Related Literature

PVs are generally recognised as informal or colloquial in tone, occurring 2,000 times per million words in fiction and conversation (Biber et al., 1999). Consequently, PVs have been deemed somewhat inappropriate in academic prose and formal registers, and thus the use of single-word verbs of Graeco-Latin origin in place of PVs has been perceived to be more acceptable and encouraged among learners of academic discipline (Coxhead & Byrd, 2007). However, the ubiquity of PVs has been recognised throughout the use of the English language in which learners are expected to see one PV construction for every 150 English words they encounter (Gardner & Davies, 2007), making a strong case that learners would benefit greatly from their familiarity with the characteristics and utility of PVs (Siyanova & Schmitt, 2007). PVs are considered difficult for learners to acquire due to their structural features generally reserved for the Germanic languages (Dagut & Laufer, 1985; Gilquin, 2015) and semantic complexity arising from idiomaticity and polysemy (Hulstijn & Marchena, 1989; Laufer & Eliasson, 1993; Liao & Fukuya, 2002) Efforts in the creation of PV wordlists have been made to enhance the accessibility of PVs for learners (Biber et al., 1999; Gardner & Davies, 2007; Liu, 2011). In a more recent attempt to reduce the total number of meanings of PVs to be introduced down to a manageable size based on frequency criteria, Garnier and Schmitt (2015) succeeded in producing the Phrasal Verb Pedagogical list, or more commonly known as the PHaVE list. The list contains 150 most essential PVs, as well as carefully selected definitions for the PVs based on the percentage of usages covered by these definitions. For example, the PV *take off* is given three definitions on the list with the meaning of *removing something* showing 41% of usage while the meanings of *leaving or departing suddenly* and *leaving the ground* showing 28.5% and 14% respectively. By giving priority to high frequency meanings and disregarding the rest, the study succeeded in lowering the number of meanings for the 150 PVs to be listed down to a manageable size of 288.

2.1 Research Questions

From the review of the literature, certain points could be made in regard to English PVs. First, PVs are an indispensable part of English communication without which fluid verbal interaction as well as adequate comprehension would be near impossible. Second, since PVs are one of the distinct features of the Germanic languages that are highly polysemous and often rather figurative, L2 users of English are likely to avoid using PVs in their production in preference to the safer alternative of single-word verbs, consequently making “nonnatives sound stilted and unnatural in speech” (Siyanova & Schmitt, 2007). Finally, even though efforts have been made for the production of pedagogical wordlists of essential PVs, no wordlists have incorporated the utility of vocabulary level classification provided by the English Vocabulary Profile to the discernment of the relationship between PVs and their single-word equivalents (hereafter SVs).

The current study explored PVs from three perspectives, touching upon the vocabulary levels of lexical verbs found in common PVs, their convertibility into SVs, and polysemy exhibited by common PVs. Special attention was paid on

vocabulary level progression that was expected to occur as the vocabulary level of the constituent verbs in PVs became more advanced. To this end the following research questions were addressed:

1. What are the CEFR vocabulary levels of constituent verbs found in common PVs?
2. To what extent can PVs be converted into single-word verbs, and what are the CEFR vocabulary levels of those single-word verbs?

3 Method

3.1 A list of common PVs

For the purpose of inquiring into the vocabulary levels of constituent verbs found in common PVs, compiling a list of word combinations potentially capable of forming PVs, i.e., Verb + Particle + Preposition (if any), was necessary. Furthermore, the list would be required to contain details about what combinations were considered most essential for learners by teachers and materials developers, as well as information about vocabulary levels of the constituent verbs in the word combinations that would be found in the compilation process. To this end, three steps were taken. First, a corpus of approximately 3.5 million tokens containing texts from phrasal verbs textbooks was compiled. Second, by using a special pattern matching query for extracting word combinations of Verb + Particle + Preposition (if any), frequency data of all possible combinations from the corpus was extracted. And third, vocabulary levels which correspond to the constituent verbs in the word combinations as defined by the CEFR were assigned to the verbs.

3.1.1 Corpus compiled for the study

In order to determine what combinations of lexical verbs and particles have been considered to be essential for learners by educators, a total of 25 textbooks which had been published specifically for the purpose of introducing and describing the utility of English phrasal verbs were assembled and converted into PDF files (see Appendix A for the list of the textbooks collected for this study). The texts contained in the PDF files were stored in the corpus manager, Sketch Engine (Kilgariff et al., 2014), and consequently a corpus of approximately 3.5 million tokens was compiled. For convenience, the corpus will be referred to as the Common English Phrasal Verbs Corpus throughout this paper (hereafter CEPVC). Since the CEPVC was designed to contain texts specifically produced for the description of the most useful and common English phrasal verbs, it was presumed that frequency data extracted from the corpus would accurately demonstrate what PVs had been judged to be most common and essential for learners by educators.

3.1.2 Data extraction using CQL

With the aim of extracting specific word combinations from the CEPVC as part of the second step of the process, a special code or query language was used. The query language applicable to Sketch Engine is termed Corpus Query Language (hereafter CQL) (Jakubíček et al., 2010), and is used to set criteria for words, part-of-speech, positions, etc. that would be necessary for accurate data extraction. Since word combinations which would form PVs were of interest in the current study, the following CQL as in (1) was applied to the data extraction process, which proceeded to look for the word combinations of [any lexical verbs except for *be verbs*] + [adverbs or particles or prepositions] + [prepositions (if any)] contained in the CEPVC. Although some textbooks were found to introduce transitive phrasal verbs with an object inserted in between the verb and the particle with abbreviations such as *sb* and *sth* for *somebody* and *something* respectively, (for example, *take sb out*), the majority of phrasal verbs were not introduced in this fashion. And therefore, the inclusion of *sb* (or *somebody/someone*) and *sth* (or *something*) into the CQL was deemed unnecessary in this investigation.

(1) CQL:
[tag="V.*"&!tag="VB.?"][tag="RB|RP|IN"][tag="IN"]?

3.1.3 Filtering out non-PVs

The result from the data extraction via the aforementioned CQL was exported to a spreadsheet in Microsoft Excel. Since the data contained numerous combinations that did not qualify as PVs, a certain screening measure against non-PV combinations was necessary. To this end, only particles were selected immediately following verbs by means of the filter function in MS Excel to ensure that verbs exclusively followed by appropriate particles would remain in the data. Furthermore, relative frequency (per million) was restricted to "5 or above" to filter out those combinations that were theoretically only present in a few textbooks. Moreover, since the data obtained after the filtering process with the particles still contained word combinations such as *know about* and *study at* which would not function as PVs, another measure of identifying questionable verbs (i.e., *know*, *study*, *learn*, etc.) was performed, and their validity as PVs was examined and rejected by comparing example phrasal verbs entries in dictionaries. Finally, careful attention was paid to the deletion of several combinations that contained the particle *to* which included such combinations as *go to*, *need to*, *want to*, etc., for they did not qualify as PVs.

3.1.4 Preparation of vocabulary levels

For the assignment of vocabulary levels to the verbs extracted from the corpus as part of the third step of the process, the CEFR level classification as defined by the English Vocabulary Profile (EVP) of the English Profile Programme (Saville, 2010), was referenced. The data pertaining to verbs in the EVP database was searched online (English Profile, n.d.), and was tabulated in a spreadsheet in Microsoft Excel. Since multiple proficiency levels are given to a verb in the EVP due to the polysemous nature of high-frequency English verbs, certain measure of removing duplicates was necessary. For example, the verb *make* is presented to cover five different levels, ranging from A1 to C1, in the EVP depending on its semantic complexity in given contexts. Since forms, rather than meanings, were of interest at this stage of the study, duplicates were removed while keeping the least difficult level assigned to each verb for further analysis. Therefore, the level A1 remained tagged to the verb *make* in this study. By applying the same logic to all verbs found in the EVP, the current study proceeded to reduce the 2,317 verbs originally catalogued in the EVP to a total of 1,324 unique verbs.

3.1.5 Assigning vocabulary levels to the verbs

As part of the third step of compiling a list of common PVs, the verbs found in the data were put through the process of vocabulary level assignment by the computer programming language R (R Core Team, 2022) and the application of the open-source package Tidyverse (Wickham et al., 2019). With the help of Tidyverse, the constituent verbs in the word combinations extracted from the CEPVC and their corresponding EVP vocabulary levels were tied together for the completion of the three-step process of compiling a list of common PVs. The relative frequency on the list was to indicate the number of times per million tokens the particular combinations of verbs and particles would appear in the introductions, definitions, and example-sentences in the textbooks. The current study assumes that the higher the frequency the more likely that educators would regard the word combinations as essential PVs.

3.1.6 Counting items

By making use of the table function in R, the occurrences of each vocabulary level across the range of A1 to C2 associated with the verbs present on the list were tallied, revealing the extent of representation held by each vocabulary level in the extracted data. This process allowed the investigation to quantitively discern the overall vocabulary levels of constituent verbs found in common PVs. The value of the minimum relative frequency (hereafter RF) was incrementally raised from 5 to 7, and eventually to 10 to determine the degree of change in representation held by each vocabulary level as a

function of RF. An increase in RF would mean that a smaller number of PVs would remain on the list, but the remaining PVs would be more common. It was expected that the degree of representation held by the occurrences of A-level verbs in the common PVs would increase as the PVs became more common.

3.2 Collecting synonyms

For the purpose of investigating PV's convertibility into SVs, synonyms from *the Oxford Thesaurus of English* (Oxford University Press, 2006) (hereafter OTE) were collected digitally by means of using all verbs catalogued in the EVP (i.e., 1324 verbs) as search words. All text data containing synonyms that corresponded to each search word in the thesaurus was saved as an individual text file. Furthermore, all text files collected in this manner were processed with the help of computational efficiency provided by the programming language Python (Van Rossum & Drake, 2009) such that multiword synonyms including PVs and single-word synonyms were separated into two different files. The file containing multiword synonyms was further processed in a similar fashion to the filtering procedure of non-PVs described in 3.1.3, appropriate particles were used to extract possible PVs from the multiword synonyms. The search words and the extracted synonymous PVs were then tabulated in a spreadsheet side by side as a list, and vocabulary levels were assigned to all verbs present in the list following the same procedure with R described in 3.1.5. Consequently, a comprehensive list of single-word verbs catalogued in the EVP and their synonymous PVs with vocabulary levels corresponding to all verbs present in the list was generated.

3.2.1 Counting the types of synonyms

Maximising the filter function in MS Excel enabled the specification of particular PVs based on the vocabulary levels of their constituent verbs. This, in turn, facilitated the search capability for the corresponding SVs of those specified PVs. Consequently, specification of PVs whose constituent verbs belonged to A1 level in accordance with the EVP allowed a search for SVs which corresponded to PVs composed of A1-level verbs. The SVs identified in the process were extracted and had their duplicates removed such that types of SVs synonymous with PVs composed of A1-level verbs were revealed. Since each type of SV had been assigned a vocabulary level, it was made possible to group together the SVs based on their vocabulary levels. The number of SVs contained in each group was measured in comparison to the number of verbs contained in each level group of the EVP to calculate the percentage of representation exhibited by the SVs in each level group. Furthermore, the total number of SVs attained in the process was compared with the total number of verbs catalogued in the EVP (i.e., 1324 verbs) to reveal the extent to which PVs composed of A1-level verbs can be converted into SVs in relation to the total number of verbs listed in the EVP. The same procedure was performed on PVs composed of A1 & A2-level verbs to reveal the convertibility of PVs composed of A-level verbs into SVs. PVs composed of B1-level verbs as well as B2-level verbs were cumulatively added to the process, ultimately revealing the convertibility of PVs composed of all four levels ranging from A1 to B2 into SVs.

4 Results

4.1 A list of common PVs

The application of the CQL to the extraction of word combinations that would form PVs from CEPVC resulted in over 41,000 items which included such word combinations as *do not* and *see also*. After following the procedure of filtering out non-PVs by specifying particles following the verbs, restricting the relative frequency to "5 per million or above", and assigning vocabulary levels to the remaining

verbs with the help of R, a frequency list of 1,402 common PVs was created as shown in Table 1.

Ranking	Verb Level	Verb	Particle	Preposition	RF/million
1	A1	go	out	-	311
2	A1	go	on	-	276
3	A1	look	at	-	236
4	A1	come	in	-	233
5	A1	use	in	-	232
⋮	⋮	⋮	⋮	⋮	⋮
1398	C2	spark	off	-	5
1399	B2	cooperate	with	-	5
1400	C2	refrain	from	-	5
1401	C2	patch	up	-	5
1402	B2	sneeze	at	-	5

Table 1: PVs extracted from CEPVC.

4.1.1 Levels of verbs in common PVs

In pursuit of determining the degree of representation held by each vocabulary level associated with the verbs found in common PVs, the number of occurrences of each vocabulary level across the range from A1 to C2 present in the frequency list were tallied with the help of the table function in R. The investigation proceeded to increase the minimum relative frequency (RF) from 5 to 7, and ultimately to 10 to assess the degree of change in representation held by each vocabulary level as a function of RF. The result shows that 67% of 1,402 common PVs at a minimum RF of 5 were of A-level verb constructions. The degree of representation held by PVs composed of A-level verbs increased to 70% at a minimum RF of 7 with 999 common PVs. Finally, it was found that at a minimum RF of 10, the total number of common PVs stood at 724, and 76% of the PVs were composed of A-level verbs. The result shows that, on average, more than 70% of common PVs are composed of A-level verbs, which quantitatively confirms the intuitive notion that most PVs are constructions of basic-level vocabulary as shown in Table 2.

Level	Relative Frequency (per million)					
	~ 5		~ 7		~ 10	
A1	585	42%	442	44%	344	48%
A2	346	25%	262	26%	206	28%
B1	281	20%	187	19%	118	16%
B2	126	9%	82	8%	43	6%
C1	24	2%	10	1%	5	1%
C2	40	3%	16	2%	8	1%
Total	1402		999		724	

Table 2: Vocabulary levels of verbs in common PVs.

4.2 Convertibility of PVs to SVs

All 1324 single-word verbs registered in the EVP were used as search words for the collection of their synonymous PVs from the OTE. The identified PVs were then tabulated in a spreadsheet alongside their corresponding search words. Vocabulary levels were assigned to all verbs present in the list for the creation a comprehensive list of SVs and their corresponding PVs, which resulted in 10,899 entries as shown in Table 3. Consequently, 72 A1-level verbs and 88 A2-level verbs were determined to be capable of forming PVs. For the PVs composed of A1-level verbs, 909 unique SVs were identified as synonymous with such PVs, representing 69% of all verbs listed in the EVP. Furthermore, a total of 1073 unique SVs or 81% of the verbs catalogued in the EVP were found to be synonymous with PVs composed of A1&A2-level verbs. The addition of PVs composed of B1 and B2-level verbs to the PVs of A-level verb constructions only increased the percentage of SVs to 85% and 87% respectively. Interestingly, an addition of PVs composed of C-level verbs did not change the overall percentage of SVs synonymous with PVs. This finding shows that A-level verbs (i.e., A1 & A2-level verbs) are already capable of producing verb semantics delivered by more than 80% of verbs listed in the EVP when combined with particles, and that PVs composed of higher-level verbs account for less than 10% of verb semantics unrepresented by PVs composed of A-level verbs. The finding also shows that the A-level verbs found in the PVs, which represent 12% of the verbs in the EVP, have a significant impact or leverage in representing verb semantics which are supposedly confined in B/C-level single-word verbs. The breakdown of SVs and their corresponding vocabulary levels is shown in Table 4.

No.	SV Level	SV	Meanings	PV Verb	Particle	Verb Level
1	A1	answer	1	come	back	A1
2	A1	answer	1	get	back to	A1
3	A1	answer	1	write	back	A1
4	A1	answer	2	defend	oneself against	B1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
10896	C2	yield	4	comply	with	C1
10897	C2	yield	4	consent	to	C2
10898	C2	yield	4	go	along with	A1
10899	C2	yield	4	submit	to	B2

Table 3: A comprehensive list of SVs and PVs

EVP	A1	A2	B1	B2	C1	C2	TTL
	85	116	290	366	212	255	1324
PV	72						72
	85%						5%
SV	74	90	214	245	141	145	909
	87%	78%	74%	67%	67%	57%	69%
PV	72	88					160
	85%	76%					12%
SV	79	100	251	283	172	188	1073
	93%	86%	87%	77%	81%	74%	81%
PV	72	88	168				328
	85%	76%	58%				25%
SV	81	103	258	308	180	199	1129
	95%	89%	89%	84%	85%	78%	85%
PV	72	88	168	178			506
	85%	76%	58%	49%			38%
SV	82	104	259	313	183	205	1146
	96%	90%	89%	86%	86%	80%	87%
PV	72	88	168	178	61		567
	85%	76%	58%	49%	29%		43%
SV	82	104	259	313	184	205	1147
	96%	90%	89%	86%	87%	80%	87%
PV	72	88	168	178	61	103	670
	85%	76%	58%	49%	29%	40%	51%
SV	82	104	261	317	184	207	1155
	96%	90%	90%	87%	87%	81%	87%

Table 4: SVs' vocabulary levels corresponding to PVs of varied vocabulary levels.

4.3 Exploring polysemy

By rearranging the comprehensive list of SVs and PVs shown in Table 3 such that PVs were listed alongside their single-word synonyms, the levels of semantic complexity exhibited by polysemous PVs became accessible through the means of SVs and the vocabulary levels assigned to them. An example case with the PV *go through* is shown in Table 5 where four meanings contained in *go through* are expressed in the form of SVs. The first meaning of the PV is expressed in 11 unique SVs whose vocabulary levels range from A2 to C2. Even though assigning vocabulary levels to verb semantics can be difficult and is up to a certain level of subjectivity, the observation that B2-level verbs constitute most representation of the meaning suggest that the first sense of the PV belongs to the vocabulary level of B2, or at least belong to an intermediate level. Furthermore, the semantic level of the second meaning of *go through* can be determined by the SV that best captures the notion of *using up something* even though this type of judgement requires statistical analysis of intuition to overcome the inevitable subjectivity. As a by-product of investigating PVs, the study was able to produce a list of single-word verbs and their corresponding single-word synonyms, and it was determined that the same logic of determining levels of semantic complexity by means of single-word synonyms can be applied to the investigation of polysemous single-word verbs. An example case is shown in Table 6 where the verb *colour* is presented to have four distinct meanings. By observing the concentration of B2 and C-level verbs being synonymous with the semantics imparted by the third and fourth meanings of *colour*, it can be intuitively determined that the latter two senses held by *colour* belong to an advanced vocabulary level.

Level	PV Verb	Particle	Meaning	SV	Level
A1	go	through	1	receive	A2
A1	go	through	1	stand	A2
A1	go	through	1	experience	B1
A1	go	through	1	face	B1
A1	go	through	1	bear	B2
A1	go	through	1	endure	B2
A1	go	through	1	suffer	B2
A1	go	through	1	tolerate	B2
A1	go	through	1	undergo	C1
A1	go	through	1	sustain	C2
A1	go	through	1	withstand	C2
A1	go	through	2	spend	A2
A1	go	through	2	waste	B1
A1	go	through	2	consume	B2
A1	go	through	2	exhaust	C1
A1	go	through	2	squander	C2
A1	go	through	3	check	A2
A1	go	through	3	search	B1
A1	go	through	3	inspect	C1
A1	go	through	4	study	A1
A1	go	through	4	check	A2
A1	go	through	4	consider	B1
A1	go	through	4	analyse	B2
A1	go	through	4	examine	B2
A1	go	through	4	inspect	C1
A1	go	through	4	scan	C1

Table 5: Polysemous PVs expressed in SVs

Level	Word	Meaning	Synonym	Level
A1	colour	1	paint	A1
A1	colour	1	stain	C2
A1	colour	2	blush	B2
A1	colour	3	affect	B2
A1	colour	3	influence	B2
A1	colour	3	poison	B2
A1	colour	3	distort	C1
A1	colour	3	twist	C1
A1	colour	4	bend	B2
A1	colour	4	disguise	B2
A1	colour	4	strain	B2
A1	colour	4	distort	C1
A1	colour	4	enhance	C1
A1	colour	4	exaggerate	C1
A1	colour	4	overdo	C1

Table 6: Polysemy expressed in synonyms.

4.4 Interchangeability of single-word verbs

Table 6 demonstrates that A1-level single-word verbs behave in a similar way to that of PVs composed of A1-level verbs. By following the same methodology described in 3.2.1, it was determined that A1-level single-word verbs, which represent 6% of the entire verbs listed in the EVP, are

interchangeable with 61% of unique verbs catalogued in the EVP. Furthermore, 79% of all verbs presented in the profile were determined to be synonymous with A-level single-word verbs (i.e., A1 and A2 combined) as shown in Table 7. The results indicate greater expressiveness of PVs in comparison with single-word verbs since the semantic representation exhibited by PVs composed of A1-level verbs as measured by the number of corresponding SVs was 69% or 8% greater than that of their single-word counterparts. Interestingly, however, the interchangeability with synonyms exhibited by single-word verbs belonging to a range of vocabulary levels from A1 to B2 collectively demonstrated a 94% coverage of all verbs catalogued in the EVP, indicating that the majority of semantics required in communication can be accomplished by employing single-word verbs of up to level B2. The finding also shows that the number of SVs corresponding to PVs was capped at 87% of all verbs listed the EVP even with the inclusion of B2-level verbs as constituent verbs, while single-word verbs were seen to outperform PVs in their expressiveness after passing the B2-level threshold as determined by the number of corresponding synonyms.

EVP	A1	A2	B1	B2	C1	C2	TTL
	85	116	290	366	212	255	1324
Verbs	84						84
	99%						6%
Synonyms	77	90	198	218	117	110	810
	91%	78%	68%	60%	55%	43%	61%
Verbs	84	113					197
	99%	97%					15%
Synonyms	82	102	246	280	163	178	1051
	96%	88%	85%	77%	77%	70%	79%
Verbs	84	113	273				470
	99%	97%	94%				35%
Synonyms	84	106	265	324	188	225	1192
	99%	91%	91%	89%	89%	88%	90%
Verbs	84	113	273	356			826
	99%	97%	94%	97%			62%
Synonyms	84	108	273	343	199	240	1247
	99%	93%	94%	94%	94%	94%	94%

Table7: Interchangeability of single-word verbs with their single-word synonyms of varied vocabulary levels

5 Discussion

5.1 Summary of major findings

The current study has explored characteristics of PVs from three specific perspectives: vocabulary levels, convertibility, and polysemy. In investigating the vocabulary levels of common PVs, the majority of constituent verbs found in the 1,402 common PVs were judged to be of A-level classification as 70% of common PV-forms were found to be combinations of an A-level lexical verb and a particle occasionally followed by a preposition regardless of modifications made to minimum relative frequencies. Furthermore, certain lexical verbs and particles were found to be particularly productive in the formation of PVs. The top 20 most productive lexical verbs (i.e., *go, come, get, run, look, move, fall, take, keep, put, walk, stay, work, live, pull, grow, make, stand, bring,* and *hold*) of which *live* is the only B-level verb were collectively capable of forming 21.8 PVs on average at a minimum relative frequency of 5 per million, while the top 3 (i.e., *come, go,* and *get*) demonstrated their capability of producing 53 PV-forms on average. Likewise, the top 10 most frequent particles (i.e., *up, out, in, on, off, down, for, back, away,* and *with*) were each found to be part of more than 100 PV-forms on average while the top 3 (i.e., *up, out,* and *in*) being components of 169 PVs on average. The convertibility of PVs to SVs was investigated by means of synonyms contained in the OTE. Subsequently, PVs composed of A-level verbs (i.e., A1 and A2 combined) were found to be synonymous with 81% of all single-word verbs from a wholistic range of vocabulary levels from A1 to C2 catalogued in the EVP, prompting the study to conclude that PVs not only function as a bridge between vocabulary levels (i.e., PVs composed of A-level verbs acting as a bridge between A and B levels in semanticity specifically, while PVs composed of B-level verbs bridging between B and C level verb semantics), but also as "a free pass" allowing access to various tiers of semantic representations. In addition, the degree of verb semantics delivered solely by PVs composed of B-level verbs was determined to be relatively modest accounting for less than 10% of semantics unrepresented by PVs composed of A-level verbs, signalling the significance of A-level verbs in expressiveness when combined with particles.

Furthermore, polysemy exhibited by polysemous PVs and single-word verbs as well as possible vocabulary level classification of their various semantics was explored by considering the utilisation of single-word synonyms and their assigned vocabulary levels. Attempts at assigning vocabulary levels to verb semantics can be vulnerable to criticism as a high degree of subjectivity would inevitably be involved. However, the current study has successfully suggested a method that utilizes synonyms and their assigned vocabulary levels to provide a more objective approach in determining the levels of difficulty among the various semantics exhibited by polysemous verbs. Finally, the investigation into the interchangeability of single-word verbs with their single-word synonyms indicated that the semantic expressiveness exhibited by PVs composed of A-level verbs was greater than that of A-level single-word counterparts, while single-word verbs' semantics became greater than those of PVs after crossing the B2-level threshold.

5.2 Pedagogical implications

The current study has succeeded in incorporating the utility of vocabulary level categorisation brought forward by the EVP into clarifying the hitherto vague notion of *high-frequency* or *common* often associated with the descriptions of PVs. By observing the results obtained from the current study which indicate that PVs are vastly synonymous with single-word verbs, it stands to reason that learners would avoid PVs when the safer alternative of using single-word equivalents is readily available without taking the risk of misinterpretations and idiomaticity associated with PVs. Admittedly, PVs are not indispensable for conveying one's intentions, and single-word verbs are often more preferred in certain registers. However, since PVs are extremely common in spoken English, complete disregard for PVs in the classroom could inhibit learners' ability to comprehend details provided in situations where the use of PVs would be more appropriate, which are ubiquitous in the English-speaking community. With the knowledge from the current study that more than 70% of common PVs are composed of A-level lexical verbs, as well as the fact that 81% of single-word verbs catalogued in the EVP (or 1,073 single-word verbs) could be expressed by at least one PV composed of an A-level verb (see Table 4), certain measure of incorporating the utility of both PVs and single-word verbs into learners' lexical development could be proposed. For example, compilation of wordlists that display the relationship between PVs and SVs (single-word counterparts) could be considered. Table 5 could be such a wordlist that conveys PVs' semantic relations to their SVs, clearly demonstrating that higher-level single-word verbs could be expressed by A1-level verbs when combined with particles. By observing such wordlists, learners could clarify the meaning of newly encountered C-level verbs such as *exhaust* and *squander* in Table 5 by referring to their lower-level synonyms (e.g., *waste* and *consume*), or to their PV counterpart (i.e., *go through*), which could be construed as more semantically transparent. Additionally, semantically opaque versions of *go through*, such as the one listed as Meaning No. 4 in Table 5, could be familiarised with the help of transparency provided by B-level single-word verbs such as *consider, analyse,* and *examine*. Likewise, a collection of semantically opaque PVs such as *put up with* and *take after* could be listed and have their meanings clarified by the semantic concreteness provided by their single-word counterparts included in an example list shown in Table 8. Such a list could motivate learners to learn not only the meanings of ambiguous PVs expressed in SVs, but also the fact that A-level verbs such as *take* and *stand* could suggest *to tolerate* or *to endure*. Furthermore, the symbiotic relationship between PVs and SVs could be put to good use so as to eliminate the need for placing L1 translations alongside target

words which may only encourage memorisation of translated texts rather than the semantics of the target English words themselves. For instance, research has shown that access to external information such as dictionaries and glosses, as well as repeated exposure, foster the formation of form-meaning relationships within learners' lexicon (Hulstijn et al., 1996). Therefore, instead of relying on L1 translations, glosses that provide single-word equivalents of basic vocabulary levels corresponding to target vocabulary items may be proposed. Table 5 also indicates the potential efficiency in learning when PVs are used to good advantage as the 19 unique SVs on the list could easily be expressed by only one English phrasal verb presumably without the need for L1 translations since the verbs *go* is undoubtedly already known by learners. Furthermore, the current study has succeeded in identifying 169 single-word verbs catalogued in the EVP that are unexchangeable with PVs (see Appendix B). Such verbs include *change* and *walk*, and further study into why such verbs do not possess PV counterparts may shed light on more effective approaches to teaching and learning PVs.

Level	Single-word	→	PV Verb	Particle	Preposition	Verb Level
A1	take	→	put	up	with	A1
A2	stand	→	put	up	with	A1
B1	accept	→	put	up	with	A1
B1	support	→	put	up	with	A1
B2	tolerate	→	put	up	with	A1
B2	swallow	→	put	up	with	A1
B2	endure	→	put	up	with	A1
B2	bear	→	put	up	with	A1
B1	suggest	→	take	after	-	A2
B2	recall	→	take	after	-	A2
C1	resemble	→	take	after	-	A2

Table 8: Single-word verbs corresponding to PVs

6 Conclusion

This paper has demonstrated the utility of incorporating the vocabulary level classification provided by the EVP into investigating several characteristics of English phrasal verbs. By replacing such expressions as *high-frequency* and *common* with more precise account of CEFR-based level specifications such as, *A-level*, the current study succeeded in shedding light on the multi-faceted nature of phrasal verbs which involved vocabulary levels, convertibility to single-word verbs, and polysemy. The study has empirically confirmed the intuitive notion that phrasal verbs are combinations of basic-level verbs and particles with corpus-informed quantitative data which could be of use in encouraging learners to adopt phrasal verbs into their repertoire. Furthermore, the study has confirmed that certain symbiotic relationships between phrasal verbs and single-word verbs in vocabulary learning could be established and put to use in creating materials for pedagogical purposes. The effectiveness of phrasal verbs in assisting the development of learner vocabulary is a topic of further research. Moreover, the account of semantic transparency exhibited by phrasal verbs cannot be detached from subjectivity, which may complicate efforts in classifying what is transparent and what is opaque and in placing them along the cline of semantic transparency. However, the quantitative information regarding PVs obtained from the current study suggests that more than four fifths of all verbs indexed in the EVP have at least one PV counterpart composed of a basic-level lexical verb, and therefore, more learning resources other than L1 translations that take full advantage of PVs in vocabulary learning could be proposed and put to good use. In other words, English phrasal verbs could be one untapped resource that have been shunned by learners for too long. Further research into the relationship between phrasal verbs and single-word verbs may hold the key to drastically reducing the workload that learners have to handle, or *deal with*, when furthering their lexical knowledge.

References

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

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. Journal of Second Language Writing, 16(3), 129–147. https://doi.org/10.1016/j.jslw.2007.07.002

Dagut, M., & Laufer, B. (1985). Avoidance of phrasal verbs—a case for contrastive analysis. Studies in Second Language Acquisition, 7(1), 73–79. https://doi.org/10.1017/s0272263100005167

Darwin, C. M., & Gray, L. S. (1999). Going after the phrasal verb: An alternative approach to classification. TESOL Quarterly, 33(1), 65. https://doi.org/10.2307/3588191

Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. TESOL Quarterly, 41(2), 339–359. https://doi.org/10.1002/j.1545-7249.2007.tb00062.x

Garnier, M., & Schmitt, N. (2015). The phave list: A pedagogical list of phrasal verbs and their most frequent meaning senses. Language Teaching Research, 19(6), 645–666. https://doi.org/10.1177/1362168814559798

Gilquin, G. (2015). The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. Corpus Linguistics and Linguistic Theory, 11(1). https://doi.org/10.1515/cllt-2014-0005

Hulstijn, J. H., & Marchena, E. (1989). Avoidance. Studies in Second Language Acquisition, 11(3), 241–255. https://doi.org/10.1017/s0272263100008123

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental Vocabulary Learning by Advanced Foreign Language Students: The Influence of Marginal Glosses, Dictionary Use, and Reoccurrence of Unknown Words. The Modern Language Journal, 80(3), 327–339. https://doi.org/10.1111/j.1540-4781.1996.tb01614.x

Jakubíček, M., Kilgarriff, A., McCarthy, D., & Rychlý, P. (2010). Fast Syntactic Searching in Very Large Corpora for Many Languages. Proc PACLIC (Vol. 24, pp. 741-747). Japan.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten Years on. Lexicography, 1(1), 7–36. https://doi.org/10.1007/s40607-014-0009-9

Laufer, B., & Eliasson, S. (1993). What causes avoidance in L2 learning. Studies in Second Language Acquisition, 15(1), 35–48. https://doi.org/10.1017/s0272263100011657

Liao, Y., & Fukuya, Y. J. (2002). Avoidance of phrasal verbs: The case of Chinese learners of English. Language Learning, 54(2), 193–226. https://doi.org/10.1111/j.1467-9922.2004.00254.x

Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: A multicorpus examination. TESOL Quarterly, 45(4), 661–688. https://doi.org/10.5054/tq.2011.247707

Oxford University Press. (2006). Oxford Thesaurus of English.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Saville, N., & Hawkey, R. (2010). The english profile programme – the first three years. English Profile Journal, 1. https://doi.org/10.1017/s2041536210000061

Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. IRAL - International Review of Applied Linguistics in Language Teaching, 45(2). https://doi.org/10.1515/iral.2007.005

Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. Scotts Valley, CA: CreateSpace.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen

TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." /Journal of Open Source Software/, *4*(43), 1686. [doi:10.21105/joss.01686](https://doi.org/10.21105/joss.01686) .

A Appendices

Appendix A. The List of 25 Textbooks Referenced for the Compilation of the CEPVC

Booth, T., & Davies, B. F. (2021). English for everyone: English phrasal verbs. DK.

Burdine, S., & Barlow, M. (2008). Business phrasal verbs and collocations. Athelstan.

Cambridge Univ. Press. (2006). Cambridge Phrasal Verbs Dictionary.

Colby, H. E. (2014). 150 Useful English collocations, idioms, and phrasal verbs. BookBaby.

Dixson, R. J. (2004). Essential idioms in English: Phrasal verbs and collocations. Longman.

Emir, M., & Allans, R. (2019). Advanced English: Idioms, Phrasal Verbs, Vocabulary and Phrases: 700 Expressions of Academic Language. Independently Published.

Errey, M. (2007). 1000 Phrasal Verbs in Context. Teflgames.com.

Flockhart, J., & Pelteret, C. (2012). Work on your phrasal verbs: Master the 400 most common phrasal verbs. Collins.

Gul, I. (2020). 1000+ Phrasal Verbs With Meanings and Sentences. TheEnglishLover.com.

Harrison, J. (2003). Phrasal verbs. Stanley.

Hart, C. W. (2020). Phrasal Verbs. Barrons Educational Services.

Makar, A. (2020). 100 Practical English Phrasal Verbs. Independently Published.

McCarthy, M., & O'Dell, F. (2017). English Phrasal Verbs in Use Advanced Book with Answers: Vocabulary Reference and Practice. Cambridge University Press.

McIntosh, C. (2006). Oxford Phrasal Verbs Dictionary for learners of English. OUP Oxford.

Melvin, J. (2018). Phrasal Verbs and Idioms in Context.

Melvin, J. (2019). Phrasal Verbs : Practice Tests. English Language Academy.

Mordaunt, O. G., & McGuire, M. (2020). Phrasal Verbs for English Language Learners. Cambridge Scholars Publishing.

Parkinson, D. (2007). Really Learn 100 Phrasal Verbs: Learn the 100 most frequent and useful phrasal verbs in English in six easy steps. OUP Oxford.

Roche, M. (2020). Master English Collocations & Phrasal Verbs. Roche English Language Publishing.

Sandford, G. (2012). Amazingly Easy Phrasal Verbs! Praski Publishing .

Shepherd, D., Wagland, M., & Pinkney, R. (2015). Easy Phrasal Verbs: Learn English Through Conversations. Shaer Publishing.

Smith, D. B. (2020). English Phrasal Verbs Ultimate Collection.

Spears, R. A. (2006). McGraw-Hill's Dictionary of American idioms and Phrasal verbs. McGraw-Hill.

Spears, R. (2008). McGraw-Hill's Essential Phrasal Verbs Dictionary. McGraw-Hill.

Wyatt, R. (2006). Check Your English Vocabulary for Phrasal Verbs and Idioms. A & C Black.

Appendix B. EVP Verbs With No PV Equivalents

A1

be; change; walk

A2

boil; brush; camp; download; email; lend; matter; point; snow; surf; text; thank

B1

apologise; barbecue; blog; breathe; clap; cycle; deserve; fax; film; fry; grill; guide; hitchhike; iron; lock; owe; own; rebuild; sew; skate; ski; smell; smile; star; sunbathe; type; unpack; upload; vote

B2

alter; bark; benefit; blink; bookmark; bounce; coach; compromise; cruise; debit; doubt; enable; enquire; entitle; envy; fine; frighten; Google; gossip; guarantee; harm; kneel; link; misunderstand; participate; photograph; poison; pollute; punch; reward; rewrite; rip; rule; sentence; sneeze; sob; specialize; splash; spray; stare; steer; stroke; suspect; switch; terrify; unlock; whisper; whistle; yawn

C1

alternate; commute; distort; generalize; grade; hop; insert; modify; narrow; oblige; outnumber; outrage; presume; price; privatize; readjust; recharge; recreate; redevelop; relocate; rethink; scare; simplify; sip; smuggle; starve; summarize; surge

C2

amend; arch; blackmail; bond; cling; commemorate; diagnose; dice; drift; exemplify; filter; fluctuate; frown; gasp; gesture; giggle; glare; glue; grin; haul; hum; maximize; merit; misinterpret; misplace; moan; murmur; nest; overlap; pat; redistribute; reign; restructure; rhyme; riot; scar; shape; shrug; shudder; speculate; spit; sprinkle; spur; squeak; stain; vaccinate; weep; wink