

# Selection and Presentation of Phrasal Verbs in ESL Textbooks

Abdolvahed Zarifi  
Universiti Putra Malaysia, Malaysia

Jayakaran Mukundan  
Universiti Putra Malaysia, Malaysia

**Abstract**—The emergence of the lexical syllabus and the recognition of the key function of phrasal verbs in language acquisition and fluency gave way to the inclusion of these notoriously challenging structures in the ESL/EFL curriculum. Of main pedagogic concern, therefore, is, with the sheer number of phrasal verbs in the English language and the limited volume of course books, whether the selection and presentation of these forms is informed by research findings. Findings from the current corpus-based study revealed that that selection of these forms in the Malaysian ESL textbooks were more intuitively than empirically-based.

**Index Terms**—adverb particles, corpus linguistics, content analysis, ESL materials, phrasal verbs, textbook analysis

## I. INTRODUCTION

During the past few decades, with many researchers and educators shifting attention from syntax to vocabulary in second language education (Folse, 2004; Laufer, 1997), there has appeared a growing interest in multiword vocabulary items, including phrasal verbs. The English phrasal verb combinations are one of the most notoriously challenging aspects of English language instruction (Celce-Murcia & Larsen-Freeman, 1999; Gardner & Davies, 2007; Siyanova & Schmitt, 2007) or in Mullany and Stockwell's (2010, p. 201) words "the scourge of the learner" as they present a host of inherent difficulties. Despite their rather complicated structure and unpredictable meaning of some combination types, phrasal verbs are of high relevance for ESL/EFL learners because a grasp of them "can be a great asset to learners in acquiring a new language" (Celce-Murcia & Larsen-Freeman, 1999).

## II. REVIEW OF LITERATURE

The phrasal verb combinations have been extensively dealt with in non-corpus-based studies (Dagut & Laufer, 1985; Hulstijn & Marchena, 1989; Laufer & Eliasson, 1993; Laufer and Hulstijn, 1996; Laufer & elMorales 2000; Sjöholm 1995; Liao and Fukuya 2002; Schmitt 2007). Likewise, corpus linguistics has witnessed an encouraging body of research on phrasal verbs in general and learner corpora (Gardner and Davies, 2007; Trebits, 2009; Akbari, 2009; Von, 2007; Schneider, 2004). There is, however, a dearth of studies dealing with the use of these mysterious structures in the instructional materials (Gardner and Davies, 2007; Koprowski 2004; Zarifi & Mukundan, 2012). This small body of research, unfortunately though, provides some significant evidence supporting a remarkable divide between research findings and the inclusion of phrasal verbs in ESL materials.

Darwin and Gray (1999), for instance, developed a list of the 20 most frequently occurring phrasal verbs in the BNC. Comparing the list with the phrasal verbs in a typical ESL grammar book, they found that only three of all the phrases in the textbooks matched the 20 phrasal verbs on the list. Likewise, Koprowski (2005) studied the use of phrasal verbs and other multiword expressions in three contemporary ELT course books. Although they were all developed as British general English materials for learners at the intermediate level, not even a single phrasal verb was found to be shared by the three books. Lamenting the lack of consistency among the textbooks, she observed that ELT materials developers do not follow any principled criteria in the process of vocabulary selection. They rather arbitrarily decide upon the selection and inclusion of these items based on their own personal experience and intuition. In another recent study on the Malaysian ESL textbooks, Zarifi and Mukundan (2012) investigated the use of phrasal verb combinations in the spoken sections of the materials. Findings of the study revealed that both the selection and presentation of these combinations were inconsistent with their actual use in the BNC. They reported that textbooks contained combinations of extremely low frequency counts in general English and vice versa.

In line with the research evidence from the literature, selection and presentation of phrasal verb combinations turns out to be a major pedagogical concern. With the phrasal verbs overwhelmingly present in the language and different collections available to choose from, curriculum designers and materials developers alike often feel frustrated with which phrasal verbs and how many of them to include. This sense of frustration is complicated by the limited volume of course books and the fact that not all of these phrases are of equal use to EFL/ESL learners. That being said, it is worth

considering whether ELT developers are really taking into account the research findings and pedagogical principles in selecting and presenting the most useful phrasal verb combinations. The current study, therefore, aimed to identify all the instances of phrasal verb combinations and their constituents, determine their frequency of occurrence and investigate their distribution in the Malaysian ESL textbooks.

### III. METHODOLOGY

This is a corpus-based content analysis of the textbook materials Forms One through Form Five prescribed for use by the learners at the secondary level in Malaysia. This pedagogic corpus consists of. The BNC was also used as the reference corpus. The study involves a comprehensive data sampling (Ary, Jacob, Razavieh, & Sorensen, 2006) since all the instances of phrasal verb combinations in the corpus were included in the analysis. As to the instrumentation of the study, the WordSmith tools version 4.0 was used to hunt down the potential phrasal verbs in the corpus. In a similar way, the Zar-Test was employed to identify the different types of phrasal verbs. This test is based on the basic notion that while a preposition makes a whole unit with its following NP, and the real particle forms an integrate unit with its preceding lexical verb, the adverb particle is almost independent of the constituents it keeps company with.

### IV. RESULTS AND DISCUSSION

The concordance function of WordSmith Tools 4.0 yielded a total number of 16826 particle forms. Looking into the concordance lines, the researchers opted out a large number of cases from the data since they were not preceded by any lexical verb, hence absolutely no potential candidates for phrasal verbs. Only 9060 out of the 16826 particle forms were following a lexical verb, and were, therefore, opted for further investigation. Since these particle forms could feature either a Preposition, an Adverb or a Real particle, the researchers went on to distinguish them from one another against the Zar-Test (Zarifi, 2013). Identification of phrasal verb combinations in terms of the particle type yielded the following results:

TABLE 1  
FREQUENCY OF DIFFERENT PARTICLE FORMS IN THE CORPUS

Word tokens	Verb+particle	Verb+Prep	Verb+Adv	Verb+Real Part
302642	16826	6848	232	1980

As the table shows, an overwhelming number of 6848 (40.70%) particle elements following a lexical verb acted as Prepositions. 1980 (11.77%) of the remaining forms behaved as Real Particles, and a negligible number of 232 cases (1.38%) occurred as Adverbs.

It should be pointed out that due to some pedagogical considerations, only the combinations made up of a lexical verb followed by an Adverb Particle or a Real Particle were viewed as phrasal verbs. Therefore, prepositional components are no further dealt with in the study but in some rare cases for comparison.

TABLE 2  
GENERAL DESCRIPTIVE STATISTICS OF PVs IN THE CORPUS

PV tokens	PV Lemmas	LV Lemmas in PVs	Particle elemnets in PVs
2212	464	226	23

As table 2 reveals, combination of 226 different lexical verbs with 23 different particle forms made up a total number of 2212 tokens of phrasal verbs. These phrasal verb combinations appeared in 464 different lemmas. In other words, on the average, each lemma appeared roughly in 4.77 tokens of phrasal verbs, and each lexical verb type took part in the formation of 9.79 phrasal verbs.

TABLE 3  
TOKENS OF WORDS AND PARTICLE FORMS IN THE CORPUS  
(TOKEN FREQUENCIES NORMALIZED TO 10,000 TOKEN CORPUS AND ROUNDED TO THE NEAREST INTEGER)

Form	Tokens	No. of PVs	Normalized to 10000 tokens
One	44672	347	78
Two	48485	345	71
Three	57918	395	68
Four	72936	583	80
Five	78631	542	69

As it is shown in table 3, the most number of phrasal verbs occurred in a descending order in Forms Four, Five, Three, One and Two. However, in order for the figures to be directly comparable, they were normalized (Evison, 2010; McEnery & Hardie, 2012; Schneider, 2004). To this end, the observed frequency counts of phrasal verbs in each Form were projected to a corpus basis of exactly 10,000 words. The normalized figures enabled the researchers to directly compare the frequencies of the combinations.

Based on the above table, Form Five had the second largest and Form One had the second smallest number of combinations. However, normalization of the data indicates that Forms Five and One contained, in fact, the first lowest

and the second largest number of units, respectively. Given the complicated features of the phrasal verbs and the problems they could create for ESL learners, the presentation of the combinations, with other criteria being observed, is unjustified in terms of the developmental stages of the learners at each level. With Form One students being at the lowest level, and Form Five students at the highest level of developmental stage across the school board, we would expect just the opposite.

Table 4 presents the total number of all grammatical tags for each of the 23 particle forms in the corpus. It shows the number of times each of these forms was tagged as a particle, including Real Particles and Adverb Particles, as opposed to a preposition and/or other grammatical categories. Out of a total of 16826 particle forms in the corpus, 2212 occurrences (13.15%) turned out to act as particles forming phrasal verb structures. Most of the cases (14035 or 83.90%) served as prepositions and a small number of them (586 or 3.48%) acted as other functions such as adjectives, nouns and verbs in the corpus. Needless to say that these particle elements behaved quite idiosyncratically in terms of the functions they served. Although such elements as 'UP, OUT, DOWN and AWAY' acted almost always as particle forms, a few other items like 'BY, FOR and WITH' almost always behaved as prepositions.

TABLE 4  
FREQUENCY OF PART/PREP FORMS AND THEIR FUNCTIONS IN FORMS 1-5

Part/Prep	Freq	Part	Part%	Prep	Others
About	1108	17	1.53	1073	18
Across	57	8	14.04	38	13
After	339	48	14.16	292	0
Ahead	16	7	43.75	3	8
Along	77	20	25.98	54	11
Apart	5	1	20.00	0	4
Around	121	23	19.01	83	15
Away	194	140	72.16	17	65
Back	132	76	57.58	0	81
By	925	6	.65	905	18
Down	243	180	74.07	29	51
For	2411	2	0.08	2355	27
Forward	34	32	94.12	0	2
In	5407	190	3.51	5204	10
Off	138	94	68.12	12	29
On	1859	113	7.11	1734	40
Out	689	587	85.20	80	71
Over	209	50	23.92	113	56
Part	143	35	24.48	0	108
Round	43	5	11.63	21	17
Through	189	47	24.87	137	4
Up	632	509	80.54	16	111
With	1855	22	1.20	1833	0
Total 23	16826	2122(13.15%)		14035(83.90%)	586(3.48%)

Table 5 presents the frequency occurrence of Particle/Preposition elements and their function in each Form. As it is shown, the proportion of particles as a grammatical category to the total times of occurrences of particle/preposition elements in all the five Forms appeared to be almost similar, ranging from 11.85% to 14.57%. Chi-Square statistics of the normalized data in table 3 also revealed that there was no meaningful difference between the five textbooks in terms of the observed frequency of phrasal verbs at 0.05 level of significance. Despite the similarity between the textbooks with regard to the total number of particle elements, the proportion among the different particle elements within and across the textbooks was not consistent. For instance, elements like 'IN' (5407) and 'ON' (1859) appeared to have a strikingly higher frequency of occurrence in comparison with units such as 'OUT' (689), 'UP' (632) and 'OFF' (137); they, nonetheless, had a very poor percentage of occurrences as particle forms. They occurred as particles only 3.51% and 7.11%, respectively. The particles 'OUT, UP, and OFF', on the other hand, served 85.20%, 80.54% and 68.12% as particles, respectively. One point should be mentioned about the frequency count of FORWARD as a particle. Despite its high frequency in phrasal verb constructions, FORWARD behaved more as an adverb than as a real particle in most of the cases. This happened simply because FORWARD, as a directional particle element, can freely combine with all the verbs of motion like MOVE, GO, DRIVE, RUN, etc.

TABLE 5  
FREQUENCY OF PART/PREP FORMS AND THEIR FUNCTION IN EACH FORM

Form	Part/Prep type	Part Freq	Part%	Prep Freq	Freq as others	V+Part/Prep
One	22*	347	14.57	1915	117	2381
Two	22*	345	12.99	2245	65	2655
Three	23	395	13.03	2540	95	3030
Four	23	583	13.92	3452	154	4187
Five	22*	542	11.85	3876	155	4573

\*There was no record for 'ahead' in Form one; 'apart' was also absent from Forms Two and Five.

With particles having a number of confusing functions as real particles, adverbs, prepositions, nouns, adjectives, etc. in the language, it comes as a shock why these forms were so highly frequent in Form One and so infrequent in Form Five. More surprisingly, while phrasal verbs have been reported to be grammatically, semantically and orthographically complicated and challenging to ESL learners (Sawyer, 2007), the proportion of these combinations across the five Forms was far from pedagogically justified. Pedagogically speaking, one would expect an increase in the rate of presentation of these fuzzy forms in line with increase in language proficiency level of the learners; however, the frequency counts of these forms in the Malaysian ESL materials was just the reverse. That is, the highest proportion of these forms occurred in Form One and the lowest occurred in Form Five. This state of imbalance of the presentation of particle elements and phrasal verbs across the textbooks highlights the criticism often made of the instructional materials, namely the inclusion and organization of teaching materials in textbooks is largely based on the assumptions and intuitions of the writers (Mukundan, 2004; Moon, 1998; Koprowski, 2005).

The sheer number of prepositions in the corpus was due to the fact that some forms like 'WITH, FOR and BY' tend to serve almost exclusively as prepositions in the English language. Some others like 'ON and IN' show more tendency towards prepositions than towards particles (O'Dowd, 1994). Interestingly enough, the elements that function more as prepositions are overwhelmingly more frequent than those that function more as particles in general English. For instance while 'IN, FOR, and WITH' have frequency counts of 1944328, 883599, and 659332 in the BNC, respectively, the frequency counts for 'OFF, OUT, and UP' are 67880, 197149 and 207521, respectively. Furthermore, those other elements that could function as particles were infrequently preceded by any lexical verb in the corpus. In other words, they failed to act as particles depending upon the context in which they appeared. Moreover, not any sequence of verb + particle form signals a phrasal verb. There were noticeable sequences of this configuration that behaved as prepositional rather than phrasal verb combinations. Still in some other cases, these elements occurred in noun phrase combinations. For instance, 'UP' which mostly functions as a particle both in general English and in the corpus, does not form a phrasal verb in the utterance "Are you taking mother to Dr Chan for a CHECK-UP this morning?" Likewise, there occurred a number of sequences like "I CARRIED the small children OUT of the bus" that feature prepositional verbs rather than phrasal verbs. As a result, the particle forms acted either as prepositions or other forms like nouns, verbs or adjectives most of the time.

The particle forms behaved quite idiosyncratically with regard to the functions they carried in the corpus. Some forms like 'PART and APART' never acted as prepositions; some items like 'WITH and FOR' almost always acted as prepositions; and some others such as 'FORWARD, BACK, and OUT' served rather exclusively as particles. There were still a few other forms like 'UP, DOWN, and AWAY' which acted more noticeably as particles than as prepositions, and there existed some other items like 'BY, IN and ON' which tended to act more preferably as prepositions. The overall totals indicate that these forms appeared about 13.15% of the time in the phrasal verb combinations.

Although the statistics appeared to be satisfactory regarding the occurrence of these particle forms when compared against the BNC, it is, nevertheless, a bit misleading. As it will be shown later in table 9, the high frequency of occurrence of these forms was mainly due to the fact that Malaysian ESL textbooks, like other ELT materials, turned out to contain some phrasal 'teddy bears' for instructional functions such as 'FILL IN, ZOOM IN, SOUND OUT, LOOK UP, CHECK OUT, WRITE DOWN, etc.' which were used in the instructions given to the students at the beginning of the exercises. These phrases were far from pedagogically exploited and are most likely to escape the learners' attention just because the formats of the exercises in themselves indicate what the students are required to do. For instance, seeing a list of words followed by a couple of sentences with blank spaces, the students, with no attention to the instruction containing the combination 'FILL IN', could figure out that the blanks are required to be *filled in* with the words given. Reduction of these forms from the analysis reveals some degree of discrepancy between the pedagogic corpus and the natural use as empirical corpus-based evidence indicates that phrasal verbs as a grammatical category appear once per 150 words in actual use (Gardner & Davies, 2007), that is, about one and half times as much as their frequency rate in our pedagogic corpus. The observed imbalance can, however, be interpreted in terms of the make-up of the BNC since it comprises a number of different genres such as conversation, fiction, academic prose, etc.

Table 6 reveals the ten top frequent particle forms in phrasal verb combinations. As it is shown, 'OUT and UP' were the most highly frequent particles in the corpus with a frequency count of more than 500 cases each. The presentation of these two particles in the corpus seemed to be in keeping with their use in general English as Gardner and Davies (2007), Kennedy (1998) and Armstrong (2004) reported 'UP and OUT' as the most frequent particles in the BNC and Brown corpora. The particles 'IN, DOWN and AWAY' were the other high frequent forms occurring more than 100 times in the corpus. As it was already mentioned, despite the large number of times that 'IN (193)' was used as a particle, this rate was quite negligible compared with the number of times the element occurred in the corpus. It served as a particle for less than 5% while other low occurring particles like 'OFF (94) and BACK (76)' occurred with a noticeably higher ratio of 68.12% and 57.58%, respectively.

TABLE 6  
TOP TEN HIGHLY FREQUENT PARTICLES IN THE CORPUS

Part	Freq as Particle					Total Freq	Cum Freq	Cum freq%	Rank
	F1	F2	F3	F4	F5	F 1-5	F 1-5	F 1-5	F 1-5
Out	101	110	108	160	108	587	587	29.56	1
Up	66	95	93	118	137	509	1096	55.19	2
In	35	13	43	29	70	190	1286	64.75	3
Down	20	33	32	49	46	180	1466	73.82	4
Away	26	24	21	32	37	140	1606	80.87	5
On	16	15	11	42	29	113	1719	86.56	6
Off	22	8	15	26	23	94	1813	91.29	7
Back	13	13	7	26	17	76	1889	95.12	8
Over	8	6	5	13	18	50	1939	97.63	9
Through	0	2	5	28	12	47	1986	100	10

Table 6 also reveals that particle forms followed different patterns of behavior in terms of frequency counts within and across the corpus. For instance, while ‘OUT’ was more frequent than ‘UP’ in Forms One through Four, it was less frequent in Form Five. Likewise, although THROUGH was more frequent than ‘OFF, BACK and OVER’ in Form Four, it failed to occur in Form One at all. In addition, ‘OUT and UP’ were, to a higher degree, more frequent than other items. Phrasal verb tokens involving these two particles accounted for 50.11% of all the combinations in the corpus. To put this observation into a more practical perspective, acquisition of the phrasal verbs containing ‘OUT and UP’ as their particle elements would enable the students to understand and use a larger number of phrasal verb constructions of the textbooks language than if they acquired the remaining 21 particles and their associated verb components which covered only 49.89% of all the occurrences in the corpus.

TABLE 7  
TOP TEN PRODUCTIVE PARTICLES

Part form	No. of LVs combined with each Part
Up	88
Out	79
Down	48
Away	41
Off	36
Back	34
On	29
In	24
Over	17
Around	14

Table 7 presents the top ten productive particle forms. Although all of these particles but for AROUND were also reported as the top frequent forms, they fell in different rank orders in terms of productivity. For instance, while ‘OUT’ was more frequent than ‘UP’ as a particle, it was, nevertheless, less productive combining with 79 different lexical verbs. ‘UP’ was, on the other hand, used with a larger number of lexical verbs, combining with 88 various lexical verb items across the five Forms. This observation also held true for particles ‘IN and ON’ and ‘OFF and BACK’. Despite the higher frequency counts of ‘IN and ON’ (see table 4), they, notwithstanding, occurred with fewer lexical verbs than ‘OFF and BACK’.

TABLE 8  
MOST PRODUCTIVE LVs IN FORMING PVs

LV	Part	Part type
Go	14	about, ahead, along, around, away, back, down, off, on, out, over, round, through, up
Get	12	about, along, around, back, down, in, off, on, out, over, through, up
Come	11	about, across, along, back, down, forward, in, on, out, over, up
Take	11	along, away, back, down, in, off, on, out, over, part, up
Turn	11	around, away, back, down, in, off, on, out, over, round, up
Look	11	after, ahead, around, back, down, forward, out, over, round, through, up
Put	10	across, away, back, down, forward, in, off, on, out, up
Move	9	about, along, around, away, back, forward, in, on, over
Bring	7	about, along, back, down, in, out, up
Send	7	away, back, down, in, off, out, up

Table 8 presents the most productive lexical verbs in forming phrasal verb sequences in the corpus. ‘GO’ turned out to rank first, combining with 14 different particles like ‘ABOUT, AHEAD, ALONG, AROUND, AWAY, BACK, DOWN, OFF, ON, OUT, OVER, ROUND, THROUGH, and UP’. Next in rank was ‘get’ combining with 12 particles. Other highly productive lexical verbs included ‘COME, TAKE, TURN, LOOK, PUT, MOVE, BRING, and SEND’. It is important to point out that almost all the potential verb-particle structures of a few of these lexical items were covered in the corpus. For instance, ‘GET, LOOK and TURN’ appeared in the corpus with nearly all their potential particles in general English, hence a good quantitative, though not equally qualitative, presentation of these units. In other words,

although they occurred with the most possible particle forms, not all the different word meanings of each of these forms were presented and adequately recycled. On the other hand, the potential structures with a few other lexical verbs turned out to be restrictively represented. For instance, while ‘GO’ was used with 14 different particles in the corpus, it occurs with 18 different particles in the real language use. Likewise, ‘COME’ was used with only 11 out of the 17 particles it could potentially combine with.

TABLE 9  
TOP 20 PV LEMMAS IN THE CORPUS AND THEIR BNC RANK ORDER

PV	Freq	Corpus rank	BNC rank
Find out	138	1	9
Fill in	122	2	*
Carry out	66	3	2
Write out	58	4	*
Look up	49	5	26
Look after	44	6	*
Pick out	43	7	75
Write down	35	8.5	*
Take part	35	8.5	*
Pick up	32	10	4
Cut down	31	11	*
Go out	30	12	7
Put up	29	13.5	33
Wake up	29	13.5	*
Go through	28	15.5	73
Throw away	28	15.5	*
Give up	27	17	24
Set up	26	18	3
Check out	25	19	*
Take up	24	20	19
	899		

To look at the data from another angle, the ten top lexical verbs which comprised only 4.46% of all the lexical verbs in phrasal verb constructions in the corpus accounted for about 31.90% of all the phrasal verb tokens. In other words, roughly about 1 out of 3 phrasal verb lemmas in the corpus incorporated one of these 10 forms as their lexical verb component, suggesting the noticeable productivity of these verbs and their remarkable tendency towards combining with particle forms. The use of these lexical verbs and their associated phrasal verbs were not appropriately spaced and graded across the five Forms, however. For instance, there occurred no phrasal verb with the lexical verb ‘GET’ in Form 2, and there happened only one phrasal verb with ‘TURN’ in the same Form. Therefore, the presentation and sequencing of these combinations appeared to be more intuitively than pedagogically motivated (Mukundan, 2004; Sinclair, 1991).

Table 9 presents the top 20 phrasal verb lemmas in the corpus along with their frequency rank orders in the BNC. These combinations altogether made up 4.33% of all the phrasal verb lemmas but 41.09% of the phrasal verb tokens in the corpus. Presentation of these units in terms of recycling was, however, far from satisfactory, with some being overwhelmingly over-repeated at the expense of some others. For example, the units ‘FIND OUT’ and ‘FILL IN’ alone accounted for 21.40% of all the instances of these high frequent forms. In addition, the frequency counts of the shared items between the two corpora are not consistent. For instance, ‘LOOK UP’ and ‘GO THROUGH’ that ranked 5th and 15.5th in our corpus in the same order, rank 26th and 73rd, respectively, in the BNC corpus (Gardner & Davies, 2007). Furthermore, there is a huge extent of discrepancy between the presentation of these forms in the corpus and the BNC. As it is shown, 9 out of these top 20 combinations were not even among the top 100 phrasal verbs in the BNC, providing empirical evidence in favor of the often-made observation that inclusion of lexical items in textbooks are not informed by research findings (Koprowski, 2005). The inclusion or exclusion of these items was, it is contended, rather unwittingly intuitively motivated. Implied here is that the learners were denied sufficient opportunity to come across almost half of the forms that are highly frequent in natural use of the language.

TABLE 10  
LEMMA FREQUENCY OF PVs IN FORMS 1-5

PV Lemma	Frequency	Item No.
Find out	138	1
Fill in	122	1
Carry out	67	1
Write out	58	1
Look up	49	1
Look after	44	1
Pick out	43	1
Take part, write down	35	2
Pick up	32	1
Cut down	31	1
Go out	30	1
Put up, wake up	29	2
Go through, throw away	28	2
Give up	27	1
Set up	26	1
Check out	25	1
Take up	24	1
Log on, deal with	23	2
Clean up, zoom in	20	2
Go back, take over	19	2
Act out	18	1
Go on, speak up, warm up	17	3
Come up, make up	15	2
Leave out, write away	14	2
Look forward, sound out	13	2
Come on, run away, switch off, wrap up	12	4
Fill up, get up, sit down	11	3
Break down, end up, take off	10	3
Go away, grow up, hand in, keep up, put on, reach out	9	6
Bring about, cut off, read through, sign up, take down, ...	8	7
Come across, come back, get along, help out, read out, ...	7	11
Bring up, check on, go ahead, open up, set out, stand up, ...	6	10
Break out, call up, hurry up, start off, sum up, take on, ...	5	14
Build up, burn off, call off, cheer up, come forward, ...	4	28
Blow up, eat up, put off, try out, watch out, wonder off, ...	3	39
Clear up, fall down, set off, speak out, turn out, work out, ...	2	76
Back off, call on, dress up, mark down, shy away, top up, ...	1	227

Table 10 lists down the identified lemmas of phrasal verbs and their frequency of occurrence in the corpus (see Appendix). As the table shows, the whole textbooks corpus contained 464 cases of lemmas out of which 227 (48.92%) items were hapaxes, occurring only once in the corpus. Out of the 464 lemmas, 394 instances (84.91%) had a frequency count of less than seven. If we subtract the highly frequent cliché forms (23 forms) like ‘FILL IN, FIND OUT, LOOK UP, SOUND OUT, ZOOM IN, etc.’ used in the instructions on the exercises, there would remain only about 46 cases (10.12%) with a frequency count of seven or more occurrences. It is pedagogically disappointing that the learners were denied the chance of revisiting the combinations sufficiently across the corpus to consolidate their learning.

In need of special attention would, perhaps, be the frequent combinations in the ESL corpus that are infrequent in general English and the other way round. For example, ‘CLEAN UP, THROW AWAY, CUT DOWN, WAKE UP, and LOOK AFTER’ which were among the most frequent forms in this pedagogic corpus are not included in the 100 top phrasal verbs in general English. Moreover, much to the surprise of the researchers, some of the phrasal verb combinations that are listed down among the top 100 lemmas in general English such as BREAK OFF, COME OFF, SET DOWN, HOLD UP, HOLD OUT, SIT BACK; SET ABOUT and COME THROUGH were missing from the corpus, but some highly infrequent forms like ‘WHAM BACK, RUSTLE OUT, SPROUT OUT, FLICK AWAY, PELT DOWN, LAZE AROUND, etc. did appear. The BNC query cropped out frequency shots of 0, 1, 4, 9, 15, and 20 for each of these combinations, respectively. More surprising was that most of these highly infrequent combinations like ‘WHAM BACK, LAZE AROUND and POKE ABOUT’ were introduced in the Form One textbook.

Given the inevitability of an enormous knowledge of vocabulary for the ESL learners’ success in both their academic career and international communication, the researchers hold back to stand against the inclusion of the wide range of phrasal verbs in the corpus; however, what the researchers are trying to make an almost firm stand against is why relatively infrequent phrasal verbs should be presented to the exclusion of frequent forms at least in the lower level Forms. This position is in full agreement with Biber and Conrad’s (2001, p. 335) observation “dramatic differences in frequency should be among the most important factors influencing pedagogical decisions”. This is considerably important as psycholinguistics indicates that word frequency affects word familiarity which, in turn, serves as a major factor in word recognition (Alderson, 2007).

Despite the criticism leveled above against the inclusion of the infrequent forms, different interpretations can be raised as to the overrepresentation of a few of these forms in the Malaysian ESL textbooks, however. To begin with, the



high frequency of some combinations can be interpreted in light of the writers' tendency to organize the selection of lexical items including phrasal verbs thematically (Koprowski, 2005). They probably began with a topic and then went on to introspectively include lexical items that appeared to fit in. A case in point is the phrasal verb 'CLEAN UP' which got repeated for five times in lesson 11, Form 1, dealing with the topic of a 'caring society' in its dialogue section. Second, the overuse of some other forms appeared to be culturally motivated. For instance, due to the tropical weather conditions and environmental issues in Malaysia, and the role that jungles play in this connection, 'CUT DOWN' was semantically relevant to the texts on these issues, hence highly overused in the textbooks. Third, the high frequency rate of some forms like 'FILL IN, WRITE OUT, WRITE DOWN, SPEAK UP, WARM UP' and some others could be attributed to the relevance of these forms to the bookish and scholastic register of the language (Cornell, 1985). These are among the common phrases that are likely to crop up in instructions on different types of exercises in textbook materials. Fourth, some others like 'THROW AWAY and GO THROUGH', it can be argued with less certainty, were unwittingly intuitively overused because the textbook developers had probably no access to the empirical findings on the frequency counts of phrasal verbs. As a result, these low frequent combinations were overused at the expense of some other forms which could be pedagogically more valuable to the learners.

Among others, frequency of occurrence and recycling of different language phenomena are important factors affecting language learning. However, as table 8 shows more than 85% of the phrasal verb combinations in the corpus had a frequency count of less than seven. Thornbury (2002) reiterates that words with a minimum occurrence of seven times over spaced intervals stand a good chance of being remembered. Likewise, Celce-Murcia and Larsen-Freeman (1999) consider it pedagogically sensible for different aspects of language to be recycled and elaborated on over a period of time. Indeed, for the ESL learners to master phrasal verbs, the most notoriously challenging aspect of the language, they need not only to be exposed to the most frequent forms which are, by the same token, the most useful ones but also to frequently revisit them at spaced intervals in the follow-up activities and lessons.

With that being said, care should be exercised in pedagogical interpretation of the lexical items of seven and more frequency counts. In other words, just because a phrasal verb gets repeated seven or more times (Thornbury, 2002), it cannot be concluded that it is optimally appropriate for learning since chances are that each occurrence presents a single specific meaning of the target combination.

Despite the above mentioned shortcomings of the Malaysian ESL textbooks in dealing with phrasal verbs, one promising characteristic of the use of these combinations, among others, could be the number of shared units among the five Forms. Although Koprowsky (2005) found not even a single lexical phrase including phrasal verbs shared by all the three course books he investigated, and only a negligible number of seven items shared by any of them despite their being targeted for the same level students, the empirical evidence from the present pedagogic corpus painted a very different picture. Data analysis revealed that the textbooks, albeit designed for different levels, shared a noticeably large number of phrasal verbs. For instance, a total number of 16, 23, 50, and 69 combinations were shared by five, four, three and two Forms, respectively. To put it into perspective, these shared combinations appeared to be appropriately recycled across the different levels. This is pedagogically significant for recycling, as a teaching technique, brings about consolidation of learning (Ur, 2006). Likewise, Celce-Murcia and Larsen-Freeman (1983) consider it sensible for different language items to be recycled through revisiting old items, elaborating on them and comparing and contrasting them with new items as they emerge.

## V. CONCLUSION

Findings revealed that both the selection and presentation of the phrasal verb combinations in the Malaysian ESL textbooks were not in good keeping with the use of these forms in actual language use. The researchers hold back to reason that such incongruity is a drastic pedagogic deficiency on the part of the textbooks but tend to argue that not only pedagogical factors like difficulty level, learnability, availability, etc. but also profitability issues like usefulness and encounter probability should be given due weight in the selection of these fuzzy forms as well. Ranali (2003) is infact holding the same stance by arguing that corpus-based findings need "to be creeping into ELT slowly over time" (2003, p. 3). Similarly, with the inherent difficulty of the phrasal verb combinations in mind, we would make a firm stand for the appropriate recycling of these forms in the instructional materials. In other words, we tend to concur with McPartland (1983, p. 155) who stated "the frequent occurrence of a phrasal verb in the input seems to accelerate the acquisition process, overriding semantic, syntactic, and phonological complexity".

## REFERENCES

- [1] Alderson, C. (2007). Judging the word frequency of English words. *Applied Linguistics*, 28(3), 383-409.
- [2] Biber, D., & Conrad, S. (2001). Quantitative corpus-based research: Much more than bean counting. *TESOL Quarterly*, 35(2), 331-336.
- [3] Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course*. Boston: H & H Publishers.
- [4] Darwin, C., M., & Gray, L. S. (1999). Going after the Phrasal Verb: An Alternative Approach to Classification. *TESOL Quarterly*, 33(1), 65-83.
- [5] Evison, J. (2010). What are the basics of analyzing a corpus. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 122-136). London: Routledge.



- [6] Folse, K. (2004). Vocabulary myths: Applying second language research to classroom teaching. Ann Arbor: University of Michigan Press.
- [7] Gardner, D., & Davies, M. (2007). Pointing Out Frequent Phrasal Verbs: A Corpus-Based Analysis. *TESOL Quarterly*, 41, 339-359.
- [8] Koprowski, M. (2005). Investigating the usefulness of lexical phrases in contemporary coursebooks. *ELT Journal* 59(4), 322-332.
- [9] Laufer, B. (1997). The lexical plight in second language reading: words you don't know, words you think you know and words you can't guess *Second language vocabulary acquisition: A rational for pedagogy* (pp. 20-52).
- [10] McEnery, T., & Hardie, A. (2012). *Corpus linguistics*. Cambridge: Cambridge University Press.
- [11] McPartland, P. (1983). The acquisition of phrasal verbs by non-native speakers of English. *CUNYForum*, 14, 150-156.
- [12] Mukundan, J. (2004). A composite framework for ESL textbook evaluation. Universiti Putra Malaysia, Serdang.
- [13] Mullany, L., & Stockwell, P. (2010). *Introducing English language: a resource book for students*. Oxon: Routledge.
- [14] O'Dowd, E. (1994). *Prepositions and particles in English: A discourse-based, unifying account* University of Colorado.
- [15] Ranali, J. M. (2003). *ELT coursebooks in the age of corpus linguistics: constraints and possibilities*. Birmingham: University of Birmingham.
- [16] Schneider, E. W. (2004). How to trace structural nativization: particle verbs in world Englishes. *World Englishes*, 23(2), 227-249.
- [17] 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, 119-139.
- [18] Thornbury, S. (2002). *How to teach vocabulary*. London: Addison Wesley Publishing Company.

**Abdolvahed Zarifi** was born on July 6, 1966 in Choram, a small town in the southern province of Kohkilooyeh and Boyrahmad, Iran. He started off his matriculation studies in ELT at Martyred Ghasemi TTC in Gachsaran in 1985. He holds a bachelor's degree in English Literature and a master's degree in English language teaching from Shiraz University, Shiraz, Iran. He has recently been awarded a pass with Distinction by the Universiti Putra Malaysia for his PhD Thesis. His research areas of interest include ESL/EFL teaching, story schema, textbook analysis and corpus linguistics. He is particularly interested in studying the English phrasal verb combinations.

**Jayakaran Mukundan**, PhD is Professor of English Language teaching at the Faculty of Educational Studies, Universiti Putra Malaysia. His main research interest area is ELT Materials. He has won Gold Medals for software development at the British Invention Show, UK and IENA, Germany. He is a Director on the Extensive Reading Foundation Board and Visiting Research Fellow at Leeds Metropolitan University, UK.