Pakistan Languages and Humanities Review
www.plhr.org.pk

RESEARCH PAPER

Gaps in ESL Syllabuses: A Comparative Corpus-based Analysis of Language Constructions in Pakistani English Textbooks and British Academic Registers

[1]Tahir Qayyum* and [2]Dr. Shahida Khalique

1. PhD. Scholar, Department of English, University of Azad Jammu and Kashmir, Muzaffarabad, AJ&K, Pakistan
2. Assistant Professor, Department of English, University of Azad Jammu and Kashmir, Muzaffarabad, AJ&K, Pakistan

| ***Corresponding Author** | tq.khan@hotmail.com |

ABSTRACT

The study aims at finding out the differences of the two constructions between Pakistani intermediate English textbooks and British academic registers. Phrasal verb (PV) and Infinitive Object (IO) constructions have often been problematic for second language learners. Pakistani English Textbooks (PETB) corpus is compiled by converting them into soft copies to extract text using Google Cloud Vision API. The data was uploaded to Sketch Engine where British academic corpora were preloaded. The comparison highlights significant register variations between native and non-native use of verb phraseology. The PETB corpus has least constructions of PVs and IOs and most of them represent general senses, rarely found in academic discourse. PETB is closer spoken register than written academic English. Such gaps can be addressed by providing learners comprehensive exposure of language patterns used frequently by native speakers. This study may aid learners internalizing native patterns for effective academic communication.

| **KEYWORDS** | Academic Communication, Constructions, Infinitive Objects, Phrasal Verbs, Register Variation |

Introduction

This study (a part of PhD dissertation) examines the possibility and scope of developing lexical syllabuses for second language learners, based on corpus evidences from the text produced by native speakers of English. The goal of this study is to recognize and scrutinize lexical and phrasal verb patterns from Pakistani English textbooks (PETB) and compare them with British National Corpus (BNC) Academic registers, both written and spoken, represented by the British Academic Written English (BAWE) corpus and the British Academic Spoken English (BASE) corpus. This comparative analysis focuses on verbal constructions to determine how the lexical choices in Pakistani English textbooks align with international academic standards and whether they cater to the needs of learners aiming to acquire academic proficiency in English.

Firth (1957) has rightly commented that words are known by the company they keep. This speaks of constructions and pattern analysis of these phraseological constructions. It is well known fact that words vary in frequency and so do the patterns and for practical pedagogical purposes, measuring and using lexical frequencies need to go much deeper. Thanks to recent work in the analysis of large corpora, it is now possible to do this, either using recently developed tools or the built-in interfaces of British National

Corpus and Corpus of Contemporary American English. Computational linguistic theories and approaches (Corpus linguistics and Construction Grammar, for example) have shown that for each word, some phraseological patterns are more frequent than others. Pattern frequency is not the only criterion for inclusion of a word or a phrase in a lexical syllabus. Other important criteria are idiomaticity, cognitive salience and specificities such as academic needs. In this paper, only two verb phraseological constructions are brought into discussion i.e. phrasal verb and infinitive object constructions with only three lexical verbs which are frequent across three corpora. They are significant from language pedagogical perspectives.

Literature Review

Over the past few decades, there has been an increasing interest of researchers in corpus-based language teaching beside its influence on other areas of linguistic inquiry (Biber & Reppen, 2002; Granger, 2002). The recent works on language teaching particularly in English Language Teaching (ELT) have greatly employed and relied upon corpora-based approaches whether they have been used in dictionary creation (McCarthy, 2008) or in Construction Grammar (Burton, 2012) by challenging conventional approaches (Hunston and Fancis, 1999). Corpus linguistics is a powerful tool for identifying features in learners' language, marked with over-use of certain linguistic features such as phrases, and word frequency (Schmitt;2000, Granger, 2002). Nevertheless, not much attention has been paid to textbooks in the past. The teaching materials have largely been up to the course book writers and syllabus designers' intuitions. Thompson (2000) and Hoey (2005) argue that the major focus of corpus-based approaches has been on academic written material other than textbooks. It has been surprising for Skolverket (2006), and Konstantakis & Alexiou, 2012) due to the vital role they play in ELT. Aligning textbooks with daily life has recently got attention with particular focus on lexical words to the New General Service List proposed by Browne (2013).

Tyler (2012) precisely asserts that vocabulary automization essentially need words to appear frequently in different contexts. It also gets support from Aitchison (2012) that words are more a semantic network rather than isolated entities. The more they appear in different contexts the more they internalize the connections (Nation, 2008). It implies that language learning materials should be frequently exposed to learners in several contexts (Cameron, 2001). The frequency of a word, a phrase or a pattern in the textbook varies, depending upon the size and level of the book however, Cameron (2001) and Nation (1990) believe it ranges from 5-6 (minimum) to (what Waring & Takaki, 2003) suggest twenty occurrences per textbook.

It has always been a daunting task to ascertain vocabulary size and type of words one must know to attain a desired proficiency. There have, however, been studies which indicate the high frequency words help in successful language learning and have a wider text coverage (Nation & Beglar, 2007; Nation, 2006). To note, the AWL (Coxhead, 2000) consists of 2000 most frequently used lexical words without segregating their categories which includes verbs too. There are several verbs in the list that can be used as a noun, at the same time and with the same form. In cases like this, the lexical verbs do not get an isolated treatment. The point here is not to undermine the AWL (Coxhead, 2000) or AGL (West, 1953), rather we aim to address lexical verbs as a separate class that plays a central role in a clause (Biber, et al., 1999). Mayer (1999) also maintains that verbs like 'show' and 'find' are central to scholarly research even though they are prone to everyday language ambiguities. To address this, Paquot (2007) includes important reporting verbs like aim, argue, cause, claim, effect, or suggest in her Academic Keyword List (AKL), that somehow brings LVs into discussion separately unlike Coxhead's (2000) list.

Researchers often use corpus linguistics to assess and evaluate ELT course materials and have been reporting discrepancies between contents included in textbooks and actual use of language. Kettermann (1995) has reported such gaps between instructional rules of grammar and real language use. Nevertheless, Course books play a crucial role in academia, providing far-reaching research and guiding principles across various disciplines (Wells, 1992; Hyland, 2012).

Considering the role of lexical verbs there has been a consensus that they play central roles in Academic English. They are used to express several functions including personal stance, reviewing literature, quoting, and summarizing and help language learners shape ideas and spot on their work within their respective discipline (Granger & Paqout 2010). However, lexical verbs have not been the central point in second language research. Hyland (1999) highlights the less paid attention towards lexical verbs by adding that only reporting verbs could draw researchers' attention because they have been widely utilized in academic writing. The overview of all lexical verbs has been provided by Williams (1996) nonetheless that has remained limited to medical field. Despite verbs are vital to clause structure, lexical specificity has frequently been explored in nouns than in verbs (Cruse, 1977; Rivero, 1977; Halliday & Matthiessen, 1999; Langacker, 2004). Tomasello and Merriman (1995) have rightly pointed out the lack of researchers' interest in lexical verbs despite their natural and necessary occurrence in any clause structure. They advocate for more attention to verbs, leading to a rise in publications on this topic (Tomasello & Merriman, 1995; Altenberg & Granger, 2001; Lee, 2003; Newman & Rice, 2004; Kipper et al., 2007; Ramchand, 2008).

All the same, a little effort has been carried out in Pakistani context. Hussain and Zehra (2020) while discussing lexical bundles in Pakistani textbooks argue that very little research revolves around exploring the nature and frequency of lexical bundles in academic books. In Pakistan, English majority of textbooks include literature for language learning. The presentation of vocabulary as course material needs careful consideration. Literature-based course books may and often contain obsolete or very infrequent vocabulary, which Lipinski (2010) believes, should be avoided. Nation (2006) and Nation and Beglar (2007) believe that higher frequency words improve text comprehension. Nation's (2006) claims that a substantial portion of texts and speech could be covered by the most frequent words, supporting the idea that learning these words can substantially improve language skills.

Material and Methods

Pakistani English Textbook (PETB) corpus is compiled by scanning two books of Federal Board of Intermediate and Secondary Education (Book I and Book II). These books were transformed to machine readable form by scanning them with digital scanner to change them into JGP format by using Google Cloud Vision API. These files were subsequently uploaded to Sketch Engine which saves it as a corpus and applies built-in tools to identify the required features of the given text. Target and reference corpora can, therefore, be easily compared and analyzed. The two significant levels of analyses have been employed. Firstly, it considers the top twelve lexical verbs extracted from PETB and are compared with same number of LVs from British Academic registers. They have been compared to know the frequently used LVs in PETB, BAWE and BASE. Secondly, out of those twelve LVs only those have been extracted which form patterns either with particles to get into Phrasal Verb (PVs) construction.

Reference corpora (BNC academic) have been extracted from preloaded corpora of Sketch Engine and preferred over alternatives like the Corpus of Contemporary American English (COCA) or COCA-AC and others due to its predominant advantages for this particular research as emphasized by Štěrba (2018). The BNC offers all-inclusive representation of British English, aligning well with the Pakistani textbooks chosen as a case study. Moreover, the classification system provided by Lee (2001) allows for nuanced distinctions among various spoken and written registers. Štěrba (2018) has rightly pointed out that Liu (2011) addresses register comparison using COCA, which categorizes texts into only five genres, the BNC's classification, on the other hand incorporates seventy diverse register categories. And finally, the analysis has utilized the BNC as it is freely available for analysis. Moreover, since the study accounts for frequency-based verbs, phrasal verbs and their constructions, the Sketch Engine interface already has it incorporated along with significant association measures. Sketch Engine also incorporates corrections for text type and parts of speech tags, as well as the implementation of Lee's Classification (O'Keeffe et al., 2007).

Sketch Engine has been used as an analytical tool which offers a very user-friendly interface with built-in programs for variety of interpretations. A corpus of Pakistani English Textbooks was created as PETB by uploading 47 labelled files from A01 to A22 and B01 to B25 of the book 1 and book 2 respectively. Once the corpus has been created, it appears on the dashboard for either to compare it with any reference corpus (available on Sketch Engine for free) or do perform separate operations. After the lexical lemmas of verbs have been extracted, they have carefully been looked upon with the combinations they appeared. After the completion of the first step of taking out the most frequently used lexical verbs, verb patterns were observed for each lexical lemma of verb. Only three have been discussed in this paper. For phrasal verb constructions the current study, mainly has relied on Gardener and Davies (2007) identification of particles and definition of phrasal verbs. However, for infinitive object constructions, our observation of their specificity has been relied upon. This pattern accounts for register variations as well as marks native not-native distinction.

Results and Discussion

The compiled corpus consists of 79499 tokens. For 1506 total LV lemmas, 9389 hits were found with minimum frequency of 1 to the maximum frequency of 294. These LV lemmas are categorized according to their absolute frequencies aligning with the relative frequency range.

Table 1
Comparative frequencies of LVs in PETB, BASE and BAWE

method name: poswordlist			method name: poswordlist			method name: poswordlist		
corpus: user/tq.khan/PETB			corpus: preloaded/base2			corpus: preloaded/bawe2		
LV	Frequency	Rltv. Freq.	LV	Frequency	Rltv. Freq.	LV	Frequency	Rltv. Freq.
say	294	3234	go	8590	4890	use	17983	2157
go	207	2277	get	8513	4846	make	12100	1451
make	204	2244	know	6268	3568	show	9621	1154
come	161	1771	say	6211	3536	see	9531	1143
see	159	1749	think	5997	3414	take	8554	1026
take	157	1727	see	3610	2055	give	7535	904
know	145	1595	look	3369	1918	find	7426	891
give	144	1584	make	2774	1579	provide	6355	762
use	134	1474	mean	2770	1577	become	5598	672

tell	120	1320	want	2770	1577	need	5171	620
think	116	1276	come	2663	1516	must	4880	585
get	108	1188	use	2503	1425	suggest	4844	581

The data reveals that PETB includes nine out of twelve LVs included BASE list. It implies that 75% of the LVS are exactly the same though they are ranked differently on the basis of their frequencies. When compared with BAWE it has 50% of the same LVs with striking differences in relative frequencies. It is also important to note the most frequent LVs in PETB include communicative verbs (say, tell, see), action verbs (make, take, give, use), motion or dynamic verbs (go, come) and cognitive verbs (think, know, get). The relative frequencies are varied though. "Say", for instance has the highest hits in the compiled corpus of PETB with 294 hits. The BASE corpus differs in three LVS ("look", "mean", and "want"), with PETB and BAWE. When compared to PETB, it does not include "take", "give", and "tell". It is interesting to note that "say" in BAWE does not appear in the list even in 30 most frequent lexical verbs. BAWE shows relative frequency of "say" less than four hundred hits which is eight time lower than that of PETB. Similarly, "go", shares the closest rank of its appearance in BASE whose frequency is more than double of PETB "go" with relative frequency of 4890 and remains on the top of the table. When compared to BAWE, interestingly, this entry misses out too. In BAWE, "go" is not even included in 35 most frequently used lexical verbs. The lexical verb "make" has almost similar and balanced distribution with a little difference of higher frequency in PETB (2244) compared with the BAWE and the BASE. The LV, "See" has an analogous representation with slight difference of relative frequencies when compared to BASE and a higher one when compared to BAWE. However, the LV, "see" is among the top six most frequently used LV lemmas across the corpora. "Know" is also one of the most frequent lexical verbs used in PETB. When checked with its relative frequency, it has 457 hits per million ways less than PETB and nine times lower than that of BASE. "GIOe" is only lexical verb missing from the BASE top twelve lexical verbs but present in PETB and BAWE with former having higher relative frequency. The LV "Use" also shares distribution and frequency with BASE with a slightly lower rank of BASE. Contrarily, "Use" is the most frequently used lexical verb in BAWE. The LVs "think" and "get" also share characteristics in terms of distribution in PETB and BASE but are not found in the top ranked twelve most frequently used lexical verbs in BAWE. The frequencies have striking differences though. The lexical verb, "think" has almost thrice higher frequency in BASE when compared to PETB. The lexical verb "tell" in the PETB with relative frequency of 1320 is the only one neither in top ranked LVs of BASE nor in the BAWE.

Figure 1   Display of Frequencies across Corpora
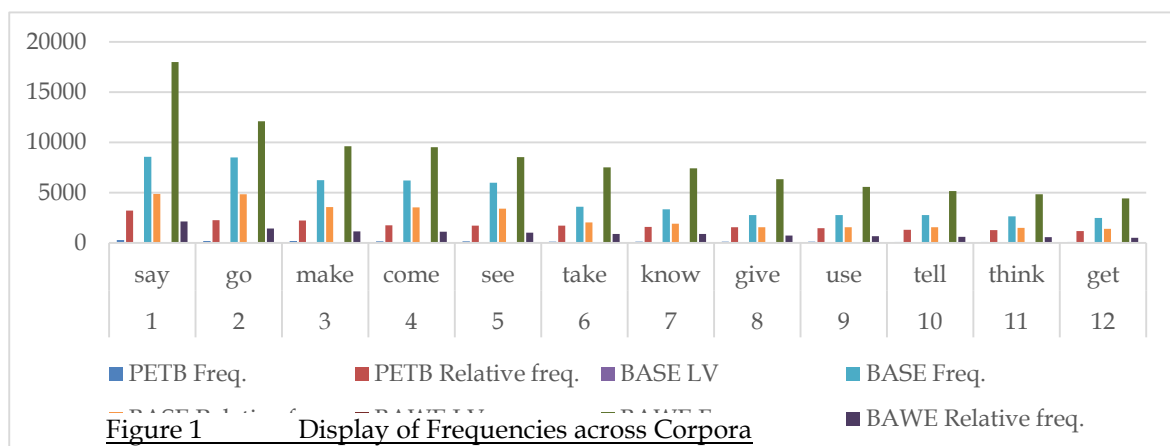
It shows that there are significant similarities between PETB and BASE with striking differences in PETB and BAWE as well as BASE and BAWE for lexical verbs' presence, distribution and frequencies. PETB includes well balanced lexical entries of

communicative and action verbs with high frequencies, it has some notable differences with BAWE. "Say", for example has the highest frequency in PETB and BASE but significantly less frequent in BAWE. Similarly, "go" is infrequent in BAWE when compared with BASE. "Make" shows a balanced distribution in all three corpora with slight variations of frequencies. "See" is consistently among the top LVs in all corpora. The LV "Know" has higher frequency in PETB and BASE but completely missing in the top 12 LVs in BAWE. Also, BASE includes verbs like "look," "mean," and "want," which are absent in PETB and BAWE. The presence of LV "Tell" in PETB among the most frequent twelve LVs and absence in BASE and BAWE shows its probable redundant use. Some significant dissimilarities in relative frequencies are also noted, such as "think" having a much higher frequency in BASE compared to PETB. "Use" is highly frequent in BAWE but not as much in PETB and BASE. Both PETB and BASE have lexical verbs "say," "go," "make," "see," and "know" among their most frequently used LVs. Also, "Make" and "see" exhibit analogous distribution patterns across all three corpora.

There are certain lexical verbs such as "look," "mean," and "want" which are exclusively frequent in BASE unlike PETB and BAWE. The LV "Tell" exclusively appears among the most frequent LVs in PETB but has a much lower frequency in BASE and BAWE. Frequency of the LV "say" is much lower in BAWE compared to PETB and BASE whereas "Go" is notably frequent in BASE but almost absent among top ranked LVs in BAWE. "Think" has a significantly higher frequency in BASE compared to PETB which includes it in top ranking however in BAWE it is not much frequent.

These findings indicate that while there are some common high-frequency verbs shared across the corpora, there are also notable differences in the presence and frequency of specific verbs. It is evident from these top ranked lexical verbs that Pakistani English Textbooks are closer to British Academic Spoken English as both share the most frequent verbs and their relative frequencies though there are some differences in terms of their relative frequencies. It, implies in other words that Pakistani English textbooks of intermediate have less in common with BAWE.

The second part of the analysis accounts for Phrasal Verb (PV) and Infinitive Object (IO) constructions. The inclusion of PVs in academic texts shows how the learners better equip themselves with Norms and Exploitations (as suggested by Hanks, 2013) of the constructions.

Table 2
Comparison of Phrasal verb construction with LV "Come"

corpus: preloaded/base2				corpus: preloaded/bawe2				corpus: user/tq.khan/PETB			
Keyword	Collocate	Freq	Score	Keyword	Collocate	Freq	Score	Keyword	Collocate	Freq	Score
come	particles after X	505	19	come	part_intrans	309	10.1	come	particles after X	9	5.59
	out	183	11.66		about	40	11.43		up	4	10.84
	up	177	11.58		up	139	10.84		down	2	10.64
	along	36	10.95		back	42	10.38		off	1	10.14
	across	26	10.57		along	6	9.15		out	2	9.66
	down	38	10.49		down	18	9.06	come	infinitive objects of X	12	7.45
	in	19	10.07		through	5	8.94		tell	2	11.67
	through	6	8.4		over	7	8.76	come	dust	1	11.3
	on	7	8.3		round	4	8.59		scoff	1	11.3
	off	8	8.3		off	7	8.34		spot	1	11.3
	over	4	7.71		out	34	8.32		cash	1	11.3

come	infinitive objects of X	78	2.93	come	part_trans	15	0.49		greet	1	11.3
	mind	4	10.6		out	9	7.48		mean	1	11.3
	see	6	9.15						realize	1	11.19
	look	5	8.65						expect	1	11.19
	be	9	6.73						speak	1	10.91
									take	1	10.3

These above structures give comprehensive insight and helps identifying gaps where ESL learners fall short of the native like proficiency. Comparison LV "come" in PV and IO Construction

The table shows that in PETB, the LV only combines with four particles nine times with only "up" having four hits. BASE has the highest total frequency of 505 followed by BAWE (309). The log Dice Score in all three corpora shows the typical association of LV with the respective particle. Although, "down, off and out", in PETB, have strong log dice score, the frequency is significantly low. If we look at the total PV constructions in BAWE and BASE, there are ten constructions equally found in both corpora, more than twice of PETB. There are some differences between BASE and BAWE too for instance, the presence of "across" and "round" for which the latter is present in BAWE with comparatively low log dice. The PV "come across" has a strong association of collocates despite the fact that it, sometimes, is not used as a phrasal verb. Similar conclusion can be drawn for particles "in" and "back".

Similarly, the IO construction is significant too. The results show that BAWE rarely has such constructions. However, in BASE, we find 78 such instances with low log dice score of 2.93. It implies that the association is quite random. There are 12 such instances in PETB with stable log dice of 7.45 showing stronger association of appearance between them. However, in this construction, all 12 instances have been found once in the text except one "come to tell" with two hits. If we look at BAWE, with no such instances, it is even more important to examine what necessitates it. The apparent justification seems to be its formal nature. Academic written English has a strict and accurate sentence structure. The very reason seems a plausible justification.

Table 3.
Comparison of LV "get" in PV and IO Construction

corpus: preloaded/base2				corpus: preloaded/bawe2				corpus: user/tq.khan/PETB			
Keyword	Collocate	Freq	Score	Keyword	Collocate	Freq	Score	Keyword	Collocate	Freq	Score
Go	particles at X	562	6.54	go	part_intran	482	16.6	go	particles at X	35	16.9
	on	120	12.29		on	227	12.92		on	6	12.06
	down	65	11.17		back	90	11.26		over	6	11.83
	up	110	10.85		through	15	9.92		away	4	11.5
	off	45	10.7		about	16	9.66		in	3	11.3
	along	33	10.7		down	30	9.59		out	7	11.06
	out	94	10.65		up	42	9.01		off	2	10.21
	through	30	10.59		out	42	8.55		down	2	9.93
	over	19	9.83		off	7	8.02		through	1	9.79
	away	16	9.73		along	4	7.97		along	1	9.71
	around	15	9.5	go	part_trans	31	1.07		around	1	9.64
	in	10	9.01		down	11	9.38		up	2	9.39
	across	5	8.05		back	4	8.75	go	infinitive objects of X	29	14

Go	infinitive objects of X	3720	43.3	up	7	7.14	sleep	3	11.58
	be	921	12.19				jail	2	11.04
	do	257	10.72				take	2	10.57
	have	201	10.54				snow	1	10.09
	talk	140	10.17				fetch	1	10.09
	get	133	9.97				school	1	10.09
	go	127	9.97				complain	1	10.04
	say	90	9.44				try	1	10.04
	look	83	9.39				deal	1	10
	give	60	8.96				maintain	1	10
	use	56	8.85				bear	1	10
	take	49	8.65						
	come	46	8.6						
	happen	45	8.6						

The LV "get" is one of the most frequently used lexical verb across corpora and peculiar in academic discourse with variety of functions. If we look at Phrasal Verb constructions, PETB has eight such instances with 7.41 of total log dice way too high compared to both BAWE and BASE. There are only five constructions in PETB with "up" thrice "out" twice. Phrasal verbs like, "get along" and "get around" have strongest association with log dice of more than eleven still they are not frequent. It is also interesting that PVs "get up" and "get out" are used with no semantic variations. The BASE has 124 total appearances of "get" with particles forming phrasal verbs. However, it has eight different phrasal verb constructions with varying semantic nuances.

In the BASE, LV "get" for instance, has 124 hits with several particles, signifying its frequent use in construction of PVs. The BASE shows, however, a high frequency of specific particles such as "across" (11 times, with log score of 10.99) and "out" (45 times, with log score 9.97), proposing the everyday usage of such PV constructions. In contrast, in BAWE "get" appears with several particles a relatively higher log dice score excluding "across" and "away", implying their absence in written academic English is register specific. Particularly, the particle "up" marks a reasonable frequency. The phrasal verbs having particles "along" and "around" have exceptionally high scores with drastically low frequencies in PETB. The rest of the PV constructions are low in frequencies too. At the same time, absence of IO construction from BAWE demonstrates that formal written text is direct, simple and relies on finite verbs instead of making structures complex with infinite verbs however, the pervasive use of IO construction in BASE is evidence of its presence in academic spheres. Looking at the IO constructions themselves, the stark difference between PETB and BASE can be observed. The PETB on one hand does not correspond with BAWE in a sense that later is devoid of IO constructions completely and with BASE in a sense that IO constructions are way less in terms of frequency and variations. There are, for instance, only four IO constructions in the PETB compared to 600 in the BASE. "Get to know" and "got to be" are quite common in everyday life and contextually dependent. The comparison of two constructions in the table below, further elucidates the significance of the already discussed patterns.

Table 4
PV and IO constructions with verb "Go" across Corpora

corpus: preloaded/base2				corpus: preloaded/bawe2				corpus: user/tq.khan/PETB			
Keyword	Collocate	Freq	Score	Keyword	Collocate	Freq	Score	Keyword	Collocate	Freq	Score
get	particles after X	124	1.46	get	part_intrans	65	2.63	get	particles after X	8	7.41

	Across	11	10.99		around	6	10.25		along	1	11.41
	Away	8	10.53		back	21	9.78		around	1	11.19
	Through	7	10.16		up	15	7.78		up	3	10.45
	Out	45	9.97		down	4	7.26		out	2	9.67
	Down	10	9.39		out	10	6.66		down	1	9.67
	On	6	9.15	get	part_trans	35	1.41	get	infinitive objects X	4	3.7
	Up	22	8.9		back	13	10.43		bed	1	12.41
	Off	6	8.77		around	6	10.37		deal	1	12.19
get	infinitive objects X	600	7.05		out	7	7.09		get	1	11
	Be	111	10.1		up	4	6.33		do	1	10
	Go	30	9.89								
	Have	36	9.81								
	Do	40	9.46								
	Know	17	9.3								
	Think	15	9.18								
	Get	20	9.08								
	Make	15	9.04								
	Put	10	8.73								
	Prove	8	8.67								
	Try	9	8.62								
	Say	13	8.57								
	Start	8	8.53								
	Give	9	8.51								
	Look	10	8.46								
	Keep	7	8.39								
	Calculate	6	8.28								
	Take	8	8.27								
	Decide	6	8.26								
	Come	7	8.24								
	See	7	8.04								
	Learn	5	8.01								

In the BASE corpus, the lexical verb "go" collocates with particles for a total of 562 times, much higher than PETB. The most frequent particle "on" (120 times, score 12.3) is followed by "down" (65 times, score 11.2), "up" (110 times, score 10.9), and "out" (104 times, score 10.6). The BAWE shows a comparatively low total 482 phrasal verb constructions with intransitive and 31 with transitive usage patterns totaling 513 total occurrences as phrasal verbs. Relying on Wordnet for word senses, "go on" has been reported with five senses and all of them have been observed extensively by the native speakers in academic settings. In PETB, the phrasal verb "go on" has been used six times. The stronger association makes it a case to cross check the way it has been used through the corpus. However, the concordance shows that "go on" in PETB has only been used either referring to continuation (four times) or a happening (twice). The PV "go back" reveals interesting results. It is neither found in BASE nor in PETB though it has extensively been found in the BAWE. It has 94 total hits with 90 hits as an intransitive particle and four as transitive particle with stronger association of log dice around 12. Again, it authenticates our earlier claim that the PETB is closer to BASE as compared to BAWE. The phrasal verb "go back" has a variety of interesting semantic interpretations. Wordnet gives us 4 senses, and all are present in BAWE. The higher log dice scores advocate a more formal and emphatic use in academic text. These findings highlight the importance of phrasal verb constructions.

Phrasal verb "go over" is yet another example, where the PETB aligns with BASE due to its presence in both corpora which is not listed in the BAWE. In PETB, the lexical verb "go" collocates with particles 35 times to get into phrasal verb construction. Despite some higher collocational scores, the phrasal verbs are in significantly low frequency giving strength to the assumption that the limited use of high probability collocates does not offer sufficient exposure of phrasal verb constructions which are very frequent in other academic corpora. It also implies that this is not a deliberate choice for PETB writers or course designers. Textbooks cover or are supposed to cover a wide range of language content within the provided space, and the lower frequency might simply be a result of such inherent limitations rather than a precise academic choice. It may also refer to a point that most textbooks include literary content which primarily does not aim at assisting language pedagogical purposes. It is even more pertinent to mention that the literary content is not produced for language instruction at this level. The already available literary text, even one produced by native speakers, may not have all the essentials to be used for language instructions. It does not necessarily capture the complex and varied nature of real-world language.

Conclusion

This research has explored the usage patterns of phrasal verbs and Infinitive Object constructions across three corpora by identifying the twelve most frequent Lexical Verbs. The variations in the frequency and senses of these constructions highlight academic implications and the malleability of language use. BASE reflects the dynamic nature of spoken English. BAWE underlines certain PVs more strongly, signifying formal and categorical practices in native academic writing and the same has been advocated by Gardener and Davies (2007). The PETB corpus exposes a limited use of PVs whereas IOs are more frequent. The limited frequency of PV constructions in textbooks reveals that ESL learners in Pakistan may have lesser exposure to a comprehensive range of phrasal verbs. This could potentially restrict learners' ability to use these patterns effectively in academic communication. PETB is closely aligned with the BASE than BAWE in terms of LV frequencies, PV and IO constructions.

The data from the three corpora provide valuable insights into the adaptability and contextual variations of PVs and IOs in English. The flexible nature of the BASE sharply contrasts with the more organized and formal usage in the BAWE. The limited use in pedagogical contexts (PETB) highlights the challenges of providing comprehensive exposure to learners within the constraints of textbook content. The formality and precision of the BAWE shows conventional academic writing, where the choice of constructions is guided by the need for explicit and accurate expression. It also shows that there is a potential gap in natural language exposure for learners. Language instruction should aim at giving learners comprehensive exposure to language patterns depending on frequency and usage in both spoken and written contexts as emphasized by Biber et al. (1998), Hunston (2002), Hoey (2005), Nation (2006) and Nation and Beglar (2007), Gardner and Davies (2007), Liu (2011) and Garnier and Schmitt (2014). Textbooks should incorporate patterns extracted from both spoken and written corpora to improve learners' adaptability. Language instruction must focus on high-frequency PVs and IOs, emphasizing their strong associations and natural occurrence in language. Future research can also find out the impact of variations on the usage of PVs and IOs across different genres and registers. By addressing the ruptures in instructional resources and concentrating on high-frequency patterns, language teaching can benefit learners for effective communication in both speaking and writing. Future research can further explore the cognitive and contextual aspects of PVs and IOs use, contributing to a better understanding of language pedagogy.

Recommendations?????

References

Aitchison, J. (2012). *Words in the mind: An Introduction to the Mental Lexicon* (4th ed.). Wiley-Blackwell.

Altenberg, B., & Granger, S. (2001). The Grammatical and Lexical Patterning of Make in Native and Non-Native Student Writing. *Applied Linguistics*, 22(2), 173-194.

Biber, D., Conrad, S., & Reppen, R. (1999). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.

Biber, D., & Reppen, R. (2002). "What Does Frequency Have to Do with Grammar Teaching?" Studies in Second Language Acquisition, 24(2), 199-208. doi:10.1017/S0272263102002048

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

Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, 24(2), 199-208.

Browne, C. (2013). The new general service list: Celebrating 60 Years of Vocabulary Learning. *The Language Teacher*, 37(4), 13-16.

Burton, S. (2012). *Corpus-based Approaches to Language Teaching*. Routledge.

Cameron, L. (2001). *Teaching Languages to Young Learners*. Cambridge University Press.

Cobb, T. (n.d.). *Compleat Lexical Tutor [website]*. Montréal: Université du Québec à Montréal.

Cornell, A. (1985). Realistic Goals in Teaching and Learning Phrasal Verbs. *International Review of Applied Linguistics in Language Teaching,* 23(4), 269-280.

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213-238.

Cruse, D. A. (1977). The Pragmatics of Lexical Specificity. *Journal of Linguistics*, 13(2), 153-164.

Gardner, D., & Davies, M. (2007). Pointing out Frequent Phrasal Verbs: A Corpus-Based Analysis. *TESOL Quarterly*, 41(2), 339-359.

Granger, S. (2002). A Bird's-eye View of Learner Corpus Research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching* (pp. 3-33). John Benjamins.

Granger, S., & Paquot, M. (2010). *Lexical Verbs in Academic Discourse: A Corpus-Driven Study of Learner and Expert Writing*. Continuum.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999). *Construing Experience Through Meaning: A Language-Based Approach to Cognition*. Continuum.

Hoey, M. (2005). *Lexical Priming: A New Theory f Words And Language*. Routledge.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hunston, S., & Francis, G. (1999). *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. John Benjamins.

Hyland, K. (1999). *Disciplinary Discourses: Social Interactions in Academic Writing*. Longman.

Hyland, K. (2012). *Disciplinary Identities: Individuality and Community in Academic Discourse*. Cambridge University Press.

Kettermann, B. (1995). *Lexical Collocations and their Impact on Language Learning*. Peter Lang.

Kipper, K., et al. (2007). A Large-Scale Classification of English Verbs. *Language Resources and Evaluation*, 41(1), 1-29.

Konstantakis, N., & Alexiou, T. (2012). *Vocabulary Learning and Teaching: Insights from Research*. Cambridge Scholars Publishing.

Langacker, R. W. (2004). *Foundations of Cognitive Grammar: Volume II: Descriptive Application*. Stanford University Press.

Lee, D. (2003). *Cognitive Linguistics: An Introduction*. Oxford University Press.

Lipinski, J. (2010). *Frequency-based Analysis of Language*. Springer.

Liu, D. (2011). The Most Frequently Used English Phrasal Verbs in American and British English: A Multicorpus Examination. *TESOL Quarterly*, 45(4), 661-688.

Mayer, R. E. (1999). *Multimedia Learning*. Cambridge University Press.

McCarthy, M. (2008). *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge University Press.

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Heinle & Heinle.

Nation, I. S. P. (2006). How Large a Vocabulary is Needed for Reading and Listening? *The Canadian Modern Language Review*, 63(1), 59-82.

Nation, I. S. P. (2008). *Teaching Vocabulary: Strategies and Techniques*. Heinle Cengage Learning.

Nation, I. S. P., & Beglar, D. (2007). A Vocabulary Size Test. *The Language Teacher*, 31(7), 9-13.

Newman, J., & Rice, S. (2004). *Patterns in Language and Linguistics*. Cambridge University Press.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). From Corpus to Classroom: Language Use and Language Teaching. Cambridge University Press.

Paquot, M. (2007). *Academic Vocabulary in Learner Writing: from Extraction to Analysis*. Continuum.

Ramchand, G. (2008). *Verb Meaning and the Lexicon: A First-phase Syntax*. Cambridge University Press.

Rivero, M. L. (1977). *On the syntax and semantics of verb-particle constructions*. Indiana University Linguistics Club.

Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge University Press.

Schmitt, N. (Ed.). (2004). *Formulaic Sequences: Acquisition, Processing, and Use*. Amsterdam: John Benjamins.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Basingstoke: Palgrave Macmillan.

Siyanova, A., & Schmitt, N. (2007). Native and Nonnative Use of Multi-word vs. One-word Verbs. *International Review of Applied Linguistics*, 45(1), 109-139.

Skolverket. (2006). *Curriculum for the Compulsory School, Preschool Class and the Leisure-time Centre*. Skolverket.

Štěrba, J. (2018). Phrasal Verbs in the British National Corpus and ELT Textbooks. Masaryk University.

Thompson, G. (2000). *Introducing Functional Grammar*. Arnold.

Tomasello, M., & Merriman, W. E. (1995). *Beyond Names for Things: Young Children's Acquisition of Verbs*. Lawrence Erlbaum Associates.

Tyler, A. (2012). *Cognitive Linguistics and Second Language Learning: Theoretical Basics and Experimental Evidence*. Routledge.

Waring, R., & Takaki, M. (2003). At What Rate do Learners Learn and Retain New Vocabulary from Reading a Graded Reader? *Reading in a Foreign Language*, 15(2), 130-163.

Wells, G. (1992). The Centrality of Talk in Education. In K. Norman (Ed.), *Thinking voices: The Work of the National Oracy Project* (pp. 283-310). Hodder & Stoughton.

West, M. (1953). *A General Service List of English Words*. Longman.

Williams, J. (1996). *Focus on Form in Classroom Second Language Acquisition*. Cambridge University Press.