

HANDLING ENGLISH LEXICON IN EFL CLASS: A CORPUS-BASED PERSPECTIVE

¹Abrar Hussain Qureshi ²Behzad Anwar ³Athar Munir

¹University of Lahore Pakpattan Campus, Pakpattan.

²Department of English, University of Gujrat, Gujrat.

³Department of English, Govt. Emerson College Multan

Corresponding Author: abrarqureshi74f@gmail.com

ABSTRACT: *Strategies for vocabulary development fall into two main categories: learning techniques for vocabulary directly and learning new words indirectly. Students learn new vocabulary items, and nearly about 10 percent of these words come from direct vocabulary instruction. The rest come from their everyday language experiences and reading a wide variety of texts in an autonomous environment. In order to make classroom vocabulary learning strategies more interesting and motivating, there is a great need of innovations in the learning techniques of vocabulary in an EFL classroom. The paper reinforces the fact how corpus-based learning of English vocabulary in an EFL class room in Pakistan can improve the condition. As researchers conclude that classroom learning of English vocabulary can have the biggest impact on the overall language proficiency of the non-native learners.*

Key Words: Vocabulary, Corpus, EFL Learners, Vocabulary learning strategies in Pakistan

1. INTRODUCTION

According to Carter McCarthy, “Lexicon building recently has been identified to be the most distinct event in general language proficiency.” [1]. Laufer & Melamed states: “It is also identified by most learners to be one of the significant challenges of language learning.” [2]. Vocabulary items play a vital role in foreign language learning. Katamba claims that “words are the building blocks of a language since they label objects, actions, ideas without which people cannot convey the intended meaning.” [3]. Bauer claims that “little can be conveyed without grammar but nothing can be conveyed without vocabulary.” [4]. Sinclair (1991) is of the opinion that corpus is one of the modern and most enlightening among various strategies that contribute to lexicon building, for a corpus makes an easy and quick analysis of great amounts of linguistic data possible. The vocabulary teaching method based on corpus helps to develop students to learn the foreign language independently. It builds their confidence even in the use of those vocabulary items.

The use of corpus in an EFL classroom is a new phenomenon in applied linguistics in Pakistan. Based on a large collection of authentic text and potential soft wares, corpus helps in analyzing accurately the greatest amounts of linguistic data in a short time, and the learners are provided with a new approach to learn a language independently.

2. BACKGROUND OF THE STUDY

Language teaching practice is becoming more learner-centered in the field of language teaching, learner autonomy has been an ongoing issue of foreign language educators in Pakistan. The development of individualized study methods and the autonomous learning ability on part of the students is an important indicator of the successful reform of the teaching model in the past decade in Pakistan. As an assistant tool in language learning, corpus makes an easy and quick analysis of the greatest amounts of linguistic data possible, and the learners are provided with a new approach to learn a language independently. Sinclair States: “As one of the major components of language system, vocabulary is especially

essential for English as Foreign Language. Pakistani college students generally believe that the learning of large amount of vocabulary is one of the most challenging as well as necessary tasks. Among various approaches that contribute to vocabulary acquisition, corpus is one of the latest and most enlightening, for a corpus makes an easy and quick analysis of great amounts of linguistic data possible.” [5].

3. RESEARCH METHODOLOGY

With the dawn of science and technology, a new approach towards language learning and teaching has emerged. There have been many attempts to integrate computers as tools in language classrooms to facilitate the learning. The application of the computer corpus in the EFL teaching and learning has marked a fundamental shift both in methodology and in ideology with regard to linguistic studies and language learning.

A latest Pakistani English corpus was used to accomplish this research. Data taken from this corpus helped in suggesting the list of highly frequent words. Their underlying meanings were determined with the help of concordance and collocations. Though Pakistani English Corpus is not a big collection of English text as The Bank of English, compiled by University of Birmingham United Kingdom, however, it was used as a point of departure hoping that a large corpus of Pakistani English would be compiled for linguistics purposes. In order to process the corpus, Wordsmith Tools, software prepared by Mike Scott and published by university of Oxford were used. It is a comprehensible software package mainly used in linguists specifically in modern lexicography. This software helped in processing the corpus data of Pakistani English.

4. DATA ANALYSIS

In order to enhance the lexical proficiency of the non-native learners of the English language and provide a guide for the future study, a word list of the most frequent words from Pakistani English was retrieved with the help of Wordsmith Tools

Table:1 Most Frequent Words in the Corpus

N	Word	Freq. %	Texts%	Lemmas	Set
1	THE	336,865	6.51	934	99.89
2	#	154,696	2.99	920	98.40
3	OF	151,358	2.93	933	99.79
4	TO	131,571	2.54	904	96.68
5	AND	125,844	2.43	907	97.01
6	IN	108,707	2.10	934	99.89
7	A	94,484	1.83	933	99.79
8	THAT	50,065	0.97	901	96.36
9	ON	48,765	0.94	933	99.79
10	FOR	45,543	0.88	933	99.79
11	SAID	39,931	0.77	832	88.98
12	IS	38,027	0.74	899	96.15
13	WAS	34,226	0.66	899	96.15
14	BY	32,258	0.62	904	96.68
15	HE	30,654	0.59	888	94.97
16	WITH	28,186	0.54	900	96.26
17	AS	27,257	0.53	900	96.26
18	IT	27,046	0.52	900	96.26
19	AT	26,318	0.51	903	96.58
20	BE	24,369	0.47	897	95.94
21	FROM	23,113	0.45	912	97.54
22	NOT	20,627	0.40	925	98.93
23	HAS	19,667	0.38	898	96.04
24	HAD	19,073	0.37	879	94.01
25	HAVE	18,734	0.36	883	94.44
26	ARE	18,719	0.36	918	98.18
27	HIS	16,884	0.33	875	93.58
28	WERE	16,619	0.32	881	94.22
29	AN	16,608	0.32	893	95.51
30	THEIR	16,261	0.31	891	95.29
31	THIS	15,928	0.31	890	95.19
32	BUT	15,757	0.30	891	95.29
33	THEY	15,649	0.30	887	94.87
34	WOULD	15,064	0.29	873	93.37
35	WILL	14,123	0.27	860	91.98
36	PAKISTAN	13,491	0.26	796	85.13
37	WHICH	12,760	0.25	879	94.01
38	BEEN	12,580	0.24	876	93.69
39	ALSO	12,267	0.24	871	93.16
40	GOVERNMENT	11,679	0.23	779	83.32
41	ITS	11,344	0.22	856	91.55
42	WHO	10,872	0.21	869	92.94
43	I	10,833	0.21	766	81.93
44	ALL	10,245	0.20	862	92.19
45	AFTER	10,166	0.20	866	92.62
46	ONE	9,759	0.19	865	92.51
47	OR	9,412	0.18	848	90.70
48	US	9,218	0.18	751	80.32
49	AGAINST	9,188	0.18	828	88.56
50	WE	9,047	0.17	842	90.05
51	TWO	8,700	0.17	846	90.48
52	PEOPLE	8,694	0.17	835	89.30
53	UP	8,352	0.16	856	91.55
54	OVER	7,787	0.15	842	90.05
55	WHEN	7,772	0.15	853	91.23
56	OUT	7,714	0.15	855	91.44
57	THERE	7,530	0.15	843	90.16
58	NO	7,525	0.15	863	92.30
59	POLICE	7,515	0.15	574	61.39
60	MORE	7,408	0.14	833	89.09
61	ABOUT	7,252	0.14	847	90.59

July-August

62	IF	7,160	0.14	835	89.30
63	OTHER	7,067	0.14	845	90.37
64	PRESIDENT	7,017	0.14	754	80.64
65	YEAR	6,991	0.14	774	82.78
66	NEW	6,762	0.13	793	84.81
67	COURT	6,705	0.13	647	69.20
68	THEM	6,450	0.12	835	89.30
69	SOME	6,101	0.12	828	88.56
70	WHILE	5,814	0.11	826	88.34
71	PER	5,808	0.11	528	56.47
72	TIME	5,743	0.11	824	88.13
73	COULD	5,726	0.11	827	88.45
74	OUR	5,709	0.11	739	79.04
75	SEPT	5,592	0.11	437	46.74
76	KARACHI	5,502	0.11	578	61.82
77	SO	5,499	0.11	823	88.02
78	ONLY	5,481	0.11	829	88.66
79	COUNTRY	5,453	0.11	760	81.28
80	ANY	5,432	0.11	808	86.42
81	SHOULD	5,412	0.10	783	83.74
82	FIRST	5,408	0.10	767	82.03
83	CAN	5,366	0.10	769	82.25
84	HER	5,175	0.10	619	66.20
85	POLITICAL	5,144	0.10	657	70.27
86	INTO	5,138	0.10	804	85.99
87	LAST	5,126	0.10	809	86.52
88	WORLD	5,112	0.10	689	73.69
89	THAN	5,106	0.10	802	85.78
90	MINISTER	5,105	0.10	703	75.19
91	GENERAL	5,040	0.10	728	77.86
92	PERCENT	4,962	0.10	322	34.44
93	SHE	4,962	0.10	606	64.81
94	THREE	4,871	0.09	751	80.32
95	KHAN	4,852	0.09	640	68.45
96	UNDER	4,826	0.09	789	84.39
97	MUSHARRAF	4,797	0.09	474	50.70
98	BEING	4,662	0.09	822	87.91
99	RS	4,637	0.09	305	32.62
100	HIM	4,620	0.09	751	80.32
101	THESE	4,596	0.09	769	82.25
102	DAY	4,582	0.09	787	84.17
103	MILLION	4,582	0.09	617	65.99
104	TOLD	4,567	0.09	712	76.15
105	CHIEF	4,502	0.09	726	77.65
106	YEARS	4,497	0.09	801	85.67
107	DURING	4,467	0.09	773	82.67
108	YOU	4,432	0.09	537	57.43
109	REPORT	4,397	0.09	629	67.27
110	SUCH	4,294	0.08	792	84.71
111	MARKET	4,180	0.08	441	47.17
112	NOW	4,150	0.08	780	83.42
113	WHAT	4,074	0.08	733	78.40
114	INDIA	4,018	0.08	544	58.18
115	MADE	4,011	0.08	807	86.31
116	BEFORE	3,981	0.08	798	85.35
117	BECAUSE	3,966	0.08	768	82.14
118	HIGH	3,943	0.08	742	79.36
119	NATIONAL	3,911	0.08	767	82.03
120	ISLAMABAD	3,819	0.07	541	57.86
121	MAY	3,816	0.07	724	77.43
122	PARTY	3,812	0.07	591	63.21
123	MOST	3,755	0.07	751	80.32
124	EVEN	3,643	0.07	738	78.93

125	BETWEEN	3,590	0.07	726	77.65
126	MANY	3,517	0.07	739	79.04
127	CITY	3,511	0.07	612	65.45
128	TAKE	3,507	0.07	773	82.67
129	DO	3,431	0.07	754	80.64
130	ELECTION	3,426	0.07	477	51.02
131	LIKE	3,395	0.07	708	75.72
132	THROUGH	3,385	0.07	761	81.39
133	WHERE	3,357	0.06	760	81.28
134	MEETING	3,348	0.06	570	60.96
135	SECURITY	3,338	0.06	588	62.89
136	STATE	3,311	0.06	655	70.05
137	PRICES	3,308	0.06	315	33.69
138	HOWEVER	3,297	0.06	706	75.51
139	LAHORE	3,296	0.06	509	54.44
140	WELL	3,275	0.06	729	77.97
141	MILITARY	3,263	0.06	494	52.83
142	POWER	3,248	0.06	621	66.42
143	OFFICIALS	3,236	0.06	618	66.10
144	FOUR	3,216	0.06	688	73.58
145	HERE	3,189	0.06	680	72.73
146	OFF	3,165	0.06	683	73.05
147	INTERNATIONAL	3,132	0.06	673	71.98
148	CASE	3,114	0.06	655	70.05
149	ADDED	3,098	0.06	619	66.20
150	SINCE	3,096	0.06	714	76.36
151	ACCORDING	3,078	0.06	699	74.76
152	DID	3,063	0.06	740	79.14
153	STAFF	3,062	0.06	537	57.43
154	CENT	3,057	0.06	287	30.70
155	THEN	3,017	0.06	733	78.40
156	ANOTHER	2,994	0.06	744	79.57
157	MY	2,982	0.06	505	54.01
158	BANK	2,954	0.06	403	43.10
159	WEEK	2,941	0.06	588	62.89
160	JUSTICE	2,937	0.06	497	53.16
161	BILLION	2,934	0.06	401	42.89
162	DOWN	2,906	0.06	709	75.83
163	SEPTEMBER	2,896	0.06	759	81.18
164	NEXT	2,888	0.06	698	74.65
165	DISTRICT	2,877	0.06	425	45.45
166	VERY	2,863	0.06	700	74.87
167	MR	2,862	0.06	405	43.32
168	INCLUDING	2,861	0.06	718	76.79
169	FRIDAY	2,855	0.06	438	46.84
170	SHARIF	2,827	0.05	417	44.60
171	PUBLIC	2,808	0.05	699	74.76
172	FORMER	2,805	0.05	664	71.02
173	MAKE	2,804	0.05	732	78.29
174	ALI	2,787	0.05	602	64.39
175	THOSE	2,783	0.05	699	74.76
176	BACK	2,771	0.05	694	74.22
177	LAW	2,769	0.05	613	65.56
178	NAWAZ	2,767	0.05	436	46.63
179	FIVE	2,761	0.05	671	71.76
180	SOUTH	2,699	0.05	525	56.15
181	BOTH	2,683	0.05	712	76.15
182	MONTH	2,660	0.05	629	67.27
183	PRIME	2,645	0.05	600	64.17
184	JUST	2,640	0.05	678	72.51
185	OIL	2,614	0.05	319	34.12
186	SECOND	2,600	0.05	617	65.99
187	OFFICIAL	2,572	0.05	597	63.85

188	DEVELOPMENT	2,533	0.05	553	59.14
189	FOREIGN	2,519	0.05	572	61.18
190	DUE	2,518	0.05	674	72.09
191	HELD	2,507	0.05	656	70.16
192	ROAD	2,502	0.05	466	49.84
193	ARMY	2,498	0.05	484	51.76
194	WAY	2,482	0.05	706	75.51
195	DR	2,481	0.05	516	55.19
196	COUNTRIES	2,480	0.05	580	62.03
197	WORK	2,478	0.05	653	69.84
198	SINDH	2,473	0.05	376	40.21
199	LOCAL	2,471	0.05	588	62.89
200	GET	2,469	0.05	678	72.51

First top most frequent two hundred words have been initially retrieved from the corpus of Pakistani English. The list may gradually be enlarged according to the proficiency and the level of students. The list will help the learners to master the storehouse of English Language in an economical manner.

5. FINDINGS

The paper mainly establishes and exemplifies the authentic application of corpus in vocabulary building of the non-native learners of the English Language. It also motivates the teachers and learners to build vocabulary under actual language environment so that they can master the correct usage of English vocabulary. Corpus may be used in various dimensions of vocabulary building. Firstly, by developing a word frequency list, the learner can have a clear idea of the relative importance of the vocabulary. Secondly, the large number of examples from the corpus provides the learners with extended information about the meanings and usage of the target vocabulary items, which is an established advantage over traditional methods of vocabulary building. Thirdly, concordance lines will help in discovering the meaning of the word list. The empirical study in the paper proves that the corpus-based technique is helpful in vocabulary building and can lead towards autonomy at the same time.

6. CONCLUSION

This paper suggests language corpora can enhance the quality of vocabulary teaching and learning in second or foreign

language classrooms of Pakistan. By presenting benefits of language corpora to the pedagogy, it is hoped that this paper can be helpful to both teachers and learners who are struggling to search for an efficient way of teaching and learning vocabulary. Using corpus in an EFL class room in Pakistan is a new phenomenon. This study will also motivate the researchers to search further dimensions of lexicon building by developing specialized corpora in Pakistani context.

REFERENCES

- [1]. McCarthy, M. (1996). *Vocabulary*. Oxford, Oxford University Press.
- [2]. Laufer, B. & Melamed, L. (1994) 'Monolingual, bilingual and "bilingualized" dictionaries: Which are more effective, for what and for whom?' In Martin, W. et al. (eds.). *EURALEX 1994 Proceedings, Papers submitted to the 6th EURALEX International Congress on Lexicography in Amsterdam, The Netherlands*: 565-576. Amsterdam: Vrije Universiteit.
- [3]. Katamba, F. (1993). *Morphology*. Modern linguistics series. New York: St. Martin's Press.
- [4]. Bauer, L.(1998). *Vocabulary*. London and New York: Routledge.
- [5]. Sinclair, J. (1991), *Corpus, Concordance, Collocation*, Oxford: Oxford University Press.
- [6]. Schmitt, N. (2000) *Vocabulary in Language Teaching*. Cambridge University Press