

Application of Corpus Analysis in Teaching Foreign Languages

Ataboev Nozimjon Bobojon o'g'li

Ph.D. student, Uzbek State World Languages University



Abstract - The article deals with the application of corpus analyses in the process of teaching foreign languages in the classroom. At the current time, the usage of corpora and the quantitative results gained from them are being applied in almost all spheres of the linguistic studies as well as teaching methodologies. That's why the process of teaching foreign languages in EFL classes is taken into consideration as one of the most applicable fields for corpus database. Moreover, the practical tasks have been worked out in order to show the essence of corpus-based teaching methods in learning foreign languages. By all means, it is noteworthy that the EFL teachers always feel a great need for the authentic material for assessing their students' language skills as well as motivating them by supplying with adequate language data with the help of the corpus analyses.

Keywords - Language Corpus, Corpus Analysis, Target Language, Authentic Materials, Language Data, Vocabulary, Grammar, Punctuation.

Coming out of the above mentioned information, the use of the corpora varies from one another. It can be in the field of lexicography, lexicology, phonetics, grammar and the teaching pedagogy as well. We have learned several scholars' works devoted to the application of corpus analysis in the classroom/ teaching. Now, below their points of views are presented.

The first scientist whom we talk about is the German linguist, Ute Römer. He divided the use of corpora in language learning and teaching into two types: [Römer 2008]

indirect applications: hands-on for researchers and materials writers;

Here, the author includes the specialized corpora and the research results which can help the language learning process.

Direct applications: hands-on for teachers and learners (data driven learning, DDL).

Whereas, the direct way of use indicates the important features of the general corpora from which every learner

can do searches so as to get the necessary information like grammar constructions and others. Moreover, the teachers are able to visit the corpora in order to get some authentic materials for teaching purposes.

The next linguist who gave valuable information concerning the application of corpus analysis in teaching and the researches done on it is the American scholar, Casey Mari Keck. According to her, Corpus linguistics, through the use of computer technology, allows researchers to analyse large collections of language data (corpora) much more quickly and systematically than was previously possible. As a result, corpus-based findings have emerged that both enrich and challenge previous notions about language use. Since the early 1990s, researchers have become increasingly interested in applying the findings of corpus studies to language teaching. The work of Biber and colleagues (Biber and Reppen, 2002; Leech, 2001; Conrad, 1999; 2000) highlights a number of ways in which frequency information and register variation might inform the design of teaching materials and syllabi. Research in phraseology, such as Sinclair's (1991) concept of semantic prosody and Hunston and Francis' (1999) descriptions of

pattern grammar, has challenged traditional approaches to vocabulary and grammar instruction. Most recently, L2 corpora collected from language learner populations have been used to identify learner difficulties, such as the under or overuse of linguistic features, which deserve the attention of language teaching research. [C. Keck 2004]

The Chinese linguist, Professor Amy B M Tsui is the last in our list of scholars whose works have been studied by us. In her points of view, many researchers feel that corpus linguistics (CL) could revolutionize language teaching, by fundamentally changing the ways we approach all areas of pedagogy, including materials development, curriculum design, teaching methodology and teacher training. At the same time, the connections between CL and language teaching have only recently been made, and thus it is likely that a gap still exists between the teaching applications discussed in CL literature and the primary concerns of language teaching researchers and practitioners. To bridge this gap, corpusbased findings need to be interpreted in conjunction with findings from other areas of research on language teaching and second language acquisition. It is important then, to explore the major areas of interface between corpus linguistics and language teaching research.

So far, corpus-based studies have focused on four main types of description and analysis: [A. Tsui 200]

lexical collocation by examining the frequency and context of occurrence of linguistic items (Sinclair 1991; dictionary of collocations based on the Brown Corpus);

syntactic patterning based on co-occurrence of grammatical word-class tags;

genre analysis based on the co-occurrence of groups of linguistic items and processes (Biber 1998);

discourse structure and cohesion in spoken and written English (McCarthy 1996).

To summarize, the findings of the above studies have been particularly word-based studies, because a corpus is the real source of a great number of words and is able to supply both the teachers and learners with sufficient materials with the results in exact numbers.

We think that the exact results and examples in huge amount are very helpful for those who are teaching any foreign language or learning it, because the role of assumptions made according to these real linguistic numerous data is very important to provide the language learners with clear evidence. This is the way how the

teachers are able to make searches for the examples of the words being taught during the classes.

Now, let's see some of the examples which show the efficiency of the corpus analysis in teaching languages. Here, we have divided the application of corpus analysis in the classroom in three categories in accordance with the purpose and the audience of the users. In our opinion, they can be classified as following:

- using a linguistic corpus as teacher:
 - for textual materials;
 - for collocational examples;
 - for the task evolution on grammar, phonetics, and vocabulary.
- using a linguistic corpus as a language learner:
 - for self correction;
 - for getting more examples to acquire knowledge about the word;
 - for downloading an authentic material.
- using a linguistic corpus as a linguist or a student majoring in linguistics, translation and philology:
 - for compiling corpora with specific aims to investigate a language or a part of the language;
 - for making scientific assumptions about the target language;
 - for the authentic evidences concerning the language use.

It is time to illustrate each of the mentioned categories of users in teaching field one by one with the appropriate examples.

Let's start from the TEACHERS. This group of people are mostly eager to look up for original texts in order to supply the students with the materials based on the real language application by the native speakers. In fact, the native speakers are not always available for the teachers to record their voice to make listening tasks or to ask them to write a piece of text for constructing a reading task. That's why, the corpus of the target language as the real representative is the only choice for the educators. A corpus is considered to be much more superior than a single native speaker because the corpora represent whole nations' speeches. Here, we have organised several tasks based on the application of linguistic corpus.

Task 1

Subject: Vocabulary

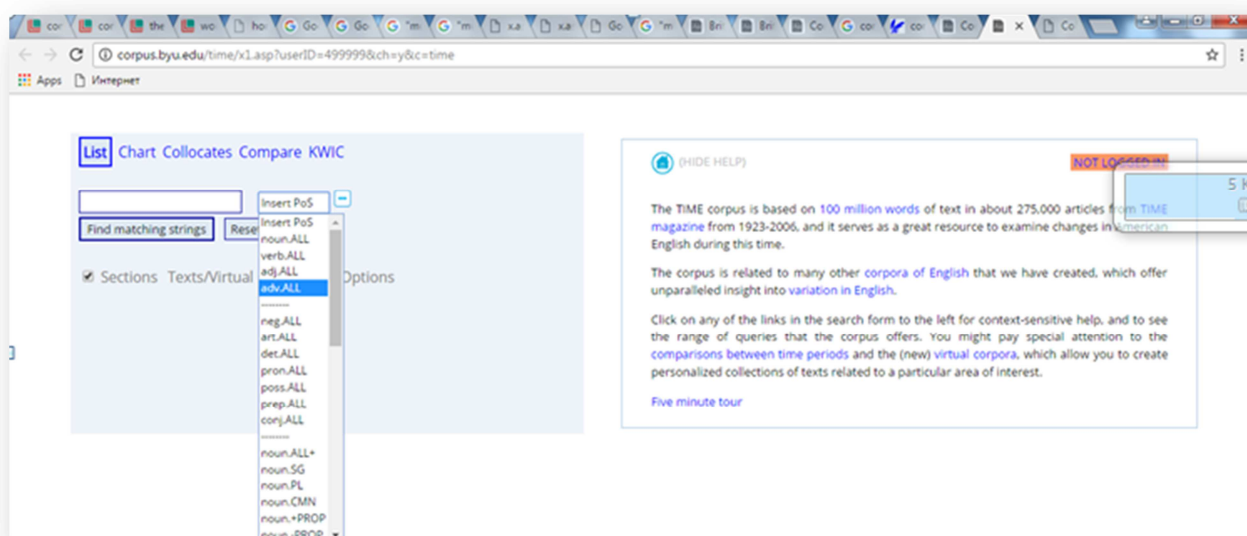
The name of the activity: Examples from concordance;

The aim: To familiarize the students with the new words' usage;

The stages of the activity: The teacher asks the students to find out real examples according to the new theme provided this day. Here, the teacher will recommend the learners a corpus and announces the limit of examples (at least 10 examples for each word). Visiting there, the language learners can find lots of examples which need to be read in order to choose. This is very beneficial for them to understand the topic vocabularies. Additionally, the teacher is supposed to ask the students to translate the found examples into their native language. This is also very useful to the knowledge seekers.

Here, to find examples is the simplest job to do in a corpus and it takes only several seconds. In fact, one word can have several meanings simultaneously, namely as a noun, a verb and an adjective as well as others. In this point of view, the users of corpora can indicate the parts of speech before making the search for it as it is shown in the picture 1. This will help them to do their task clearly as it has been required by the teacher. For making the other types of searches, one should follow the instructions for using a corpus.

Picture 1



Task 2

Subject: Grammar

The name of the activity: Mistake correction

The aim: To introduce some common grammar mistakes to the students;

The stages of the activity: The teacher asks the students which of the examples is true “much large” or “much larger”. To accomplish the task, the students have two options:

To make double searches and clarify which one is used more among the native speakers. As it is shown in the picture

2 and 3, the search was done twice for each of the examples and the results could show that having 599 examples, “much larger” is true comparing to “much large” owning only one example.

To make one search for both of them in the form of comparison and the result will be in the compared version of two phrases. Here, a glance at the picture 4 exposes the collected data which gives a clear view of the comparison. According to this search, the students can come to conclusion that “much large” is not used by the native speakers, however, the use of the phrase “much larger” is applied more frequently.

Picture 2

new British National Corpus (BYU-BNC) ⓘ ⓘ ⓘ

SEARCH FREQUENCY CONTEXT ACCOUNT

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

	CONTEXT	FREQ	
1	MUCH LARGE	1	

0.750 seconds

Picture 3

new British National Corpus (BYU-BNC) ⓘ ⓘ ⓘ

SEARCH FREQUENCY CONTEXT ACCOUNT

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

	CONTEXT	FREQ	
1	MUCH LARGER	599	

0.438 seconds

Picture 4

new British National Corpus (BYU-BNC) ⓘ ⓘ ⓘ

SEARCH FREQUENCY CONTEXT ACCOUNT

SEE CONTEXT: CLICK ON NUMBERS (WORD 1 OR 2)
SORTED BY RATIO: CHANGE TO FREQUENCY
WORD 1 (W1): MUCH LARGER (599.00) WORD 2 (W2): MUCH LARGER (0.00) NONE

	WORD	W1	W2	W1/W2	SCORE
1	MUCH	601	1	601.0	1.0
2	THE	204	0	408.0	0.7
3	.	188	0	376.0	0.6
4	,	179	0	358.0	0.6
5	THAN	153	0	306.0	0.5
6	A	271	1	271.0	0.5
7	AND	104	0	208.0	0.3
8	OF	168	1	168.0	0.3
9	IS	77	0	154.0	0.3
10	TO	69	0	138.0	0.2
11	IN	59	0	118.0	0.2
12	WITH	47	0	94.0	0.2
13	ON	42	0	84.0	0.1
14	ARE	39	0	78.0	0.1
15	IT	36	0	72.0	0.1
16	THAT	33	0	66.0	0.1
17	BUT	29	0	58.0	0.1
18	MORE	25	0	50.0	0.1
19	HAVE	24	0	48.0	0.1
20	SCALE	24	0	48.0	0.1

5 KB/s

Task 3.**Subject:** Grammar**The name of the activity:** analysing the punctuation**The aim:** To introduce the effective use of putuatinal marks to the students;

The stages of the activity: The teacher gives a word or word comination which is supposed to be used with special punctuational marks. For example, the teacher might ask students to identify the use of the putuational marks with the word “however”, that is, a conjunction.

By this task, the students can have a general point about this words' collocations and the application of the punctuations following it. The results of the search have shown the following view in the picture 5. It introduces to the

learners that mostly the word "however" comes between the sentences requiring double commas in both sides.

Picture 5

	Year	Source	Context	Highlighted Words	Collocations
1	1993	FIC	KansasQ	A B C	was nothing at all , a mere annoyance noticed however that the handkerchief he put to his mouth , he examined carefully
2	2006	ACAD	Education	A B C	common than how they would teach in the candidates responses however were statements asserting how they would think . Declarations
3	2012	MAG	America	A B C	phrase , " if you wish to be perfect however has also given rise to the idea that there must be
4	1996	MAG	Smithsonian	A B C	prospering steadily until the mid-19th century : in 1897 however the British Consul in what was then the Niger Coast
5	1990	ACAD	EnvirAffairs	A B C	the constitutionality of the fines provision. 55 in 1988 however Congress amended section 106 to provide a " good faith "
6	2005	ACAD	MechanicalEng	A B C	the late 1990s . # In the fall of 2000 however Congress stepped in and passed the Transportation Recall
7	2008	MAG	RollingStone	A B C	He was once a UFC star : in 2005 however two days after Lindland won a bout , the UFC fired
8	1991	SPOK	PBS_Newshour	A B C	realizes that what existed before can no longer exist again however he insists , and I think others insist , rightly so
9	1992	MAG	PopScience	A B C	using superbugs in the field . In the years ahead however genetically engineered microorganisms will probably play a
10	2014	MAG	Newsweek	A B C	n't know how to keep it . " # Allan however cries foul at this line of reasoning . " I've
11	2011	ACAD	MiddleEastQ	A B C	moderates are those who might be the West allies however only as long as they do not have to give up
12	2003	SPOK	CNN_Event	A B C	n't really a top concern for Iraqis : For Americans however in Iraq , it is a top concern , because it
13	2006	MAG	Astronomy	A B C	NGC 891 would be the highlight object . in Andromeda however it must defer to M31 . Still , it's worth
14	2000	ACAD	AmerStudies	A B C	is sacrilege. (n3) In one way or another however Indonesian feminist scholars have shown how the
15	2000	ACAD	AmerIndianQ	A B C	part of the day's work . " 93 Apparently however Hogsett did not take the war whoops simply as " part
16	1990	ACAD	EnvirAffairs	A B C	under the Fourth Circuit Court of Appeals approach however such a defendant would need to establish that proportionate

Taking into account the mentioned tasks and their efficiencies, we think that the use of corpus analysis by the teachers in the foreign language classrooms would be more successful. In our point of view, the application of corpus analysis in teaching is a modern trend and, that's why, the interests by the learners would be higher because of being computer aided. We are strongly convinced that every lesson designed with the use of corpora will be much superior to any other traditional language teaching classrooms. It is because of the unlimited source of data concerning the target language provided by the corpus.

REFERENCES:

- [1] Ute Römer, I. *Origin and history of corpus linguistics _ corpus linguistics vis-a-vis other disciplines*// 7. *Corpora and language teaching*: P. 112-122, - 2008 Hannover (Germany)
- [2] Casey Mari Keck *Corpus linguistics and language teaching research: bridging the gap* Language Teaching Research 2004; 8; 83 Northern Arizona University.
- [3] McEnery, Tony and Andrew Wilson. 1996. *Corpus linguistics*. Edinburgh: Edinburgh University Press.
- [4] Biber D., Conrad S., Reppen R, *Corpus Linguistics. // Investigating Language Structure and Use*. – Cambridge: Cambridge UP, 1998.
- [5] Geoffrey N. Leech, *Corpora: The Linguistics Encyclopedia*, ed. by Kirsten Malmkjaer. – Routledge, 1995.
- [6] Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- [7] Leech, Geoffrey. 1992. *Corpora and theories of linguistic performance*. In J. Svartvik (ed.). 105–122.
- [8] Amy B M Tsui *What teachers have always wanted to know - and how corpora can help*, 2009 The University of Hong Kong