Corpus Concordance for Enhancing EFL Learners' Vocabulary Retention

Dr. Ayman Mohamed El-Esery

Assistant Professor of Curriculum & TEFL, Qassim University

Uqlat Asoqour Campus, Saudi Arabia

E-mail: elesery_ayman@yahoo.com

Abstract

The snappy and dependable utilization of Concordancers forwarded them for use in studies. Since Lexis start language learning, advocates have tried to utilize concordancing programs in English language teaching. Computer programs are utilized as a regular part of language teaching, yet they don't take much place in the instructing of Lexis. This study first point of interest is the utilization of concordancing programs and gives methods that can be managed in ELT and represents sample uses of concordancing programs in teaching Lexis after reviewing literature related to concordance in ELT. Thirty EFL students enrolled at the Uqlat Asoqour College of Science and Arts, Qassim University represented the sample of this study. Results showed that the students using an online concordance learning approach had a higher level of retention in terms of word meaning and word usage in comparison with those using the regular vocabulary learning method.

Keywords: Corpus, Concordance, vocabulary retention, data-driven learning, CALL, computer software in ELT

1. Background

1.1 Learning Vocabulary

Vocabulary learning is a key piece of language learning. Learning words can thought to be the most essential part of language learning (Knight, 1994). Candlin (1988) expressed that "… the investigation of vocabulary is at the heart of language instructing as far as building syllabuses, the assessment of learner performance, and the procurement of learning resources. However, vocabulary learning is a very difficult process in which many skills are included. They are vocabulary acquisition, vocabulary retention, and vocabulary transfer (Schneider et al., 2002). Wei (2007) includes, vocabulary is still a principal test and the flow point of convergence of teaching and research. Over the previous decade or so, lexicology has been deeply influenced by technological developments.

1.2 Corpus Linguistics

Computer sciences have introduced crucial contributions to language pedagogy which have been observed within applied linguistics in constructing, processing, and analyzing language corpora (Johns, 1990). The term corpora refers to the electronic authentic language database(s) available on the Internet or stored in personal computers (Hasselgard, 2001; McEnery & Wilson, 2001). It is also defined as the body of e-texts considered to be representative or a sample of a language (Ball, 1996) or a collection of materials that has been made for a particular purpose, such as texts being analyzed for their linguistic features (Richard, J. Platt, & H. Platt, 1992). Language corpora can be either written texts using extracts from different media like fictions, newspapers, books, business letters, or magazines, published or unpublished school writings, etc.; or spoken texts involving any recorded and transcribed formal or informal conversations, radio and TV shows, weather broadcasts, business meetings, or even birthday parties, etc.

1.3 Concordancing Programs

Dyck (1999) presents three ways language educators can react to inquiries concerning word meaning; the prescriptive route in which the immediate clarification is given either by a grammarian or looking up a dictionary, but this method is short lasting. In the second way clarification is based on training and intuition. The third is is called concordancing which is a piece of corpus semantics. Peachey (2005) characterizes concordancer as a bit of programming, either introduced on a PC or got to through a site, which can be utilized to access, search and analyze language from a field or a corpus. Cncordancers gives words and their correct use in different authentic texts. Flowerdew (1996) characterizes concordancing as a method for getting to a corpus of content to demonstrate how any given word or expression in the content is utilized. As seen above definitions concordancers fill an essential need; concordancers give users the opportunity to see where and how and a given context a word is used.

1.4 Advantages of using a Concordance Program

Authenticity, transparency and learner autonomy are the three main advantages of using concordancers. "Authenticity" refers to the what should be acquired and learned by language learners. "Transparency" reflects the points of strength and weakness of learners since they can form different corpora by themselves. "Learner Autonomy" stands for the learners' ability to use corpora to learn themselves. Krieger (2003) asserts that Corpus linguistics presents an objective view of language than that of intuition, introspection, and anecdotes. With the suitable analytical instruments, an investigator can become aware of not best the patterns of language use, however the amount of usage, and contextual explanations that affect variability. For beginners who probably depend on finding out from concrete as an alternative than intuition or inference, making use of concordance program in language educating is more mighty and long-lasting.

A different advantage of concordancing is that it targets to bridge the gap between freshmen's present language capacity and anticipated ability (sun, 2007). That is finished by means of giving the official texts to scholars and while analyzing them, they are going to focus on the constructions or phrases which the instructor directs them. Moran and Diniz (2005) see this

knowledge through given that the truth that the concordancers show so many examples within the context of the equal word, language pupils are doubtless to infer the which means and patterns of the phrase. Davies (2012) states an advantage of concordancers wich is parallel concordancing from two languages; the parallel concordancer finds the identical sentences within the translated text. It is a giant possibility for analysers as they may be able to see a precise phrase or phrase utilized in two distinctive languages.

Concordancing programs are a real contribution to language learning not only for integrating technology to language courses, but in addition to motivating learners to concentrate attention on the course. Concordancing software help learners to have the sort of essential activities to concentrate on and computer systems will do away with any form of ambiguity that can be raised in the classroom. Regarding materials, concordancing programs by no means expire or end up old fashioned. Teachers will have to be able to make use of concordancing software to make their educating better, more enjoyable and lengthy- lasting.

2. Method

2.1 Research Design

This study followed a true experimental design, the researcher randomly selected and assigned subjects to two experimental groups (E1 and E2), and for capturing the initial differences between the groups a pre-test was administered.

2.2 Research Questions

The research questions of the study are given below:

1). Do web-based concordancing activities have any effect on vocabulary learning among EFL learners in comparison to the regular learning method?

2). Do web-based concordancing activities have any effect on the retention of vocabulary among EFL learners in comparison to the regular learning method?

2.3 Subjects of the Study

Participants in the course were 30 EFL students (E1=14 and E2=15) enrolled at Uqlat Asoqour College of Science and Arts, Qassim University. These learners were supposed to have a three-hours-per-week Listening and speaking III course, during which this experiment was administered at the faculty lab designated for this purpose.

2.4 Procedure

This study was undertaken in four stages: (1) pre-testing, (2) treatment, (3) post-testing, and (4) retention testing. The experiment included two participant groups. The first group (Experimental 1, E1) studied the vocabulary of listening and speaking III through concordance and corpora based activities, while the second group (Experimental 2, E2), acquired the academic vocabulary through the regular method, i.e. pronouncing words and repetition, writing new words on the board, introducing their meanings, explaining their parts of speech, using them in meaningful contexts.

Since both courses were designed to be taught for seven weeks, there were seven sessions in each course. At each of these sessions, seven English words were taught. All sessions require nearly 80 minutes to complete the specific activities. Each of these weekly sessions includes three main activity types. The initial part of the sessions were a learner guideline describing

what the learners were supposed to do during the session and giving them the list of the words they should study. The second part of the sessions was unique in terms of experimental one (E1). While the second section of the session included a concordance web page, however the learners were asked to work with the Corpus of Contemporary American English, which is probably the biggest corpus in the world with nearly 400 million words. This web page is also a significant example of an online data-driven learning environment, because it provides users and language researchers with many linguistic search opportunities and capabilities such as restricting your queries with text and genre types and production time. The third part of the sessions in the courses were designed to assess the students in terms of definitional and productive vocabulary knowledge types.

Vocabulary post-test was administered immediately after the treatment, to measure the vocabulary meaning and usage among the subjects. The same vocabulary test with some arrangements was administered to the same subjects as the second post-test three weeks later to measure the subjects' vocabulary retention.

2.5 Testing Word Meaning and Usage

A vocabulary test was used by the researcher in order to measure the subjects' ability to get the word meaning and usage. The vocabulary test included two main parts: (a) Part one was word meaning where the testees had to identify the correct meanings of words. (b) Part two was related to word usage where the subjects were asked to use the words to complete the extracts.

2.6 Test Reliability & Validity

Cronbach's alpha was computed for checking test reliability. Cronbach's alpha was 0.56 suggesting that the items had relatively high internal consistency. In order to investigate the dimensionality of the test, *factor analysis* was used; table 1 includes the resulting output.

Table 1. Factor Analysis for word meaning/usage Test

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.133	56.639	56.639	1.133	56.639	56.639
2	.867	43.361	100.000			

Looking at the previous table it is obvious that the eigenvalue for the first factor (meaning) was quite larger than the eigenvalue for the next factor (usage), 1.133 versus 0.867. Additionally, the first factor accounted for 56.6% of the total variance. This result meant that the test parts were unidimensional and that the test had construct validity.

3. Results

The data gathered throughout the study were analyzed using the Statistical Package for Social Sciences (SPSS version 22). Since the main research goal of the current study was to determine the effect of treatment on the independent variable, the most convenient statistical method was decided as Covariance Analysis (Ancova). The pre-test of the experimental groups were taken

as the covariate of the study. In addition, the statistical assumptions of the Ancova were checked before initiating the final analysis process and assured to pursue the analysis.

3.1 Word Meaning Test Report

Table (2) depicts the results of descriptive statistics of the pre, post and retention test results of word meaning test. The participant groups are labeled as E1 (Online Concordance Use Group) and E2 (Regular teaching group). According to the means of the pre-test results, both groups performed at a very similar rate (E1: = 12.06/ E2: = 12.2). These results were meaningful, as they revealed that there was no significant discrepancy between the groups pre-experimentally in terms of pre-learning backgrounds. However, the means of the post-test results showed a positive variance in favor of group E1 which studied through corpora and concordance based activities (E1: = 16.46/ E2: = 15.6). On the other hand, the retention test results pointed out that while there was a decrease in both groups in comparison to those of the post-test, the decrease in the group studying through Online coincidence was less than the one studying through the regular method (E1: = 15.4/ E2: = 13.13). The obtained descriptive statistical findings are illustrated in Table 1.

Table 2. Descriptive Statistics of Pre-test, Post-test and Retention Test Results of word meaning test

	Group	pretest	posttest	retentiontest
E1	Mean	12.0667	16.4667	15.4000
	N	15	15	15
	Std. Deviation	2.71153	1.72654	1.29835
E2	Mean	12.2667	15.6000	13.1333
	N	15	15	15
	Std. Deviation	3.01109	1.95667	1.45733
Total	Mean	12.1667	16.0333	14.2667
	N	30	30	30
	Std. Deviation	2.81723	1.86591	1.77984

The covariance analysis (Table 3) conducted on retention test results by taking pre and post test scores (to determine whether the variance in the means of post-test on word meaning test was statistically significant) proved that the difference in the corrected retention test results observed in favor of online concordance learning group was statistically significant {F = 18.059, p<.05}. Moreover, a mean difference of more than 1.3 points was observed in favor of online concordance learning group. In other words, the experiment showed that the students using online concordance learning approach had a higher level of retention in terms of word meaning in comparison with those using the regular method.

Table 3. Covariance Analysis Results of word meaning Test in comparison with Pre and Post Test Results

Source	Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	39.009[a]	3	13.003	6.396	.002
Intercept	58.695	1	58.695	28.872	.000
pretest	.465	1	.465	.229	.636
posttest	.046	1	.046	.023	.881
Group	36.713	1	36.713	18.059	.000
Error	52.858	26	2.033		
Total	6198.000	30			
Corrected Total	91.867	29			
a. R Squared = .425 (Adjusted R Squared = .358)					

3.2. Word Usage Test Report

Table 4 depicts the results of descriptive statistics of the pre, post and retention test results of the word usage test. According to the means of the pre-test results, both groups performed at a very similar rate (E1: = 11.73/ E2: = 11.93). These results were meaningful, as they revealed that there was no significant discrepancy between the groups pre-experimentally in terms of pre-learning backgrounds. However, the means of the post-test results showed a positive variance in favor of group E1 which studied through corpora and concordance based activities (E1: = 16.4/ E2: = 15.73). On the other hand, the retention test results pointed out that while there was a decrease in both groups in comparison to those of the post-test, the decrease in the group studying through Online coincidence was less than the one studying through the regular method (E1: = 15.66/ E2: = 12.93). In other words, it can be concluded that the maintenance of learning through online concordance approach has a more long-term effect on learning. The findings obtained are shown in Table 3.

Table 4. Descriptive Statistics of Pre-test, Post-test and Retention Test Results of word usage test Report

	Group	pretest	posttest	retentiontest
E1	Mean	11.7333	16.4000	15.6667
	N	15	15	15
	Std. Deviation	3.17280	1.84391	1.79947
E2	Mean	11.9333	15.7333	12.9333
	N	15	15	15
	Std. Deviation	2.76371	1.86956	1.03280
Total	Mean	11.8333	16.0667	14.3000
	N	30	30	30
	Std. Deviation	2.92532	1.85571	2.00258

The covariance analysis (Table 5) conducted on retention test results by taking pre and post test scores (to determine whether the variance in the means of post-test on word meaning test was statistically significant) proved that the difference in the corrected retention test results observed in favor of online concordance learning group was statistically significant {F = 30.677, p<.05}. In other words, the study showed that the students using online concordance learning approach had a higher level of retention in learning in comparison with those using the regular method in vocabulary usage.

Table 5. Covariance Analysis Results of word usage Test in comparison with Pre and Post Test Results

Source	Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	65.205[a]	3	21.735	11.060	.000
Intercept	103.032	1	103.032	52.429	.000
Posttest	6.569	1	6.569	3.343	.079
Pretest	4.487	1	4.487	2.283	.143
Group	60.286	1	60.286	30.677	.000
Error	51.095	26	1.965		
Total	6251.000	30			
Corrected Total	116.300	29			
a. R Squared = .561 (Adjusted R Squared = .510)					

5. Conclusion

Both of the experimental groups performed better in terms of word meaning and usage according to the statistically positive difference between pre and post test results. However, the post test results showed no significant difference between the groups. The retention test results indicated that online concordance learning group performed better than the group that learned through the regular method.

This paper has focused on the importance of online corpus concordance learning approach to vocabulary instruction. The discussions raised point out the role and place of explicit vocabulary instruction in vocabulary development, as well as issues of possible pedagogical implications and specific methods and techniques for fostering vocabulary development. The paper has also presented an important issue; the role of technology in bringing a more systematic and data-based approach and innovative methods and techniques to vocabulary instruction and learning. The Data-driven learning approach can be implemented in intensive English language programs focusing on vocabulary learning. New corpora can be compiled for the sake of language learning purposes. Pre-service teachers should be provided with the required methodology to exploit the data-driven learning approach in order to assess learners, design of materials and development. Vocabulary and grammar teaching materials designed and developed according to data driven learning approach should be incorporated into foreign language classrooms. Corpora and concordance based language learning materials and activities can also be delivered to learners through learning systems.

References

Ball, C. N. (1996). *Tutorial Notes: Concordances and Corpora*. Retrieved from http://www.georgetown.edu/cball/corpora/tutorial.html

Candlin, C. N. (1988). Preface. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teachin*g. New York NY: Longman.

Davies G. (2012) *Information and Communications Technology for Language Teachers (ICT4LT), Slough, Thames Valley University*. Retrieved from http://www.ict4lt.org/en/en_mod2-html

Dyck, G. N. (1999). *Concordancing for English teachers. Paper presented at TESL Manitoba, Canada*. Retrieved from http://home.cc.umanitoba.ca/~gdyck/conc.html

Flowerdew, J. (1996). Concordancing in Language Learning. In M. Pennington (Ed.), *The Power of CALL (Athelstan*) (pp. 87-102).

Hasselgard, H. (2001). Corpora and their use in research and teaching. Retrieved from http://www.folk.uio.no/hhasselg/UV-corpus.htm.

Johns, T. (1990) From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data- driven Learning. *CALL Austria, 10*, 14-34.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal, 78*(3) 285-299.

Krieger, D. (2003) Corpus Linguistics: What It Is and How It Can Be Applied to Teaching. *The Internet TESL Journal, IX*(3). March 2003. Retrieved from http://iteslj.org/Articles/Krieger-Corpus.html

McEnery, A., & Wilson, A. (2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Moran K, Diniz L. (2005). Teaching with concordances. *On CALL 2005, 22*(2).

Peachey, N. (2005). *Concordancers in ELT. In British Council teaching English*. Retrieved from http://www.teachingenglish.org.uk/think/articles/concordancers-elt

Richard, J. C., Platt, J., & Platt, H. (1992). *Dictionary of Language Teaching & Applied Linguistics*. Essex: Longman.

Schneider, V. I., Healy, A. F., & Bourne Jr. L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects on foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language, 46*(2), 419-440.

Sun Y. C. (2007). Learner Perceptions of a Concordancing Tool for Academic Writing. *Computer Assisted Language Learning, 20*(4), 323–343.

Wei, M. (2007). An examination of vocabulary learning of college-level learners of English in China. *Asian EFL Journal, 9*(2), 93-114.

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).