The Influence of Data-Driven Exercises Through Using a Computer Program on Vocabulary Improvement in an EFL Context

Touraj Talaei
Taravat Language Institute, Borujerd, Iran

Zahra Fotovatnia
Assistant Prof., Islamic Azad University, Najafabad Branch
fotovatnia@yahoo.com

Received 22 February 2012; accepted 5 July 2012

Abstract

The present study was conducted to evaluate data driven learning (DDL) combined with Computer Assisted Language Learning (CALL) as an approach to improving vocabulary knowledge of Iranian postgraduates majoring in teaching English, English literature and translation. The purpose was to help language learners get familiar with DDL as a student-centered method taking advantage of a computer program for enhancing the knowledge of vocabulary. This research paper analyzed the efficiency of DDL from different angles including the participants' ability to remember the word meanings and make meaningful and grammatical sentences after doing DDL exercises using the Microsoft Power Point program. The results of the study showed that this technique could significantly enhance the participants' vocabulary knowledge in all the above mentioned aspects. The analysis of the questionnaire given to the participants to elicit their attitudes towards DDL showed that most of them were interested in this technique. The results of this article can be helpful for the language teachers who teach the learners at advanced levels of language proficiency and advanced students who would like to learn new items of vocabulary on their own.

Keywords: Linguistic corpus, Computer Assisted Language Learning (CALL), Data-Driven Learning (DDL)

Vocabulary is regarded as one of the fundamental components of any language. Through using the words of the language, people convey what they mean. Therefore, vocabulary learning is something to which a lot of attention should be paid. Many studies have been conducted about useful methods and techniques for vocabulary improvement in L2. The learners of second or foreign languages can take advantage of the results of such studies to find an appropriate method or technique for themselves. Even language teachers can exert the new methods and techniques that have been introduced in these studies in order to create an atmosphere of excitement in their classes and get the best out of their teaching. With the use of computers, teachers can exert mnemonic techniques through presenting the exercises that involve the use of visual programs. For instance, many concrete vocabulary items can be taught to the learners by showing their pictures on a computer monitor. The teachers can design a computer program for testing whether or not the learners have learned the new items of vocabulary in every session. Even natural instances of language collected in the form of a corpus can be used as the basis for teaching and learning in and outside regular classrooms.

Corpus refers to a large collection of naturally occurring language produced by native speakers gathered from both spoken and written instances of language use (Richards & Schmidt, 2002). One of the advantages of corpus is its capability in familiarizing language

learners with different patterns in the target language. With the use of corpus, one has the opportunity for analyzing natural language use and finding the differences between formal and informal language. We can also learn about the features of spoken as compared with written language. At the present time, corpora are mostly computerized but in the past, they were produced in printed versions. The advent of computers has made us able to search for special words and patterns among corpora much faster and easier.

Collecting large amounts of authentic language is a time consuming activity which has been simplified with the use of computer programs at the present time. Besides searching for keywords and their surrounding contexts, these programs can also calculate the frequency of vocabulary items. The ability of computer programs in the field of corpus linguistics is not limited to the ones mentioned above. These programs have other abilities such as finding collocations related to special vocabulary items. Since different wide corpora are available for teachers and researchers, there is no need for them to produce a corpus by themselves.

Corpora can be used both indirectly and directly. If corpora are used indirectly, it means that they help teachers and syllabus designers in making decisions about what materials to teach and when to teach them. Barlow (1996) notes that corpora can influence course design and determine the content of the materials which are going to be taught in language classrooms. On the other side, if they are used directly, it means that they are used directly by learners and teachers for the purpose of getting familiar with the use of language in the real world or as Fligelstone (1993) says, they help teachers in the teaching process.

Corpora are not used only for the mere analysis of languages. Nowadays, they also play the role of linguistic materials for language classes. Woolard (2000) expressed that corpora and the use of concordancing which were used only for the purpose of research about the language in the past, is being used as an important instrument for language teachers in their classes at the present time. Having access to corpora may be thought of as something difficult for language learners, but many authors who have done investigations about the use of corpora have published articles on the fact that corpora are easily accessible to the learners (e.g. Fox 1998; Kettemann, 1995; Stevens, 1995; Tribble, 1997; Wichmann, 1995).

While the indirect use of corpus focuses on its influence on syllabus design and linguistic materials for the purpose of teaching, the direct use of it focuses more on the teachers and learners. In the direct approach, instead of relying on researchers as the providers of linguistic materials, language learners and teachers themselves study the corpus in order to discover language patterns, word meanings, etc. (Bernardini, 2002). Tim Johns pioneered the direct use of corpus for learning grammar and vocabulary. Johns (2002) believed that the learners should encounter the linguistic data directly and tried to foster the role of learners as linguistic researchers. Johns (1997) also regarded the learner as language detective and called every learner as a Sherlock Holmes. This method, in which there is an interaction between corpora and students, is known as DDL. The activities of DDL are usually taken from corpora.

Some investigations have been done about the efficiency of using corpora directly for language learning. For example, Watson (2001) conducted such an investigation. In his research, the students used the Internet as a source of linguistic corpus. In this study, the students were required to correct two mistakes they made in writing a report as their assignment. The mistakes were indicated by the teacher and the students were supposed to search the Internet to find examples of use for the indicated words as a corpus and induce rules in order to self-correct their mistakes. The results showed that students could correct their mistakes in 78% of instances.

Although many investigations have been done on the usefulness of corpora in language teaching and learning, the findings are not used widely in language classes. As Tribble (2000) said it seems that only few teachers use corpora in their classes for language teaching.

Seidhofer (2002) noted that many teachers and students are unaware of the influence corpus linguistics has on language description and development of language materials. Mukherjee (2004) remarked that the findings of investigations that have proved the advantages of corpora have not fostered their use in real language classes. Braun (2005) concluded that although corpora are the buzzword in language research departments, they have not become a major part of teaching practice yet.

The materials taken from corpora can be presented to language learners through computer programs. DDL is a technique in which computer programs can play a significant role. In DDL, a corpus can be used as a source of information for the students. The teacher can find the frequent vocabulary, grammar, collocations, and idioms in a corpus in order to produce data-driven exercises. The students themselves should discover the rules and meanings and this responsibility for learning can make it enjoyable for them. With pressing the buttons of the keyboard, the learners can observe new items of vocabulary and grammatical points on the screen of the monitor and in this way, without the direct assistance of any teacher; they can improve their language proficiency.

Learners' attitudes towards different methods of language learning are worth considering. Whenever a new method proves to be successful, it may be the result of the learners' satisfaction with that method. Consequently, attitude questionnaires can be distributed among the learners in order to ask their ideas about any treatment they receive for improving their knowledge of language. CALL as a modern means for presenting data-driven exercises to the learners has its own fans.

In some investigations about CALL, the participants of the studies have been asked to answer some questionnaires about their attitudes towards this method of learning. For instance, Finkbeiner (2001) gave 100 questionnaires to undergraduate EFL learners in order to understand their ideas about CALL. He collected 82 questionnaires and realized that the learners had a positive attitude towards it. In another study conducted by Ayres (2002), 157 language learners participated in a computer-based learning program and the results of the questionnaires showed that the participants appreciated learning through CALL. Allum (2002) noted that learners had positive feelings about this type of leaning and said that CALL should become part of regular classes.

Although many studies show positive attitudes towards CALL, some others show the opposite of this finding. For example, in a research done by Shaw and Marlow (1999), the participants did not like learning the language through using computers. In other words, they preferred traditional methods over CALL because in traditional methods there is personal contact with real humans. Such negative attitudes may result from teachers' and learners' adherence to traditional methods for a long time so that traditional methods have become a permanent habit of language learning for the learners. Unfortunately, many language teachers in Iran are technophobic and prefer to exert traditional methods for language teaching in their classes. Such an opinion has also been injected into the minds of a lot of Iranian students. Therefore, in the present study, the performances of two Iranian groups of students have been compared with each other. In one group, the learners were exposed to a traditional method and in the other group, they received data-driven exercises through the use of a computer program to see which method is more successful. Also to the best of the researchers' knowledge, almost no study has been done on the same topic and no study looked for the learners' attitude towards DDL.

In this study the following questions were addressed:

1. Does the use of DDL have any positive effect on the learning of word meaning and word grammatical features for advanced learners of English?
2. Is DDL interesting and motivating for the students? What are the learners' opinions about DDL and the presentation of such type of material through the computer?

Methodology

Participants

The experiment was conducted with a sample of 40 Iranian university students, studying English as their major at M.A. level. The reason for choosing M.A. students of English was that according to the related literature, DDL seems to be more appropriate for the students at higher levels of language proficiency. Eighteen of the participants were males and twenty two were females. All of the participants were native speakers of Farsi, aged between 24-45. The participants were selected from the population of EFL learners, studying English in universities of Borujerd, Najafabad, Tehran, Kish and Kermanshah. The criterion for choosing these 40 participants was the results of Nation's Vocabulary Test (productive version) given to them.

Instruments

In this study, six types of materials were used for data collection including Nation Vocabulary Levels Test (1983), pre-test, post-test, power point slides, printed information taken from a monolingual dictionary and a questionnaire.

Nation's vocabulary test. The productive type of this test is composed of words at different frequency levels orderly 2000, 3000, 5000, academic and 10000. This test was used to ensure the homogeneity of the participants who were later divided into experimental and control groups. There are 18 items at each level and the participants are supposed to complete an incomplete word used in context. The first letters of incomplete words can help them to guess the complete word. The participants should have answered at least 16 out of 18 items in the productive test on each level in order to approve they don't have problem in understanding vocabulary at that frequency level. The purpose of administering this test was to determine the participants who had problem with the words at the frequency level of 10000 of English vocabulary.

Pre-test. As it was mentioned in the above section, the participants who had problem with the words at the frequency level of 10000 were diagnosed, but it could not mean that they were not familiar with any vocabulary item at that frequency level. Consequently, a pre-test was designed by the researchers to select only the words at the frequency of 10000 that were unfamiliar to all the participants. The pre-test consisted of 100 vocabulary items from the frequency level of 10000 which was diagnosed as an unknown area of vocabulary for the participants according to the results of the Nation's Vocabulary Test given to them. The participants were supposed to write the Farsi or English meaning of the items they knew in front of each one and leave the unknown words blank. Based on the results, 15 words whose meanings were unknown to all the participants were selected for the treatment and posttest.

Post-test. In the post test, the participants were supposed to write a meaning for each of the 15 words whether in Farsi or English and make a sentence in English for each one. The participants were asked to answer the items they knew and leave the others blank.

Power point slides. Fifteen slides were prepared using PowerPoint, one slide for each word. In each slide, 4 sentences taken from British National Corpus (BNC) were used. The sentences were sorted in a way that the target words in all sentences were placed exactly under the one in the above sentence. At the lower part of each slide, 4 definitions were

presented. Out of these definitions, 3 were distracters and only one showed the meaning of the target word.

Printed information taken from a monolingual dictionary. The same 15 words which were used in the power point slides were printed on a piece of paper with one definition and one sentence or phrase example taken from Oxford and Longman dictionary in front of each vocabulary item.

Questionnaire. After the post test, a questionnaire including 20 items developed by the researchers was given to the participants in the experimental group. Before giving this questionnaire to the participants, it was piloted with 10 other university students majoring in English at the M.A. level. It was also given to 2 Ph.D. holders in language teaching for the sake of approving its construct validity. The reliability index of the questionnaire was measured by SPSS version 15 (.78). Each item of the questionnaire was measured on a 5-point Likert scale including *(1) strongly disagree, (2) disagree, (3) no idea, (4) agree, (5) strongly agree*. The questionnaire included questions about data-driven exercises which had been presented as the treatment of the study to the participants of the experimental group. At the end of the questions, the participants were required to express any idea they had about DDL not mentioned in the questionnaire.

Procedure

For collecting data in this study, several measures were taken. First, a sample of Nation's Vocabulary Test including productive questions was administered to the participants who were university students majoring in English teaching, translation and literature at M.A level. Based on the scores, 40 students who were homogeneous in vocabulary knowledge were selected. The reason for choosing these 40 students as homogeneous ones was that they all had problems in answering the questions related to the frequency level of 10000.

Second, the pre-test was administered to all the participants of the study. Because these words had been taken from a frequency level of English vocabulary which had been diagnosed as the level the participants had problem with, most of them could only answer a small number of the items and left most of them blank.

Third, the participants of this study were randomly divided into two groups of 20 students in each group to receive the treatment. There were 10 male and 10 female participants in group 1 and there were 8 male and 12 female participants in group 2. For the treatment of the study, 15 out of 100 words in the pre-test which were unknown to all of participants were selected as the items to be used in the treatment.

The treatment for the control group consisted of 15 vocabulary items. The participants received these 15 vocabulary items with their English definitions on a piece of paper. They were supposed to study the words in order to learn what they mean and how to make sentences using those words. The allotted time was 20 minutes.

The treatment for the experimental group consisted of the same 15 words which were used for the control group. In this group, the treatment was not given to the participants in the form of printed papers. Instead, some power point slides were prepared and the new words were shown to the participants on a computer monitor. It was told to the participants that with each time pressing the Enter button of the keyboard, a slide is shown on the screen. For each word at first, the participants could see the new word and its pronunciation in front of it.

In the next step, with each time pressing the Enter key, a sentence taken from British National Corpus (http://www.natcorp.ox.ac.uk/) in which the new item of vocabulary had been used was shown on the screen. For each word, 4 sentences of this type had been prepared. So, after 4 times pressing the Enter key, 4 sentences were shown under each other.

The sentences had been concordanced. In other words, the sentences had been arranged so that the new item of vocabulary in each sentence had been placed exactly under the one used in the upper sentence. The new word had also been typed in a different color from the whole sentence. Both concordancing and the use of a different color for the new word could attract the attention of the participants and raise their consciousness about its different features. It could also help them to get familiar with the immediate context and the collocations used with each item in the sentences presented to them.

At this time, the participants could guess the meaning of the new word. In order to make sure that the participants were able to understand the exact meaning of the target word, they could press the Enter key again. In this way, 4 definitions were shown on the screen. Three of these definitions were distracters and just one of them went with the correct meaning of the target word. It was not a multiple-choice question but only a guide by means of which the participants could find the correct meaning in order to empower their guess about the meaning of the word. The reason why the meaning was presented among some other distracters was to give this opportunity to the participants to discover the meaning by themselves and create an atmosphere of student-centered learning. The distracters were not highly related to the target word; so it was very easy for the participants to find the correct definition. It was told to the them that they could review the slides once more in order to ensure about the correct meaning. The whole time allocated to reading and reviewing the words to be learned was 30 minutes.

Data Analysis

First, the normality tests were run to ensure the normality of distributions. As Table 1 shows, the values of skewness and kurtosis are between -2 and 2 so that the distributions are supported to be normal.

Table 1. *Normality Values of the Distributions*

	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
Number of correct meanings	40	-.447	.374	-.352	.733
Number of correct use regarding meaning	40	-.262	.374	-.623	.733
Number of correct use regarding grammar	40	.094	.374	-.286	.733
Valid N (listwise)	40				

Next, a series of independent samples *t* test was run to find out if there was any significant difference between the control and experimental groups. The descriptive statistics are shown in the following tables. Table 2 shows the number of correct meanings of words out of the context.

Table 2. *Descriptive Statistics Related to Correct Meanings*

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Number of correct meanings	Traditional	20	6.85	3.717	.831
	Data driven	20	10.55	2.964	.663

As Table 2 displays, DDG outperformed TG. The independent samples *t* test shows a significant difference between the two groups, *t* (38) = 3.48, *p* = .001. Table 3 shows the descriptive statistics of participants' performance on the second part of the production test.

Table 3. *Descriptive Statistics Related to Correct Meanings of Words in Context*

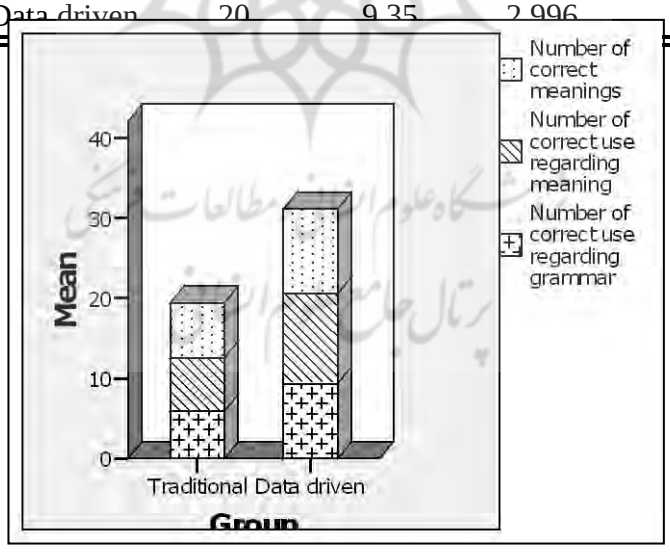
Group	N	Mean	Std. Deviation	Std. Error Mean
Traditional	20	6.65	3.514	.786
Data driven	20	11.25	2.900	.648

As Table 3 shows, DDG obtained a greater mean than TG, meaning that more correct sentences were written by the participants in this group insofar as meaning of the target words in the context was concerned. An independent samples *t* test shows the difference is significant, *t* (38) = 4.51, *p* = .000. Table 4 shows the descriptive statistics of correct sentences in terms of grammar.

Table 4. *Descriptive Statistics of Correct Sentences in Terms of Grammar*

Group	N	Mean	Std. Deviation	Std. Error Mean
Traditional	20	5.95	3.284	.734
Data driven	20	9.35	2.996	.670

As the above table displays, DDG achieved a higher mean than TG. The independent samples *t* test shows the difference is significant, *t* (38) = 3.42, *p* = .002. Figure 1illustrates the same results.



Firgure 1. Means of DDG and TG on different parts of the posttest

The questionnaire was analyzed through computing the frequency and percentage of responses to each scale. The following table shows the results.

Table 5. *Percentage of Responses to Each Scale*

Questionnaire items	Agree	No idea	Disagree
Interesting for learners	75%	20%	5%
Useful exemplars for guessing the meaning	90%	5%	5%
Difficult and complex exemplars	35%	10%	55%
Useful only for advanced learners	65%	20%	15%
Useful for learning collocations	50%	40%	10%
Useful for learning part of speech	90%	5%	5%
Useful for making correct grammatical sentences	95%	0%	5%
Time consuming	50%	25%	25%
Including difficult words plus new ones in exemplars	55%	5%	40%
Better if fewer definitions	40%	20%	40%
Attracting way for learning do to the use of power point	65%	30%	5%
Useful due to individual appearance of exemplars	70%	25%	5%
Better learning if possible to review	85%	15%	0%
Better learning if more time	70%	15%	15%
Better if more than 4 exemplars	20%	40%	40%
Better if fewer than 4 exemplars	5%	35%	60%
Better if more definitions	20%	40%	40%
Better if fewer definitions	15%	50%	35%
Better if time limitation for each slide	5%	25%	70%
Utility of definitions in guessing meaning	70%	25%	5%

The percentage values displayed in the above table can be used to show learner beliefs about different aspects of the employed technique. The first questionnaire item asked if the learners found this technique interesting for vocabulary learning. About 75% of participants agreed it was. The second questionnaire item inquired if learners found the exemplars useful in guessing the meanings of the target words. About 80% of the participants agreed the exemplars were utilized to guess the meaning of new words. The third item of questionnaire asked whether or not the exemplars were difficult and approximately 55% of the participants disagreed. The fourth item asked if DDL was only useful for advanced learners. About 65% of the participants approved of this attitude. The fifth item investigated whether this technique was useful for learning collocations. Almost 50% of the participants claimed it was. The sixth questionnaire item inquired if the exemplars were useful in conveying the part of speech for the target words. About 90% of the participants believed it was. The seventh item of the questionnaire asked if the exemplars were useful for making the learners able to make grammatical sentences using the target words. About 95% of the participants agreed it was. The eighth item inquired whether or not this technique was time consuming and 50% of the participants believed it was. The ninth item asked if the exemplars included unknown vocabulary items besides the target word. Approximately 55% of the participants agreed they included more than one new and unknown vocabulary item for them. The tenth questionnaire item inquired if it was better that fewer definitions were presented as an assistance to help the learners guess the meaning of the target words. Almost 40% of the participants agreed and 40% of them disagreed. The eleventh question was about the inclusion of computer and the use of power point. Around 65% of the participants believed that the use of power point made this type of learning attractive. The twelfth item asked if the appearance of the exemplars on the screen individually and not all at one time was useful. Among the participants, 45% agreed with it. The thirteenth question investigated whether or not devotion of more time to study the exemplars could lead to deeper learning. Around 70% of the participants believed it could. The fourteenth questionnaire item inquired if review of the exemplars could lead to better learning of the new vocabulary items. Approximately 85% of the participants believed that review could deepen their learning. The fifteenth question asked if the presentation of more than 4 exemplars could be useful. Almost 20% of the participants agreed and 40% disagreed and 40% of them had no idea about this. The sixteenth item asked if the presentation of less than 4 exemplars could be better. About 60% of the participants disagreed. The seventeenth item inquired if giving more definitions could be better and almost 30% of the participants disagreed and 40% of them had no idea. The eighteenth item asked if giving fewer definitions could enhance correct guessing and around 15% of the participants agreed, 25% disagreed and 50% had no idea about it. The nineteenth item inquired whether time limitation for each slide in the power point could be better. Almost 70% disagreed. The twentieth item asked if the definitions were helpful in guessing the meanings of the target words. Approximately 70% of the participants claimed they were useful. Finally a one-sample *t* test was run to determine the general attitude of learners toward DDL. Learners have been found to have positive attitude, $t\ (19) = 70.58$, $p = .000$.

Discussion and Conclusion

The main focus of this study was to investigate if data-driven exercises through employing power point software could improve the learners' vocabulary knowledge and if learners showed positive attitude towards it. For the sake of this purpose, the performances of the participants receiving data-driven exercises were compared with those of the participants who received a traditional type of treatment.

The first research question asked about the effectiveness of DDL as compared with the traditional method of vocabulary improvement. The results of the study showed that studying new words of English as a foreign language through data-driven input significantly enhanced leaners' vocabulary learning irrespective of the output they produced. They performed better on meaning of the words out of the context, on the meaning of the whole sentences embracing the words and on grammaticality of the sentences they wrote.

One of the factors which could help DDL be more successful in this study might be the inclusion of concordanced sentences in which the target words in the sentence examples had been sorted exactly in an imaginary column. This could attract the attention of the learners to the immediate context of the target words and make them familiar with the word collocations and prepositions. This factor may be one of the reasons for the greater success of the participants in the experimental group to produce correct grammatical sentences in the post-test. This issue is confirmed by Schmitt (2002), who declared that concordance programs can support the learners with the immediate context of word occurrence and this can make them familiar with the most frequent words that occur to the right or left of the target word.

Another factor which may have made DDL more successful than the traditional method used in the control group might be the number of the sentence examples presented to the learners in the experimental group. The participants receiving data-driven exercises were exposed to four exemplars as compared with the participants in the control group who did not receive authentic data. Repeated exposure to different contexts in which a special vocabulary item had been used could help the learners to learn the new word deeply. Nagy and Herman (1985) conducted a study about the significance of learning new vocabulary items in different contexts. They concluded that when the learners see a word for the first time, they have a chance of only 5-10% for remembering it. They also expressed that 10-12 exposures to the new words in different contexts can help the learners keep those words in their memory.

The second research question in this study asked if DDL was interesting for the learners. According to the results of the questionnaire, most of the participants believed that DDL was exciting and motivating for them. This excitement may come from the inclusion of power point as a computer program for presenting data-driven exercises to the participants. There was a question in the questionnaire asking whether the learners were interested in the inclusion of power point in regular courses and most of them agreed. This might support the conclusion that the inclusion of computer and the related software was a main factor which made this technique interesting for the learners. In a study conducted by Finkbeiner (2001) 100 questionnaires about CALL were given to 100 undergraduate EFL learners to see if they had positive attitudes towards it. The researcher of the above mentioned study could collect 82 questionnaires and found that the learners had positive feelings regarding the use of computer programs for language learning.

Many teachers have adhered to the old methods of teaching vocabulary. The results of the present study are in favor of DDL as a new method for vocabulary teaching and learning. Unfortunately, many teachers do not prefer using computers in the environment of the classroom while they follow traditional methods of language teaching. DDL through employing computer as a teaching aid was more successful than using printed papers.

If DDL is used for vocabulary learning, the students themselves will be responsible for discovering word meanings. Because in the present study the students were supposed to learn the meaning of 15 new words that they did not know before through data-driven leaning which is a student-centered method, the researcher of the study presented more than one sentence example for each word to the participants to make it possible for them to learn the meaning of the words without the direct assistance of any teacher. Because of the fact that the students could encounter the new words in different contexts instead of only one sentence example which is a norm in dictionaries, they had more opportunities to get familiar with the

correct use of the words and even their collocations. This factor could help this method to be more successful than the one used in the control group.

References

Allum, P. (2002). CALL and the classroom: the case for comparative research. *ReCALL, 14* (1), 146-166.

Ayres, R.(2002). Learner attitudes towards the use of CALL. *Computer assisted language learning, 15* (3), 241-249.

Barlow, Michael (1996), Corpora for Theory and Practice. *International Journal of Corpus Linguistics 1*(1), 1-37.

Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann &G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp.165-182). Amsterdam, NewYork: Rodopi.

Braun, S.(2005). From pedagogically relevant corpora to authentic language learning contents. *ReCALL, 17*(1), 47-64.

Finkbeiner, C. (2001). One and all in CALL? Learner-moderator-researcher. *Computer assisted language learning, 14* (3-4), 339-361.

Fligelstone, S. (1993). Some reflections on the question of teaching, from a corpus linguistics perspective. *ICAME Journal 17*, 97-109.

Fox, G. (1998) Using corpus data in the classroom. In Tomlinson, B. (Ed.), (1998) *Materials Development in Language Teaching* (pp. 25-43). Cambridge: Cambridge University Press.

Johns, T (1997), Contexts: The Background, Development and Trialling of a Concordance-based CALL Program. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100-115). London: Longman.

Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 107-117).

Kettemann, B. (1995) On the use of concordancing in *ELT. TELL&CALL 4*, 4-15.

Mukherjee, J. (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 239-250). Amsterdam, NewYork: Rodopi

Nagy, W. & Herman, P. (1985). Incidental versus instructional approaches to increasing reading vocabulary. *Educational perspectives*, 23, 16-21.

Richards, J.C & Schmidt, R (2002). *Longman dictionary of language teaching and applied linguistics.* London, Pearson Education.

Schmitt, N. (Eds.). (2002). *An introduction to applied linguistics.* New York, Oxford University Press.

Seidlhofer, B. (2002). Pedagogy and local learner corpora: Working with learning-driven data. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 213-234). Amsterdam, Philadelphia: John Benjamins.

Shaw, G., & Marlow, N. (1999). The role of student learning styles, gender, attitudes and perceptions on information and communication technology assisted learning. *Computers & education, 33*, 223-234.

Stevens, V. (1995) Concordancing with language learners: Why? When? What? *CAELL Journal* vol. 6 no. 2 pp. 2-10.

Tribble, C. (1997) Improvising corpora for ELT: Quick and dirty ways of developing corpora for language teaching. Available at: http://web.bham.ac.uk/johnstf/ palc.htm.

Tribble, C. (2000). Practical uses for language corpora in ELT. In P. Brett & G. Motteram (Eds.), *A special interest in computers* (pp. 31-41). Whistable, Kent: IATFEL.

Watson Todd, R. (2001). Building and using your own corpus and concordance. *ThaiTESOL buletin vol. 14 no. 2.*

Wichmann, A. (1995) Using concordances for the teaching of modern languages in higher education. *Language Learning Journal* no. 11 pp. 61-63.

Woolard, G. (2000) Collocation - encouraging learner independence. In Lewis, M. (Eds.), *Teaching Collocation: Further Developments in the Lexical Approach.* pp. 28-46. Hove: Language Teaching Publications.