



## BUILDING UP A LEARNER CORPUS THROUGH CREATIVE NONFICTION

### PROSE: AN EXPERIMENTAL RESEARCH

#### KURGUSAL OLMAYAN YARATICI DÜZYAZI YOLUYLA BİR ÖĞRENCİ DERLEMİ OLUŞTURMA: DENEYSSEL BİR ÇALIŞMA

**Mehmet ALTAY**

University of Gaziantep Higher School of Foreign  
Languages,  
[mehmetdeudilforum@gmail.com](mailto:mehmetdeudilforum@gmail.com)

**Assist. Prof. Dr. Filiz Y.TILFARLIOĞLU**

University of Gaziantep English Language Teaching  
Department  
[fyalcin@gantep.edu.tr](mailto:fyalcin@gantep.edu.tr)

#### Abstract

*Corpus linguistics* has become a crucial methodology in linguistic studies in the last half of the 20th century and especially two types of it are frequently applied in ELT: (1) *reference corpus* which does not serve for a particular language variety, yet it is for general purposes, (2) *learner corpus* which is the output obtained from learners (Baker, Hardie, and McEnery, 2006). The main objective of this study is to investigate if a learner corpus which accumulates the texts written in creative nonfiction prose (CN) has a relatively better contribution to EFL students' vocabulary knowledge when compared with reference corpora. The samples of this research are 95 pre-intermediate level prep students from four different classes who take basic English education at Higher School of Foreign Languages Department of the University of Gaziantep. Two of these classes were randomly selected as experimental groups: one treated with a learner corpus while the other with a reference corpus. The rest two classes were the control groups. Analysing the results of the posttest which was given right after the treatments using Independent Samples t-Test, it was seen that learner corpus has a better contribution rather than reference corpora to the vocabulary knowledge of the EFL students.

**Key words:** Learner corpus, Reference corpora, Creative nonfiction prose, Computer mediated communication, Qualitative data, Quantitative data, Empirical research

#### Öz

Derlem dilbilimi, 20. Yüzyılın ikinci yarısı ile birlikte çok önemli bir yöntem bilim haline gelmiştir ve özellikle iki tür İngiliz Dili Eğitimi'nde sıkça uygulanmaktadır: (1) belirli bir dil kullanım türünden ziyade genel amaçlar için kullanılabilen *genel derlem*, (2) öğrencilerden elde edilen verilerle oluşturulan *öğrenci derlemi*. Bu çalışmanın ana amacı, yabancı dil olarak İngilizce öğrenen öğrencilerin sözcük bilgilerine, kurgusal olmayan yaratıcı düzyazı yoluyla yazılan metinlerin bir araya getirilmesiyle oluşturulan bir öğrenci derleminin, genel derleme nisbeten daha iyi katkıda bulunup bulunmadığını incelemektir. Çalışmanın örneklemini, Gaziantep Üniversitesi Yabancı Diller Yüksekokulu'nda Temel İngilizce eğitimi alan dört sınıftan alt-orta seviyedeki 95 hazırlık öğrencisi oluşturmaktadır. Öntest sonuçlarının bu sınıflar arasında anlamlı bir fark olmadığını göstermesinin ardından, sınıflardan iki tanesi rasgele şekilde deney grupları olarak seçilmiştir: birisine öğrenci derlemi uygulaması yapılırken diğerine genel derlem uygulanmıştır. Diğer iki grup ise kontrol grupları olarak tayin edilmişlerdir. Uygulamalardan sonra verilen son testin sonuçlarının Bağımsız Örneklem t-Testi ile incelenmesiyle, genel derleme oranla öğrenci derleminin yabancı dil olarak İngilizce öğrenen öğrencilerin sözcük bilgilerine kayda değer bir katkısının olduğu görülmüştür.

**Anahtar kelimeler:** Öğrenci derlemi, Genel derlemler, Kurgusal olmayan yaratıcı düzyazı, Bilgisayar tabanlı iletişim, Nitel veri, Nicel veri, Deneysel araştırma

## 1. INTRODUCTION

As the advancement in computer and internet technology continues to expand with an incredible pace, English language teaching (ELT) is keeping its track. Today, the use of technology is a sine qua non of ELT classrooms and governments put significant budgets aside to keep its pace. Since many distinguished publishers in the market are aware of this fact, they aim to provide up-to-date materials to their customers which integrate technology into ELT classroom environment.

Keeping ELT classroom up-to-date may sound reasonable. However, not all of the English language teaching materials require an extraordinary budget and thus indulgence to publishers. As an example, corpus provides opportunities for language learners and teachers to see or hear active use of the target language, that is to say it provides examples from real life. Moreover, most of the ELT publishers use corpus themselves while producing materials such as dictionaries, course books and CDs. Therefore, there are logical reasons to integrate corpus into ELT curriculum as a supplementary resource.

A corpus can be in written or spoken form (if not both) and it has more than a dozen of type. It can be used in many fields as well as in ELT. Most of these types are based on real-life examples and these examples are obtained from various resources such as newspapers, magazines, broadcasts and books. Therefore, a corpus may be domain-specific or general (Charteris, 2004). In most general corpora, the examples are obtained from native speakers of the target language, yet the examples from non-native speakers are also very common, especially in domain-specific corpora.

When the issue is English language teaching, learner corpus is of high essence because it provides learners to see examples from learners. In other words, a learner corpus of English language is mostly based on the examples from non-native speakers, which often includes a great many of erroneous language use. Şanal (2007:77) states, “the reference corpus is much more complex in terms of lexical diversity and density than the learner corpus”. Considering, however, the fact that “both in Europe as well as in the world as a whole, English is now a language that is mainly used by bi- and multilinguals, and that its (often monolingual) native speakers are a minority” (Seidlhofer, 2007: 138), language errors or “non-native” language use is as precious as pure native language use no matter how diverse it is.

There are several corpus softwares which are far more economic when compared with most language teaching materials, or free of charge at all. It means that one can build his own corpus or a teacher can collect texts from learners in order to use it in such softwares. What is of high importance in this process is that the texts obtained from learners and gathered in a software should be relevant to the needs and the objectives of the class. They may be determined through a needs analysis, observation or past experiences. While determining the subjects of the texts, it should be considered

that most humanistic education systems such as Waldorf Education, Andragogy, and Missouri Education are highly based on experiential learning. In such systems, teachers are as facilitators, rather than controllers. Since corpus studies are chiefly a tool used to facilitate learning a language as a supplementary way, they should initially be studied in detail, and then adapted considering the objectives.

### 1.1. Corpus Linguistics

Corpus linguistics has become a crucial field in linguistic studies in the last half of the 20th century. The word *corpus* means body in Latin (plural corpora). In linguistics, a corpus is a collection of texts (a body of language) stored in an electronic database (Baker, Hardie, and McEnery, 2006: 48). These texts are generally natural texts accumulated to exemplify the reality of either general or a specific point of language. Charteris (2004) defines a corpus as:

“...any large collection of texts that arise from a natural language use; in a linguistic context, it is in contrast to other types of text that were invented specifically for illustrating a point about language”. (30)

Here it can be understood that corpora are derived from real life examples and the texts in them are not artificial that emerged just for the sake of taking their places in corpus. In contrast, they are a collection of reality themselves. Charteris (2004) supports this idea by saying:

“The notion of attested language is very important in corpus linguistics and implies that data are not invented for the benefit of a model but rather that the model emerges from large and representative samples of language. Other than this, there are no constraints on corpus composition nor are there any constraints on corpus size; these are determined by our purposes in designing the corpus in the first place.” (30)

As it is understood, corpora do not have certain constraints; depending on the purpose beyond its composition, its restrictions are evaluated. It can be concise and “contain a handful of classroom transcripts, interviews or plays” (Weisser, 2006) as well as a comprehensive one such as Brown Corpus, British National Corpus or Birmingham Corpus each of which contains over a million of words. The distinctive feature of corpus linguistics is that it is regarded as a *methodology* rather than a linguistic area of study. Supporting this fact, McEnery and Wilson (2001: 2) comes up with this explanation: “Corpus linguistics is a methodology that may be used in almost any area of linguistics, but it does not truly delimit an area of linguistics itself”.

As stated by Schmied (1996):

“It began in the English as a Native Language (ENL) context, and then, in the EFL context (and its modern expansion the EIL context), the data-based approach has been used since the heyday of error and contrastive analyses in the 1960s.” (182)

It has a key role in building dictionaries, course books, syllabuses and almost all resources in the EFL field. The methodology of European studies in cognitive linguistics in particular has tended to be more corpus-based than the early American studies, which were predominantly introspective

(Grondelaers, Geeraerts, and Speelman, 2007: 1), though several studies carried out in the States have overwhelmed this case since the start of 2000s. Not only the European side and the American side of the World are the followers of this relatively new case, but there are also numerous studies having been carried out in the Eastern World. Thus, corpus studies have become a universal part of linguistics.

Contrary to the common point of view, one of the first samples of corpus even dates back to 1350s with Amarna letters. They form diplomatic archive of correspondance between the Egyptian administration, and its representatives written in Akkadian ([http://en.wikipedia.org/wiki/Text\\_corpus](http://en.wikipedia.org/wiki/Text_corpus)). Since then, various forms of corpus have been established in order to accumulate relevant data to certain subjects in numerous fields, whereas the seeds of corpus in the linguistic studies were sown in 1960 by Randolph Quirk and his colleagues with their project called “the Survey of English Usage”, comprising one million words being used in everyday life. Later on, Henry Kucera and Nelson Francis arranged their works “Brown Corpus” in 1964 and “Computational Analysis of Present-Day American English” in 1967, which are known as the milestones of corpus linguistics (Baker, Hardie, and McEnery, 2006: 50).

Several reasons can be aligned for the importance of corpus studies, but most probably the main reason for their being the milestones in the field of linguistics is that they provide both general and specific examples from real language depending on the purpose. These examples are of high importance for language education as learners and teachers seem highly dependent for companies which provide materials; thus, the borders are limited by these companies as long as other supportive aids are not added into education. There are, of course, alternatives to delimit these borders. As a supportive aid to overcome this problem, Baik and Shim (2002) worked on the World Wide Web. Their study was to bring an awareness and understanding of the existence of varieties of English, and they mention the situation as stated below:

“Being without the support of major publishers or centralized organization for the development of such materials, there have been various practical difficulties in developing and supplying materials. We have had to depend on odds and ends collections of written texts (mostly newspaper articles) obtained through personal contacts, and the most serious limitation of such materials was undeniably the fact that there was a scarcity of audio materials.”(1)

It is crystal clear that there are other ways to be consulted and alternatives for these distinguished publishers and organizations in the EFL market, and corpus-based resources are among them. Such resources can be accessed online and/or offline, and they can even be built on one’s own. They may also be adapted as supplementary aids to existing curriculum.

There are more than a dozen types of corpus applicable for EFL contexts. Among these types, two types of it come forward in ELT: (1) *reference corpus* which does not serve for a particular

language variety, yet it is for general purposes, and (2) *learner corpus* which is the output obtained from learners (Baker, Hardie, and McEnery, 2006). Reference corpus may provide a great many of examples for the learners from the target language and these examples are generally taken from resources based on native use of language. On the other hand, a learner corpus supplies examples from learners as the data in it is based on the products of learners themselves. All in all, corpus has a rising and up-to-date value for EFL teaching with regardless of its types. The key point is choosing the appropriate type for the needs and the purposes of the context. Once the needle of a study points to a learner corpus as an appropriate method to be adapted into curriculum, the crucial requirement to be considered is to determine a suitable genre for the texts to be used selectively.

## **1.2. Creative nonfiction prose**

Although almost all people who either enjoys writing or reading have most probably been exposed to this genre of nonfiction, it is not nominally very much popular, at least among the ones who do not have a strong aim to follow the trends in literary world. However, creative nonfiction is not actually an infant in the field, though it used to be like an abandoned baby without a name in the yard of the literature at first. Nowadays, creative nonfiction prose is getting widespread among literature followers all around the world.

There have long been efforts on writing in creative nonfiction even by eminent authors such as Orwell, Baldwin, Hemingway, Tom Wolfe and many others, though most probably they have not intended to take part in creative nonfiction world. The name “creative nonfiction” has a really interesting story. The father of the name, Lee Gutkind, claims that he has been using this genre since 1970’s, yet it became known as creative nonfiction after a meeting of National Endowment for the Arts in 1983. It is also known as “literary nonfiction” and “narrative nonfiction” as it has strong ties with narration and literary language use. This genre neither fits to “essay”, nor to “journalism” and its lines have sometimes been misinterpreted.

In fact, the core elements that constructs creative nonfiction is hidden within its name; “creativity” and “nonfiction”. The writers transmit factual information to their readers through using a creative language which is often literal. Fox and Lannin (2007: 2) describe creative nonfiction as “a hybrid genre that pulls in elements of fiction (literary techniques), the writer’s perspective, and factual information”.

Before going deep into creative nonfiction, it would be better to distinguish nonfiction from fiction. The term fiction is briefly for something created by human mind (McArthur, 1992: 401). Literally, it is connected with imagination even though it is not totally out of reality. Therefore, nonfiction does not have much, if not any, to do with imagination of the human mind. Instead, it focuses on factual information.

As Kenan (2008: 3) classifies, some examples to nonfiction are gossip, legal testimony, news reports, history books, autobiography, personal letters and so on. Of course the reliability of some of these narratives might be questioned, yet they are still resources based on factuality, or at least resources seeking for factuality.

Creativity is undoubtedly more common in fiction rather than nonfiction. Figurative language including metaphor and simile; and literal meanings of words can frequently be seen in such narratives as the imagination is in the center. However, neither creativity, nor narration solely belongs to fiction. As the quotation from Lewin (1997) was emphasized by Genette (2000: 91), “Narrative is a doubly temporal sequence... There is the time of the thing told and the time of the narrative (the time of the signified and the time of the signifier)”. A narrative may be newspaper reports, history books, novels, films, comic strips, pantomime, dance, gossip, psychoanalytic sessions and many others (Kenan, 2008: 1). When compared, there are overlapping examples from both narratives and nonfictions.

What helps creative nonfiction emerge might be the instinct in human that triggers the joy, vividness and all other emotions which belong to humans and the need to this instinct in getting or transmitting information. In other words, there is no doubt that the more the creativity is hidden in a message, most probably the better it takes place in the minds of its receivers. Gutkind (2005) explains the goal of creative nonfiction writers as in the lines below:

“Although it sounds a bit affected and presumptuous, ‘creative nonfiction’ precisely describes what the form is all about. The word ‘creative’ refers simply to the use of literary craft in presenting ‘nonfiction’ that is, factually accurate prose about real people and events in a compelling, vivid manner. To put it another way, creative nonfiction writers do not make things up; they make ideas and information that already exist more interesting and, often, more accessible.”

As previously mentioned, creative nonfiction prose (CNP) is a hybrid genre; it works on information which is the core element of nonfiction by using the tools of fiction. It cannot be regarded as a traditional nonfiction just like journalism or scholarship. It offers flexibility and freedom while adhering to the basic tenets of reportage (Gutkind, 2005). Creative nonfiction writers can be poetic as well as journalistic, can use literary techniques while handling a factual issue, can describe things by adding their point of view and so on.

A creative nonfiction can be a memoir written on a journey or it can be a diary and many other examples containing experiences from real life as long as it is written in a literal (or creative) way. Boundaries are limited with the factual information provided by the writer. It may be used as an effective tool in almost all sciences as it leads learners to state themselves or their real life experiences in a creative manner. It is also known as *narrative nonfiction* and *literary nonfiction* since it has common points with narration and literary usages, and it leads the learners to stimulate their productivity through narration. In their study which adopted a similar method called “productive



narratives”, Sinclair and Healy (2007: 20) worked on narrative thinking with dynamic geometry and highlighted that productive narratives, in terms of mathematical sensemaking, are those in which learners are able to connect mathematical objects, and their paradigmatic relationships and properties, with things they already know- and care about, stories in which the mathematical is given meaning through its grounding in experienced phenomena. Therefore, methods which allow for the adaptation of narration into the education procedure may contribute to the success. Such methods including the adaptation of creative nonfiction may help learners to deal with anything happening so it contributes to the variation of the registers in their writings.

Creative nonfiction can also be used as a consciousness-raising activity in education and it assists the learning process. Gilquin and Paquot (2007: 10) applied a consciousness-raising activity in their academic survey and verify the existence of these contributions of them by saying that consciousness-raising activities have also been proposed that should help learners become more aware of register variation and of the importance of adopting a stylistically appropriate tone in academic writing.

It can be applied to students to help them “use their imaginations to tell the truth” as done in the study by Renker (1998), can be a tool to let the students forget the class discrimination and let the “less fortunate people” express themselves in autobiographies as done by Papay (2003), can facilitate the hardship of literal study as done by Skrebels (2003), or it can even help a teacher to interact with students on how to cope with a resistant student as done by Robinson (1997).

To sum up, creative nonfiction can be applied in every classroom environment by tailoring it to the needs and purposes. It is a consciousness-raising activity by enriching the registers as well as by contributing to stylistic skills. It also provides an opportunity for the students to express themselves and this situation most probably reinforces the interaction of the individuals in an educational environment. Therefore, creative nonfiction may work well as a genre of the texts which are to be used in a learner corpus for vocabulary teaching.

### **1.3. Purpose of the study**

Reference corpus and learner corpus are the two of many types of corpus which are often integrated in EFL context as a supplementary methodology. Considering the objectives of an EFL class, each may have some pros and cons. On the one hand, a reference corpus serve for general interest, that is it does not necessarily focus on a specific field. Moreover, it generally includes pieces of native language use, which are rarely erroneous. On the other hand, a learner corpus may include erroneous language uses as it is built through obtaining texts from foreign language learners. In other words, errors are welcomed and an important component of a learner corpus as they inform the learners about the mistakes as well as the correct usages.

As claimed by Baker, Hardie and McEnery, most learner corpora consist of written essays using pre-set topics produced in language-teaching classrooms (2006: 103); the topic on which the learners will write is often determined in the classroom and once the learners write on that topic, the texts are gathered to be used in the learner corpus. Therefore, the selection of topics used in a learner corpus is of high importance in an EFL context because productivity is an indispensable figure here as well as the reception. In this current study, creative nonfiction prose is used in the learner corpus as this type of prose requires both the productive skills of the learners such as while using figurative language use, metaphors and similis, and the receptive skills indicated in a variety of circumstances such as self-consciousness, environmental awareness and critical thinking.

The overall objective of this study is to investigate the vocabulary proficiency of EFL learners who are required to build a learner corpus based on creative nonfiction prose, and those who followed reference corpora. Whether a methodology which is based not only on reception but also on productivity helps EFL learners better than a methodology which is chiefly based on reception in lexical proficiency is also investigated. One further aim of this study is to reveal if studying vocabulary through using corpus, no matter which type of it is used, is a contributive methodology for an EFL class.

## **2. METHODOLOGY**

### **Participants**

The target population of the research was the students of the Higher School of Foreign Languages of the University of Gaziantep in 2008-2009 education year. As Ekmekçi (1997: 21) states, when the accessible population is found to be too broad; then, the researcher chooses a sample. Here, the researcher used cluster random sampling in order to minimize the population. The total number of the students in that education year was 1090 including the night class students. In the Higher School of Foreign Languages, students were replaced into classes depending on their proficiency levels in English. The students who were determined to be at elementary level at the beginning of the education year were placed to Level C classes while the students of pre-intermediate level were to Level B classes. Additionally, the students who were accepted to have intermediate proficiency level were replaced to Level A classes.

All three groups were aimed to reach to upper-intermediate level at the end of the education year. To conduct this research, Level B students were chosen as the accessible population; because, they were in the strata as a group in accordance with the replacement test. The number of the accessible population in this research was 422 students with 296 morning class students and 126 night class students. The number of the Level B classes was 12 morning classes and 5 night classes; 17 in total. Out of these classes, four groups from Level B were picked up for the study randomly.



The first of the four groups was a night class group namely E2 and it consists of 25 Level B students. This group was selected to participate in the research as one of the experiment groups and they were asked to follow the general corpora after the pre-test. The second group was another night class group namely C1 and it consists of 25 Level B students as well. Whereas, this group was one of the control groups and they only participated in the pre-test and post-test taking to evaluate the effects of the independent variables over the dependent variables. The next group was another night class group called C2 which has 24 Level B students. The group was the other of the control groups and followed the same process as for C1. The fourth and the last group was E1 with 21 Level B students. This group was the other experiment group and the participants were asked to take part in a learner corpus procedure. In total, 95 Level B students participated to this pre-test/post-test research study. All of the students has taken the National University Entrance Examination (ÖSS) before applying for their departments in the university. Following the ÖSS, they took an exemption test prepared by the Higher School of Foreign Languages aiming to determine the students to be excluded from taking a one-year English language class before going on in their actual departments, and they could not pass this exam. As a final stage to classify the students depending on their English language proficiency levels, the students participated in a proficiency test and those 95 students were leaded to the Level B classes.

#### **Data Collection Instrument**

A lexicon of more than 700 words was accumulated from the coursebooks of the target population. Using the vocabulary items of it, a vocabulary test was prepared and applied to the samples of four groups as pre-test and post-test. The test was made up of 40 multiple choice questions each with 5 options. Almost all the new vocabulary items were used only once as an option in the test.

During the preparation of the test, the researcher aimed to compose the questions using Collins COBUILD Resource Pack – Lexicon (2002) software which includes a corpus of five million words of texts from the Bank of English. Additionally, Cambridge Advanced Learner's Dictionary software was used as an assistant instrument. After picking up suitable examples from these softwares so as to compose the questions, the vocabulary items were carefully blended and inserted to the questions as options depending on their word forms. For instance, if a question seeks for an adjective, all the options to the question were initially aimed to be adjectives but not from other word groups. This was done so in order to reduce the probability of the students' finding the correct answer through guessing to which word form an item belongs.

On applying the pre-test, the reliability results were analysed so as to determine if the given test is appropriate as a data collection instrument for the research. Since Krombach's Alpha is accepted as one of the most consistent way of obtaining the reliability results, the researcher has chosen this type of reliability analysis for the pre-test. Therefore, the reliability is ,070 in accordance

with Cronbach's Alpha reliability statistics, which means that the test used as the instrument to collect the data is reliable for the research.

### Procedure

95 pre-intermediate level students from four different classes were required to take the vocabulary exam as the pre-test. Having seen that the mean scores were not significantly different among these groups (see Table 2), the groups were assigned randomly as experiment and control groups.

E1 is the first experiment group which was assigned to conduct a learner corpus study of 12 treatments. Meanwhile, they were provided a number of vocabulary items from the lexicon and asked to write texts based on creative nonfiction prose so that the researcher collected them after each treatment and analyse in the classroom with the students using WordSmith 5 corpus concordancer software. E2 is the other experiment group and they were required to follow a reference corpora study. A number of well-known online reference corpora were provided for them. An orientation on how to use these corpora was carried out in the language lab. Later on, the same vocabulary items related to that treatment was provided for them from the lexicon as well. Thus, they maintained their reference corpora in the language lab.

The other two groups are C1 and C2, which are determined as the control groups and had no tasks other than merely following the curriculum of their department. Once the treatments to the experiment groups were over, all of the four groups took the post-test, which is the same of the pre-test. The research design is tabulated as:

Table 1: The research design

Step	E1 (Experimental)	C1 (Control)	E2 (Experimental)	C2 (Control)
1	Cluster assignment	Cluster assignment	Cluster assignment	Cluster assignment
2	Pre-test	Pre-test	Pre-test	Pre-test
3	Treatment	---	Treatment	---
4	Post-test	Post-test	Post-test	Post-test

### Data analyses

Findings were analysed using SPSS 15.0 and Independent – Samples T Test. The pre-test results were compared and it was seen that there were not any significant difference between the groups. After the experiments, the result of the first experimental group was compared with that of the

first control group, and the result of the second experimental group with that of the second control group.

### 3. RESULTS

In this section, the quantitative data obtained from the pre-test and post test are analysed. The indications of the statistical analyses are as the followings:

Table 2: The results of the pre-test and the post-test

Tests	E2			C2			C1			E1		
	N	Mean	sd	N	Mean	sd	N	Mean	sd	N	Mean	sd
Pre-test	25	8,72	2,28	25	7,76	1,5	24	7,83	1,85	21	8,04	2,47
Post-test	25	21,88	3,12	24	13,50	2,91	25	13,76	4,85	21	33,80	1,96

As seen in the Table 2, the pre-test results of the four groups are not significantly different from one another with regard to their mean scores and standard deviations. However, the post-test results indicate that the mean of E1 (the first experimental group) overwhelms those of all other groups. On the other hand, the mean of E2 (the second experimental group) is also higher than the two control groups, but less than E1. As both E1 and E2 have significantly higher mean scores when compared to the scores of the control groups, these results indicate the significance of corpus studies in an EFL class as a supplementary methodology. More comprehensive results are given in the tables below:

Table 3: Post-test mean scores, standard deviation, t and p values for the experimental group E1 and the control group C1

GROUPS	N	$\bar{X}$	sd	t	p
C1	25	13,76	4,85	,18	,01
E1	21	33,80	1,96		

A huge gap between the mean scores of the groups C1 and E1 to the posttest is evident in Table 3. In addition, the significant value also indicates that there is a significant difference between the scores of group C1 and group E1 ( $t = ,18$ ;  $p > ,05$ ). In conclusion, it can be said that the treatment applied to the experimental group E1 must have a strong impact on this group.

Table 4: Post-test mean scores, standard deviation, t and p values for the experimental group E2 and the control group C2

GROUPS	N	$\bar{X}$	sd	t	p
E2	25	21,88	3,12	9,07	,24
C2	24	13,50	2,91		

Table 4 shows the statistical analysis of the post-test results based on the scores of the groups E2 and C2. It can be seen that there is a variation between the mean scores of the groups; however, the significant value points out no significant difference between these two groups ( $t = 9,07$ ;  $p > ,05$ ). Although the mean score of the experimental group seems to be fine, the standard deviation must have played an important role in the results, which leads to an insignificant value.

#### 4. CONCLUSION

Considering the results obtained from the analyses of the post test, the followings were concluded with respect to the hypotheses:

1. There is a significant difference between the proficiency levels of the learners who were treated by building up a learner corpus through using creative nonfiction prose and the learners who were required to follow reference corpora during the treatment period. It was concluded by comparing the significant values of both groups to the post-test results. The significant value obtained by comparing the second experimental group E2 with the second control group indicated no significant difference ( $p \leq ,24$ ) while the significant value obtained by comparing the first experimental group E1 with the first control group C1 shows a significant difference ( $p \leq ,01$ ). Therefore, the comparisons of the experiment groups with the control groups also depicts a significant difference between the experiment groups. There may be some underlying reasons for such a difference.

First of all, building up a learner corpus requires not only the competence, but also the performance of the learners. In other words, active use of the vocabulary items in real life seems to have a positive effect upon the proficiency levels. As previously mentioned, Şanal (2007) states that reference corpora provide lexical diversity, but if the learners need to learn a diverse lexicon, or they need what the course objectives require in educational contexts is the real question. Additionally (Şanal, 2007), it is possible to determine tasks and/or subjects to be handled in learner corpus studies. On the other hand, reference corpora are basically towards analysing words and their patterns, which leads to a lack of using productive skills.

Secondly, the use of reference corpora has long been criticized for missing authenticity. In this perspective, James (1992: 190) argues that the really authentic texts for foreign language learning are not those produced by native speakers for native speakers, but those produced by learners themselves. Learner corpus do not only indicates the correct use of language, but also shows the incorrect use so the learners can also observe what cannot be accepted to be correct. Therefore, learner awareness to the incorrect use of language increases.

Third, peer observation might have a positive effect on the learners and a collaborative atmosphere has been embodied in learner corpus group. Additionally, CNF as a productive prose type might have an influence over the learner corpus; and thus, the participants.

2. There is a significant difference between the proficiency levels of the learners trained through using learner corpus and the learners who merely follow the curriculum depending on the significant value of the post-test ( $p \leq ,01$ ). Corpora studies provide real life examples; therefore, a more realistic language learning environment is inevitable. However, not all curriculum works in this sense. Additionally, learner corpora brings more natural language use; and thus, they stimulate productivity as well as the competence.

3. There is not a significant difference between the proficiency levels of the students who were treated using reference corpora and the students who only followed the curriculum by means of their vocabulary knowledge. Although they seem to be weak when compared with a well-built learner corpus, reference corpora still have a contributive effect on vocabulary teaching. However, it is not a significant effect as seen after the post-test ( $p \leq ,24$ ). No matter how much authentic they are, it can be said that they provide real life language use, but providing such a language use seems not to be solely enough for a significant change in the vocabulary knowledge of the learners. Therefore, it can just be said that reference corpora study has a contribution, though it might be superficial as the mean scores of the post test show.

## REFERENCES

- Baik, M. J., & Shim, R. J. (2002). Teaching world Englishes via the Internet. *Symposium on World Englishes and Teaching English as a Foreign Language*. Seoul: Blackwell Publishing.
- Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press Ltd., pp.48-49-103.
- Charteris, B. J. (2004). *Corpus Approaches to Critical Metaphor Analysis*. Hants: Palgrave-MacMillan, p.32.
- Fox, R. F., & Lannin, A. A. (2007). *Belly Up To The Pond: Teaching Teachers Creative Nonfiction In An Online Class*, p.2. <http://www.writinginstructor.com/foxlamin> (29.09.2009).
- Genette, G. (2000). *Narrative Discourse Revisited*. New York: Cornell University Press, p.91.

- Grondelaers, S., Geeraerts, D., & Speelman, D. (2007). A case for a Cognitive corpus Linguistics. In M. G. Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey, *Methods i Cognitiv e Linguistics*. John Benjamins Publishing Company, p.149.
- Kenan, S. R. (2008). *Narrative Fiction*. New York: Routledge, pp.3-10. Gutkind, L. (2009). What is Creative Nonfiction. <http://www.creativenonfiction.org/> (19.09.2009).
- McArthur, T. (1992). *The Oxford Companion to the English Language*. Oxford, New York: Oxford, pp. 230-216-401.
- Papay, T. Y. (2003). In the Margins of the Privileged Text: Deconstructing Creative Nonfiction's Pedagogy of Entitlement. [www.eric.ed.gov](http://www.eric.ed.gov) 14.03.2010).
- Renker, F. (1998). Authority and imagined truth: notes on teachicreative nonfiction. *Paper presented at the 49th Annual Meeting of the Conference on College Composition and Communication*. Chicago.
- Schmied, J. (1996). Second-Language Corpora. In S. Greenbaum, *Comparing English Worldwide: The International Corpus of English*. Clarendon Press, p.182.
- Skrebels, P. (2003). Creative nonfiction in the classroom: Extending the boundaries of literary study. [http://education.waikato.ac.nz/journal/english\\_journal/uploads/files/2003v2n2art7.pdf](http://education.waikato.ac.nz/journal/english_journal/uploads/files/2003v2n2art7.pdf) (23.01.2010).
- Seidlhofer, B. (2007). Common Property: English as a Lingua Franca in Europe. [www.springerlink.com](http://www.springerlink.com) (08.02.2010).
- Sinclair, N., & Healy, L. (2009). *Time for telling stories: narrative thinking with dynamic*. Karlsruhe, p.20.
- Şanal, F. (2007). *A Learner Corpus Based Study on Second Language Lexicology of Turkish Students of English*. Thesis (PhD) Çukurova University.
- Weisser, M. (2005). What is a Corpus. <http://ell.phil.tu-chemnitz.de/compPhil/corpus.html> (13.05.2010).
- Wikipedia. (2010). Text Corpus. [http://en.wikipedia.org/wiki/Text\\_corpus](http://en.wikipedia.org/wiki/Text_corpus) (20.01.2011).