Impact of Text-Mining and Imitating Strategies on Lexical Richness, Lexical Diversity and General Success in Second Language Writing

Sevcan Bayraktar Çepni
YıldırımBeyazıt University, Turkey
sbayraktar@ybu.edu.tr

Elif Tokdemir Demirel
Karadeniz Technical University, Turkey
elif@ktu.edu.tr

ABSTRACT
This study aimed to find out the impact of "text mining and imitating" strategies on lexical richness, lexical diversity and general success of students in their compositions in second language writing. The participants were 98 students studying their first year in Karadeniz Technical University in English Language and Literature department. Participants in the experimental group were instructed on how to use mining and imitating strategies in their essays to be more competent in vocabulary choice while writing. On the other hand, students in the control group were only instructed on how to write their essays following lesson plans. After treatment, participants in the both group were asked to write two essays. A corpus of 15826 (control group) and a corpus of 25027 (experimental group) words were compiled in two months, and these essays were tagged using a computerized tagging system (Biber 1993). Though there is statistically slight difference in terms of lexical richness, the study revealed that compositions in the experimental group are lexically richer than those in the control group. Common nouns and general verbs subclasses were found to be used more in compositions of the experimental group. In addition to that, students in the experimental group received significantly higher grades than those in the control group.

Keywords: Text mining, imitating, lexical richness, lexical diversity, learner corpora.

INTRODUCTION
Development of vocabulary is a fundamental part of learning a second or a foreign language. Wilkins (1992: 111) declared that "little can be conveyed without grammar but nothing can be conveyed without vocabulary." Therefore, the essential need for vocabulary learning in the process of second or foreign language has gained increasing recognition in research agenda. To address this need, many studies on vocabulary learning have been conducted so far (Nassaji, H. 2006; Lee, S. H. & Muncie, J. 2006; Nation, I. S. P. 2008). These studies revealed mutual relationship between the knowledge of vocabulary and other language skills. In respect to the importance of vocabulary as a fundamental skill underlying other skills, this paper presents the relationship between vocabulary knowledge developed through reading and writing achievement in specific.

SIGNIFICANCE OF THE STUDY
The importance of competence in writing has been explicitly elaborated by many researchers. Weigle (2002: 1), for example, wrote that "The ability to write effectively is becoming increasingly important in our global community, and instruction in writing is thus assuming an increasing role in both second- and foreign- language education". To write effectively, students need to use their vocabulary knowledge properly. However, as vocabulary is generally learnt through lists and memorization, students can only achieve short term memorizations and cannot use these words properly and actively in the long term use of the language learning process. In other words, learners have difficulties in adapting their vocabulary knowledge into writing or speaking skills.

In line with what Weigle (2002) said, Gorell (1987: 53) stated "unskilled writers do not have a clear sense of form. The sense for them is all mixed up with jumbled, half-remembered rules, unsuccessful trials at writing, heavily marked papers, and insufficient and ineffectual reading". The question of what increases learner competence in the use of vocabulary for effective writing has been the focus of much research so far. As writing in a second language is different from writing in the first language, most of the research conducted on impact of L1 reading on L1writing production fails to answer questions in EFL research. It has been put that there is a

positive correlation between reading and writing in the target language; however, the question of which reading strategies for writing should be used to facilitate vocabulary use in written productions has not yet been answered completely. In order to understand how reading informs writing in terms of vocabulary and what happens when readers are also writers, students should be asked to use certain strategies called "text mining" and "imitating" to reach a conclusion about the usefulness of these strategies on vocabulary knowledge. These questions are important as they can lead to a shift in emphasis from teaching of reading and writing to the nature of learning how students use and adopt these strategies.

LITERATURE REVIEW

The relationship between vocabulary knowledge and writing competence has been the focus of many research studies. (Laufer and Nation 1995, Folse 2006, Barcroft 2007). Schellekens (2007: 103) claimed that "Many students struggle with writing and they often find it hard to write at all, even about topics that they know well, such as their home life, their children, or their job". This difficulty may stem from the fact that students' lexical knowledge is not enough to help them produce detailed expressions while writing in a second or foreign language. Relationship between reading and other skills have been focus of concern for many researchers. The power of meaningful reading to increase learner language proficiency has been highlighted by Krashen (1989: 109) who stated that "reading exposure is the primary means of developing language skills". That is, he (2004: 37-132) claimed that "reading is a powerful means of developing reading comprehension ability, writing style, vocabulary, grammar, and spelling" ; we acquire writing style, the special language of writing, by reading". Focusing on the similarities between reading and writing, Cooper (1993) suggested that reading and writing should be taught together as they are both constructive processes, share similar processes and kinds of knowledge, improve achievement, foster communication, and lead to outcomes not attributable to either process alone. Similar to Krashen's and Cooper's ideas on impact of reading on writing, some L2 writing researchers such as Grabe (1997) and Paradi (2006), who highlight the relationship between reading and writing claim that L2 learners' writing ability can be facilitated by encouraging students to use model essays as they present rhetorical elements, principles and patterns of written discourse. According to such views, reading and writing are connected as they depend on similar knowledge representations, cognitive processes, contexts and contextual constraints. Therefore, it is not unusual to think that reading and writing development has close interaction, which may lead some pedagogical combinations that may be useful in making learning more efficient.

Following Krashen's theory that holds writing skill is affected positively by extensive reading, a number of studies on reading and writing relationships in L2 have been conducted. Ito (2011) investigated relationship between L2 reading skills and persuasive essay quality in EFL Japanese high schools students. It was found that L2 reading has an effect on L2 writing quality. Özçelik (1996) sought an answer to the question whether the writing process of low level EFL students improves if it is taught through reading with the help of reading texts. Participants of the study were 20 low level prep school students. He divided these participants into two and made one group an the experimental group and one group of a the control group. Participants in the experimental group were exposed to pre-writing activities through reading with the help of reading texts but the participants in the control group was not exposed to pre-writing activities during the study. He found that reading had an effect on writing. That is teaching writing through reading text produced a significant increase in the composition profile total score of learners. Shanahan (1984) tried to find out the relationship between reading and writing by examining second and fifth graders phonic skills, reading comprehension, reading vocabulary, spelling, and prose writing. She analyzed writing samples for syntactic complexity, diversity of vocabulary, and organizational structure. She found that reading and writing measures were positively correlated, but only to a small or moderate extent. Shanahan concluded that the association between diversity of vocabulary in writing, and reading ability had an increase from second to fifth grade.

Some studies endeavored to document how reading informs writing in terms of vocabulary usage. Elgord and Warren (2014), for example, investigated acquisition of second language (L2) vocabulary from reading a connected authentic text. The study revealed that number of encounters with new words in reading helped learners gain explicit word knowledge. However, advanced learners and those with lower proficiencies differ in gaining such knowledge. While extensive reading may be sufficient to sustain vocabulary development for advanced learners, lower proficiency reading needs to be supplemented with deliberate word learning and vocabulary learning strategy training.

It is believed that students can learn about writing by "imitating" good models of written discourse. In such an approach, students are expected to internalize the style, grace, and correctness that make these works exemplary. Gorell (1987: 54) claimed that unskilled writers learn from imitation by focusing on form and structure while generating and finding the expression for their own ideas. By imitating, they learn to shape their sentences, develop their paragraphs, express their own voices, and perform many of the complicated tasks that writing

process involves. In addition to this, he asserted that "when writers read, they pick up not only meaning but also the way in which that meaning is expressed.". He also (1987) touched upon functions of imitation explaining one as having a problem solving capacity with which students make use of experience- one's own and that of others to find solutions. Applied to writing, imitation means students do not need to invent a new form every time they want to express an idea. However, role of imitation has also brought some speculations as Greene (1991: 152) said:

> One might wonder if students can articulate or apply the discourse knowledge they tacitly learn through imitation to their writing in different situations and across a number of varying tasks. Will imitation serve our students when they must transform their knowledge in order to contribute something new to an ongoing conversation in a given field?

Greene (1991) supported the view that employing this strategy students can only learn "the forms and genres and the ways of speaking that writing is a discipline demands" (Jolliffe and Brier, as cited in Greene 1991: 55). The question as to what extent these approaches teach students— individual writers— to negotiate the complex demands that a rhetorical situation places upon remains unanswered. Accordingly, Greene (1991) claimed that if students are expected to make reasonable choices and decisions in widely different rhetorical situations, abstract instruction will not suffice to them. Therefore, the term pragmatic reading to facilitate writing has been used by Greene (1991: 155) whose metaphor of "text mining" holds that in order to achieve goals in composing, writers should read purposefully and intently. Tsai (2006) defines text mining as a strategic approach that is used to dig out valuable language sources such as grammar and vocabulary. By using this strategy, writing and vocabulary skills of students are expected to improve as students pay attention to grammatical and lexical features of the texts, organization of the texts, and expressions which are unfamiliar to them. In addition to that, this strategy is thought to help students improve their reading skills, and, at the same time, build the foundation of future writing. This kind of pragmatic reading, according to Greene, is fueled by three key strategies. Three keys strategies that can inform reading are "reconstructing context", "inferring or imposing structure", and "seeing choices in language". Green (1991) sees language as a lens through which people can understand something in a particular way. Therefore, even subtle changes in language can change the ways how meaning is located, which requires a process which involves a plan, selective evaluation and organization of information in order to get a sense of the topography. So people can reflect upon one's choices and decisions about the use of this accumulated knowledge to the best effect. Mining process is like an excavation during which miner uses certain tools that are convenient for the situation to help uncover what is most desired. This means that readers who are also writers use these strategies to reconstruct context, infer or impose structure, and see choices in language. This leads readers to make informed guesses about the use of the ideas or discourse features of a given text in light of his or her goals as a writer (Greene, 1991). Mining suggests a strategic process that can be considered to be the key factor in raising student awareness on how discourse patterns organize subset information.

Intensive research on vocabulary acquisition in the process of learning a foreign language has brought the need for tools that may ease investigating large samples of students' written productions. In line with this need, developments in computer text-processing capabilities has made it possible to investigate large samples of learner writing through corpus based research. With the ease of these tools, learner corpus studies have flourished discovering more and more linguistic problems that learners have. This flourishment brought the term "Learner Corpora" in the agenda. Electronic collections of spoken or written texts that are produced by foreign or second language learners are defined as computer learner corpora. Two types of corpora have been focus of SLA research so far: corpora by learners and corpora for learners. Sylviane Granger and her team developed the International Corpus of Learner English (ICLE) in 1998 which resulted in growing interest in producing corpora that can be used to study interlanguage of learners. The general idea behind such work is that if learners' language is identified or analyzed, it may be possible to focus on teaching methods and contents in order to make teaching context more fruitful. Learner corpora can be compiled from both writing and spoken products of learners.

METHODOLOGY

This is an experimental study aiming to find out possible impact of certain reading strategies called "text mining" and "imitating" on lexical richness, diversity, and general success in learner compositions. The participants of this study, 98 students ; 39 in the control group, 59 in the experimental group, were first year students of Karadeniz Technical University in English Language and Literature department in academic year 2012- 2013. The students, who were native speakers of Turkish, were chosen with convenience sampling technique. Even though the participants chosen via convenience sampling might not represent the whole population (Paton, 2002), students who were eager to take part in the study were chosen. After attending a year of preparatory class of intensive English courses: writing, speaking, listening, reading and grammar, students advanced to first year in their departments. In order to show equality of proficiency levels of these students in

writing skill, all students were assigned to write a narrative essay before the treatment. First drafts of their essays were compiled, and randomly selected 20 essays from each group were submitted to two independent raters who are teaching writing classes to be scored by using essay scoring rubric developed by Oshima and Hogue in 2006.

An independent-samples t-test was conducted to compare the scores of compositions of the students before the treatment. The results indicated nonsignificant difference between the control group (M = 68.10, SD = 9.276) and the experimental group (M = 69.88, SD = 10.511), t (96 = -.881, p = .381). The analysis showed that students in both groups had nearly same writing performances prior to the study. In order to make sure that students understood how to apply text mining and imitating strategies, they were checked one by one when analyzing their model essays. If needed, researcher provided them with one by one training on using these strategies.

INSTRUMENT

For this study, participants were assigned to write one descriptive essay entitled "the person you admire" and one advantage disadvantage essay on "employee monitoring". They were asked to write five paragraph essays with at least 250 words.Students in the control group were instructed with power point presentation on how to write a descriptive and advantage and disadvantage essays. They were provided with charts and graphic organizers and instructions on rhetorical style and organization (see appendices 5-12 for sample student essays).

Their essays were compiled and tagged with Biber (1993) tags (see appendix 1). Bennet (2002: 14) stated that:

> When a corpus is tagged, it means that each word included in the corpus has a marker added to it that gives additional information. Often, tags are part of speech markers, enabling users of corpora to search not only for specific words, but also for specific words used as a particular part of speech.

CORPUS DESCRIPTION

Two parallel corpora of learners studying in the first year of English Language and Literature Departments were compiled in expository writing lesson. 39 essays acquired from the control group that constituted a corpus of 15826 tokens while 59 essays acquired from the experimental group which constituted a corpus of 25027 tokens. All texts in student corpus were grammatically annotated using an automatic grammatical "tagger" (a computer program developed and revised over ten year period by Biber1993. A large number of linguistic features in spoken (transcribed) and written forms are identified by this tagger. Tagging this corpus made it possible to conduct a series of more sophisticated analyses than would have been possible with an untagged corpus. Preparing student texts for the program was labor intensive and extremely time consuming. Spelling of each word is edited as not only English characters are required but also spelling of words should be correct. Each part of speech is tagged according to its classes. If a word has two functions, the one that is more commonly used is chosen to be tagged. For example, "name" has both verb function and noun function. As "name" as a noun is more used in dictionaries and corpus of native speakers, the program tags it as a name. Following table shows sentences from tagged texts.

DATA COLLECTION

At the beginning of the term, students in the experimental group were taught certain reading strategies called "text mining" and "imitating". They were trained to work on model essays to gain insights into how and where writers use words where needed. For each class and topic, they were provided with model essays. By reading and mining these essays, they tried to produce their own compositions. On the other hand, students in the control group learnt writing essays with charts and organizers; they were not exposed to reading texts or analyzing texts through text mining. They were only taught how to write an essay, and were asked to write their own essays following guidelines provided by these graphics.The data for the writing were collected via writing tasks covering the topics that were assigned to them. Two parallel corpora of students in the control group and students in the experimental group were compiled and (39 essays in the control group (a corpus of 15826 words); 59 essays (a corpus of 25027 words) in the experimental group) these essays were tagged (with Biber tagger's tag descriptions) in order to ease the analysis.

DATA ANALYSIS

Corpus-based studies often measure vocabulary richness in terms of Type-Token Ratio (TTR) in which the number of different words (types) a learner writes in a text is divided by the total number of words (tokens) in order to determine the degree of variation. However, this traditional model brings out some problems with it. Johansson (2008:63) mentions about this problem as follows:

> A problem with the TTR measure is that text samples containing large numbers of tokens give lower values for TTR and vice versa. The reason for this is that the number of word tokens can increase infinitely, and although the same is true for word types, it is often necessary for the

> writer or speaker to re-use several function words in order to produce one new (lexical) word. This implies that a longer text in general has a lower TTR value than a shorter text, which makes it especially complicated to use TTR in developmental comparisons, e.g., between age-groups, where the number of word tokens often increase with age.

As an alternative to TTR model, some models have been proposed to measure lexical richness "Theoretical *vocabulary model*" is one of these models which has been proposed (Broeder, Extra & van Hout 1986) tomeasure word types in the samples. The principle behind this measure, as Johansson (2008) says, is to pick a number of words (e.g 100 words) from a text randomly and calculate the number of word types in the sample. Therefore, theoretical vocabulary takes into account all possible ways of choosing 100 words from the text. By doing this, one can easily compare texts of different lengths with limiting number of words by random selection.
In order to find out lexical richness, this theoretical vocabulary model was used. 200 words were chosen randomly from all essays. These essays were analyzed in terms of Type-Token Ratio. To measure lexical density, tagged files were analyzed by concordancing softwareAntConc 3.2.4 which is used for carrying out corpus linguistics research and data-driven learning. It helps researchers with a comprehensive set of tools including a powerful concordancer, word and keyword frequency generators, tools for cluster and lexical bundle analysis, and a word distribution plot.

Written productions of learners were compiled and tagged. An analysis with AntConc 3.2.4 was performed to find frequencies of parts of speech. Obtained frequencies were assessed and compared with Log likelihood ratio to find statistical differences, if any, between the frequency of nouns, verbs, and adjectives. In corpus studies, when comparing different sized datasets, chi-square value has often been performed to compare word frequencies across corpora; however, Rayson and Garside (2000) stated that log-likelihood tests are considered to have higher reliability than other statistical methods. Significance difference is tested by log-likelihood ratio which computes overuse and underuse of words. If the log-likelihood ratio is ±3.84 or more a significant difference exists between the two datasets at a 5% significance level. Rayson and Garside (2000: 40) described log-likelihood ratio as follows:

> "Log Likelihood has] the effect of placing the largest LL value at the top of the list representing the word whichhas the most significant relative frequency difference between the two corpora …. Words which appear with roughly similar relative frequencies in two corpora appear lower down the list.

General success of students was evaluated by two raters. These raters independently read the essays according an essay scoring rubric (developed by Oshima and Hogue 2006). Peat (2006) suggested that, because of their explicitly defined criteria, rubrics lead to increased objectivity in the assessment of writing. Results of this grading were first analyzed in terms of reliability. This analysis was done with SPSS 16. Program and Cronbach Alpha's of these grades were found.

RESULTS

The purpose of the study was to analyze and compare written corpora of two groups of students: those in the experimental group and those in the control group. First, lexical richness of two parallel corpora were analyzed and compared. Three types of nouns in the noun category, three types of adjectives in the adjective category, finally, general verbs category were analyzed and compared. Type token ratio of each essay assessed and data from the experimental group and the control group were compared with SPSS 16. Independent samples T-test.It was found that there is no statistical significant difference between two groups in terms of lexical richness.
Nouns are one of the most frequent words that were used both in essays of the experimental group and the control group. Out of 25027 words with 6134 concordances, common nouns constitute 24,5 %of corpus of the experimental group. Percentage of singular nouns, with 4115 concordances, is 16,5 and %8 of corpus is plural nouns, with 2019 concordances; while, in the control group common nouns are %20,5,with 3260 concordances, singular nouns are %14,8, with 2351 concordances, and plural nouns are %5,7, with 909 concordances.

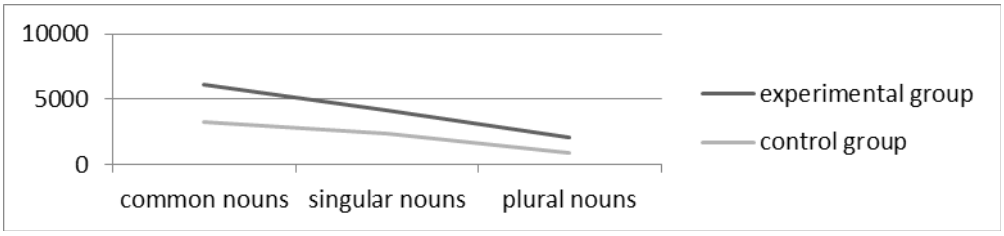

Figure 1:Noun distribution according to its subclasses.

Frequency of nouns in two corpora of students was compared with log-likelihood ratio in order to detect the results of the treatment. Concordances of singular common noun, singular noun + nominalization and plural noun + nominalization were found. The results of Log-likelihood ratio showed that there is a significant difference between writings of students in the experimental group and the control group (LL=65.54 $p < 0.05$ (critical value: 3.84). Findings indicated a significantly higher frequency of use in the experimental group corpus relative to nouns compared to the control group corpus. Depending on the findings in the noun frequency, it can be said that reading strategies, text mining and imitating had a fostering impact on students' use of common nouns in their essays.

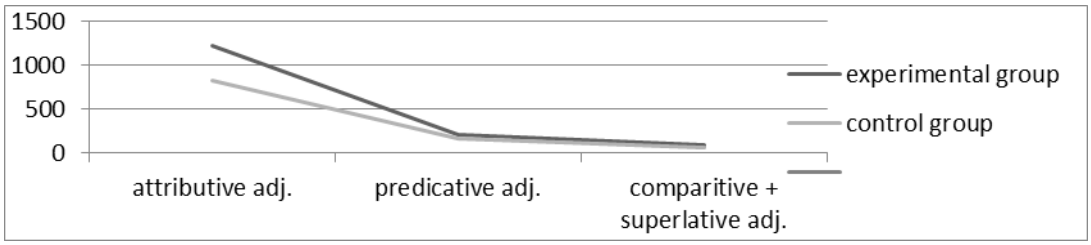

Figure 2. Adjective distribution according to its subclasses

Attributive adjectives constitute 4.9 % of corpus of the experimental group with 1226 concordances out of 25027 words while this number is 5.23% in the control group. Adjectives with predicative function constitutes 0.86% in the experimental group with 214 concordances; 1.03% in the control group. Frequencies of comparative and superlative adjectives were computed together and results showed that frequency of these adjectives is 0.36% in the experimental and 0.35% in the control group.

Log-likelihood ratio results indicated that there is no significant difference between the experimental and the control group in terms of use of frequency of adjectives in their written productions (LL = 2.13, 3.17, 0.03 < 3.84)

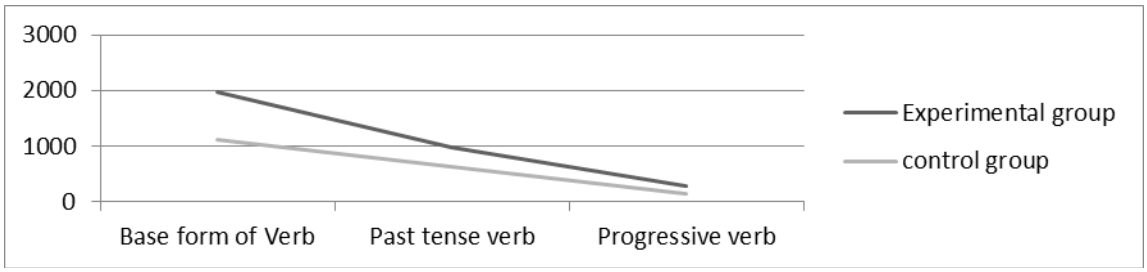

Figure 3: General verb distribution according to its subclasses

Verbs have been investigated in writing samples in terms of base form of a verb, past tense verbs and progressive verbs. Base form of verbs constitute 7,86 % of corpus of the experimental group with 1967 concordances while this number is 7,09 % in the corpus of the control group with 1122 concordances. Students in the experimental group used past tense verbs 966 times, % 3,86 of corpus, while students in the control group used it 633 times, % 4 of their corpus. Progressive verbs were used 286 times (%1,14) by the experimental group; 129 times (%0.82) by the control group.

Log-likelihood ratio showed that there is a statistically significant difference between two groups of students in terms of use of base form of verbs (LL=7.66 >3.84) and use of progressive verb (LL= 10.55 >3.84) in their essays. However, it was found that there is no statistically significant difference between two groups in terms of use of past tense verbs.

The last research question was related to general success of students in writing in terms of essay grading. Two independent raters graded papers according to Essay Scoring Rubric (developed by Oshima, Houge and Butler in 2006. To measure of internal consistency of two raters, Cronbach's Alpha Coeeficientwas calculated by using Spss 16. The alpha coefficient for the two items is .89, suggesting that the items have relatively high internal consistency. A paired-samples t-test was conducted to compare two compositions of the participants. The analysis indicated a significant difference in the scores of the experimental group (M=80.59 SD=5.857) and of the control group (M=69.17 SD=10.027), t(97)=-11.067 p=.000.

DISCUSSION AND CONCLUSION

The present study aimed to investigate the impact of text mining and imitating strategies on lexical richness, diversity and on general writing success of the students. With regard to first research question**,** the statistical analysis of two parallel corpora indicated that there is no significant difference between two groups in terms of type token ratio. As lexical richness is directly related with the quality of written and spoken language produced by learners of second language, at the start of the study, it was assumed that students in the experimental group will have better lexical richness ratio, related to this assumption, it was thought that these students will have statistically significant higher levels of lexical richness. However, tough not statistically significant, the "mean" of lexical richness of the experimental group was higher than that of the control group, which shows there is a slight difference between these two groups. Having similar findings, Laufer (1991: 445) examined written compositions of advanced L2 learners of English and found no improvement in lexical richness over two semesters. Upon this result, Laufer proposed the "active vocabulary threshold hypothesis" which assumes that "passive vocabulary knowledge may continue to develop throughout the lifespan, but "our productive lexicon will grow only until it reaches the average level of the group in which we are required to function." This model may give answers to contradictory results regarding the relationship between lexical richness of the control group who were expected to have significantly lower level of lexical richness and the experimental group vice a versa. From a pedagogical point of view, it can be concluded that lexical richness can be used as a diagnostic tool to identify vocabulary choice of students; however, it cannot be used as a tool that helps discriminate type token ratio of student writings.

With regard to second research question which was concerned with lexical diversity of student writing, the results showed that there is a statistically significant difference in terms of general nouns between the experimental group and the control group. When zoomed in on finer-grained subclasses, statistical analysis also showed that there is a significant difference between two groups in terms of use of plural and singular nouns. The results of the statistical analysis showed no significant difference between two groups in terms of use of attributive, predicative and comparative+ superlative adjectives. Statistical analysis of general verb category showed a significant difference in terms of use of base form of verbs and progressive verbs between two groups of students. As for the effectiveness of text mining and imitating, one cannot deny the value of these strategies in improving student use of general nouns in their essays. This result may stem from the fact that students who analyze model essays tend to use more nouns because of the fact that they have acquired new words incidentally from reading, and their awareness on use of nouns has increased. Similar to this result, Elgord and Warren (2014) investigated acquisition of second language (L2) vocabulary from reading authentic text. The study revealed that number of encounters with new words in reading helped learners gain explicit word knowledge. Similarly, Ponniah (2001) conducted a study on incidental vocabulary learning. He compared performance of the students who devoted their time to reading, and the students who learned meaning of words consciously to develop their vocabulary knowledge. Results of the study showed that the group who tried to learn subconsciously from their readings could use words that they have learnt in sentences while learners who spend their time learning meanings of words from dictionary could not use the previously unknown words in sentences. This study shows that words that are leant incidentally are retrieved better during writing process and learners can put their knowledge of new words into practice better when they read and learn a new word. As text mining requires learners to purposefully dig out for valuable information, it can be thought to include both incidental and intentional word learning. It can be concluded that reading strategies such as text mining and imitating may facilitate both incidental and intentional word learning. This result shows that it will be beneficial for teachers of English to provide their learners with model essays by using text mining and imitating strategies as a pre-writing activity.

General success level of the students in the experimental group has been found to be statistically significantly higher. The most obvious reason for this success may be the opportunity to mine and imitate model essays as pre writing activities which served students as a source of example for rhetorical functions of a text. When had chance to analyze these model essays and see choices in language, students' awareness on how to use what word where has increased, which resulted in better performance in their writing.

LIMITATIONS AND SUGGESTIONS

In this study, impact of reading strategies on writing performance of students, on lexical items in specific, was investigated. Taking number of students (59 in the experimental; 39 in the control group) into account, the results of the study cannot be generalized. However, this study may give a reference point for further research. It is believed that replications of this study with more students at different age levels and educational backgrounds in various ELT contexts such as compulsory service English classes and preparatory English courses will contribute to the field. This study lasted two months, if it were longer, it may have given

more significant results as means of two groups indicate that students in the experimental group use higher rate of parts of speech.

Further research might be conducted to shed light on the use of all categories of parts of speech and discourse markers.

REFERENCES

Baker, P. (2010), Sociolinguistics and Corpus Linguistics, Edinburgh University Press.
Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning, 57, 1,* 35-56.
Broeder, P., Extra, G., & van Hout, R. (1987). Measuring lexical richness and variety in second language use. Polyglot, 8(1), 1-16.
Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. Language Learning, 64, 365–414. doi:10.1111/lang.12052
Grabe, W. (1997), Discourse analysis and reading instruction, In T. Miller (Ed.), Functional approaches to written texts: Classroom applications, (2-15). Washington, DC: USA.
Greene, S. (1991), "Mining Text in Reading to Write", Journal of Advanced Composition http://www.jaconlinejournal.com/archives/vol12.1/greene-mining.pdf (19.02.2013)
Gorell, D. (1987) "Freedom to Write through Imitation", Journal of Basic Writing, 6 (2).
Ito, F. (2011), "L2 Reading-writing correlation in Japanese EFL High School Students" Gunma National College of Technology, The Language Teacher 35 (5).
Johansson, V (2008), Lexical diversity and lexical density in speech and writing: a developmental perspective, Lund University, Department of Linguistics and PhoneticsWorking Papers53 (61).
Laufer, B. (1991), "Ease and difficulty in vocabulary learning: some teaching implications", Foreign Language Annals, 23, 147-156.
Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. Applied Linguistics, 16(3), 307-322.
Lee, S. H. & Muncie, J. (2006), "From receptive to productive: improving ESL learners' use of vocabulary in a post-reading composition task", TESOL Quarterly, 40(2), 295-320.
Nation, I. S. P. (1990), Teaching and Learning Vocabulary. New York: Newbury House.
Nassaji, H. (2006), "The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success", The Modern Language
Journal, 90(3), 387-401.
Oshima, A. and Houge, A (2006), Writing Academic English, Fourth edition, Longman Pearson.
Özçelik, M (1996), A Study of teaching writing through reading to low level prep-school students, published Master Thesis. Retrieved from Turkish Higher Education National Theses Database.
Parodi, G. (2006), "Reading–writing connections: Discourse-oriented research", Reading and Writing Interdisciplinary Journal, July 1-26.
Paton, M. Q. (2002), Qualitativeresearch&evaluationmethods, ThousandOaks: Sage Publication.
Peat, B. (2006), "Integrating writing and d research skills: Development and testing of a rubric to measure student outcomes", Journal of Public Affairs Education, 12, 295–311.
Ponniah, J, R. (2011), "Incidental Acquisition of Vocabulary by Reading", The Reading Matrix, 11 (2).
Rayson P and Garside, R. (2000), "Comparing corpora using frequency profiling. In Proceedings of the workshop on Comparing corpora", 9 (1–6), Association for Computational Linguistics
Schellekens, P. (2007), The Oxford ESOL Handbook. NY: Oxford
Tsai, Jui-min (2006), "Connecting Reading and Writing in College EFL Courses Ohio State University", The Internet TESL Journal, XII (12).
Weigle, S. C. (2002), Assesing Writing, NY:Cambridge University Press.
Wilkins, D.A. (1992), Second- Language Learning and Teaching, Great Britain: The Chaucer Press.