

THE EFFECTS OF VOCABULARY RESOURCE USE ON LEXICAL RICHNESS IN L2 WRITING*

RESEARCH ARTICLE

Kıymet Merve CELEN¹, Şebnem YALÇIN²

* This study was presented at the 51st Annual Meeting of the British Association for Applied Linguistics held between 6-8 September 2018 in York, UK.

1 Öğretim Görevlisi, İstanbul Üniversitesi, Yabancı Diller Yüksekokulu, merve.celen@istanbul.edu.tr, ORCID: 0000-0002-5498-0900.

2 Doktor Öğretim Üyesi, Boğaziçi Üniversitesi, Yabancı Diller Eğitimi Bölümü, sebnem.yalcin@boun.edu.tr, ORCID: 0000-0003-2143-8055.

Geliş Tarihi: 11.02.2020 Kabul Tarihi: 06.04.2021 DOI: 10.37669/milliegitim.687806

Abstract:

This study investigates the effects of dictionary and thesaurus use on lexical sophistication and variation, two of the components of lexical richness in second language (L2) writing. After writing a take-home essay as part of their studies, 27 learners enrolled in an English preparatory program at a state university in Turkey received brief instructions in lexical sophistication and variation. Following the instructions, the learners were asked to revise their essays by using a dictionary and a thesaurus. A comparison of the take-home essays and their revised versions showed that revising take-home essays with the help of vocabulary resources led to a significant increase in lexical sophistication. Significant increases were also found for the measures used for lexical variation, except one of them. The different results obtained for lexical variation will be discussed in relation to the specific characteristics of the lexical measures used.

Keywords: lexical richness, L2 writing, revision, dictionary

SÖZCÜKSEL KAYNAK KULLANIMININ İKİNCİ DİLDE YAZMADA SÖZCÜKSEL ZENGİNLİĞE ETKİLERİ

Öz: Bu çalışma, sözlük ve eş anlamlılar sözlüğü kullanımının, ikinci dilde yazmada sözcüksel zenginliğin bileşenlerinden olan nadir ve çeşitli sözcük kullanımına olan etkisini incelemektedir. Türkiye’de bir devlet üniversitesinin İngilizce hazırlık programına kayıtlı 27 öğrenci, derslerinin bir parçası olarak ders harici bir kompozisyon yazıp teslim etmiş ve daha sonra bu öğrencilere nadir ve çeşitli sözcük kullanımı hakkında kısa bir eğitim verilmiştir. Bu eğitimi takiben öğrencilerden yazmış oldukları kompozisyonları sözlük ve eş anlamlılar sözlüğü kullanarak düzeltmeleri istenmiştir. Ders harici yazılan kompozisyonlar düzeltilen kompozisyonlarla karşılaştırıldığında sözlük ve eş anlamlılar sözlüğü kullanarak yapılan düzeltmelerin nadir sözcük kullanımında anlamlı bir artış meydana getirdiği görülmüştür. Biri hariç olmak üzere, çeşitli sözcük kullanımını ölçen araçlarda da anlamlı artışlar bulunmuştur. Kelime çeşitliliğiyle ilgili olarak elde edilen farklı bulgular, kullanılan sözcük ölçme araçlarının kendilerine has özellikleri üzerinden tartışılacaktır.

Anahtar Kelimeler: sözcüksel zenginlik, ikinci dilde yazma, revizyon, sözlük

Introduction

Knowledge of words is an important aspect of language proficiency. Words serve as the primary conveyors of meaning, which is further complemented and fine-tuned with the help of syntactic, morphological, and pragmatic knowledge of a given language. With the impractical task of learning the totality of words in English on the part of second language (L2) learners as well as those speaking English as a mother tongue (Schmitt, 2000), learning of vocabulary in English naturally stays restricted to a subset of words. Too often, however, this subset may risk resulting in generic and stylistically monotonous use of the language. The development of lexical richness in learners’ production at this point emerges as a key issue in L2 vocabulary instruction. In learning English as an L2, “building a native-sized vocabulary might be a feasible, although ambitious, undertaking” (Schmitt, 2000, 4). The fact that such an attempt can become successful has been shown by Arnaud and Savignon’s (1997) study which revealed no significant difference in the knowledge of rare words in English between the most advanced group of participants in the study and a control group of native speakers. Similarly, Lei and Yang (2020), who compared the writings of Chinese PhD candidates, native beginner students, and native experts, concluded that the nonnative group’s performance in relation to lexical richness was superior to that of beginner

students but inferior to experts. They further pointed out that nativeness does not have a critical influence when it comes to lexical richness. Given that “the tendency to simplify productive vocabulary and lack of communicative need on the part of the learner would act against the development of lexical richness” (Laufer, 1991, 441), the case for both the possibility and the necessity of improving learners’ lexical richness becomes even stronger. After all, vocabulary programs target, as part of their objectives, “bring[ing] learners’ vocabulary knowledge into communicative use” (Laufer & Nation, 1995, 308).

As it may go beyond the accurate use of form and meaning, rich and sophisticated use of lexical items may be regarded as an instructional luxury, particularly when it is juxtaposed with the cases of learners who are reported to struggle with basic or frequent vocabulary. However, “if learning a second language means moving towards the competence of the native speaker (even if such competence is not achieved)” (Laufer, 1991, 445), development of lexical richness becomes not only an important but also an exigent issue that needs closer inspection. Instead of a native speaker comparison, advanced level speakers can also be taken as a model, which might, however, emerge as an alternative no easier for some learners to attain and for some teachers to target due to a multiplicity of reasons. Therefore, it becomes rather crucial to engage L2 learners in the use of a sophisticated and wide variety of words in manageable steps. To this end, the present study aims to investigate lexical richness in L2 writing under two conditions where learners can be considered to have relatively more chances (and also ease) of producing lexically rich texts: (a) writing a take-home essay and (b) revising this essay with the help of both a dictionary and a thesaurus. By creating a need for rich use of lexis as suggested by Laufer (1991), this study compares original essays and their revised versions in relation to the changes in the variety and sophistication of the words. The following section summarizes selected studies conducted on and issues raised for the knowledge of and measurement of lexical richness. Throughout the paper, lexical richness is used as a broader term incorporating at least one of the following: lexical variation, density, diversity, complexity, and sophistication.

Definition and Measurement of Lexical Richness

Although the lexical aspects of a written work do not constitute the sole indicator of its quality (Laufer 1994; Laufer & Nation, 1995), it can be argued that “a well used rich vocabulary is likely to have a positive effect on the reader” (Laufer, 1994, 21-22). One question then is how to decide whether a certain text includes such vocabulary. Lexical richness measures, which “attempt to quantify the degree to which a writer is using a varied and large vocabulary” (Laufer & Nation, 1995, 307), serve the purpose. The broader term of lexical richness includes the use of (a) a diverse set of words (i.e., lexical variation), (b) less commonly used vocabulary (i.e., lexical sophistication), (c)

more use of content words than function words (i.e., lexical density¹), and (d) very low numbers of word errors, all of which are considered as the indicators of a piece of “good writing” (Read, 2000, 200). Lexical richness measures could assist the understanding of the links between the knowledge and use of words and also the criteria that influence how good a text is, which constitutes the two particular motives behind the attention these measures have received (Laufer & Nation, 1995). The use of lexical richness (or diversity) measures on L2 texts has been the most widely employed way of tackling subjectivity involved in the assessment of vocabulary and attaining a more precise understanding of “exactly what constitutes ‘wide’ or ‘adequate’ vocabulary” (Meara & Bell, 2001, 5). Nonetheless, in the computation of lexical statistics “subjective judgement plays an important role” (Read, 2000, 221) and the study of lexical richness is an area where differing views exist (Daller et al., 2003). The variety in the types of lexical richness measures included in studies is likely an indication of such differences. To give an example, Laufer (1991) used lexical originality (see also Linnarud, 1983 for a different calculation), namely the percentage of lexemes used by only one writer in a given group, in addition to the lexical variation, sophistication, and density measures in her analyses of lexical richness in learners’ compositions, and Lu (2012) did not include word errors as a measure of lexical richness in his study focusing on L2 oral production. These differences in the inclusion and exclusion of particular measures of lexical richness have also been accompanied by the use of different tools in the literature (e.g., Laufer & Nation, 1995; Meara & Bell, 2001).

Various calculations have also been proposed and used in order to measure the same component of lexical richness. To start with lexical variation, the type-token ratio (TTR), with its reliance on text length (Johnson, 1944; Malvern et al., 2004; Meara & Bell, 2001; Read, 2000; Saito et al., 2016), evidence repudiating its validity (Daller et al., 2003) and suggestions against its use (e.g., Vermeer, 2000), has been a measure of lexical richness (lexical diversity/variation, more specifically) which has both attracted much investigation aiming for its improvement and tended to be replaced with other measures. For instance, Laufer (1991) used the first 250 words of the essays written by her participants in order to tackle the effects of text length in her analyses. D was another measure introduced by Malvern et al. (2004) and it was a measure of lexical variation and it has been used in the field extensively. Several researchers have used (Polat & Kim, 2014; Gebril & Plakans, 2016) D as a measure of lexical diversity. Issues other than text length have also been raised with regard to the calculation of lexical variation. Whether it should be derived words or word families which should represent different words in learner production is an example of such issues (Laufer & Nation, 1995). Similar divergences have also been observed for lexical sophistication. In addition to the different labels they have been given, such as “advanced” (Laufer, 1991), “non-basic” (Laufer, 1994), or “difficult” words (Meara & Bell, 2001), sophistica-

¹ Created by Ure, 1971 by Read, 2000

ted words have been defined in relation to specific word lists. It has also been claimed that “what is labelled as ‘advanced’ would depend on the researcher’s definition” (Laufer & Nation, 1995, 309) as well as the target population investigated (Laufer, 1994). Additionally, both word frequency and familiarity (Saito et al., 2016) have been used to ascertain the sophistication level of the words found in learner production. Meara and Bell (2001) offered a different way to calculate lexical complexity, by using lambda values for the difficult words found in each 10-word sections of a text. They further argued that this approach worked more efficiently than Laufer and Nation’s (1995) Lexical Frequency Profile because it functioned successfully with shorter texts and was therefore appropriate for use with texts produced by low proficiency learners.

Lexical Richness in L2 Written and Oral Production

Lexical richness has been investigated in the literature, with varying degrees of emphasis attached to it. In addition to the attention it received in studies of language development, it has also been a well-researched aspect of L2 production in studies targeting complexity, accuracy, and fluency. Regardless of the role it has undertaken, the importance of insight it offers for improving L2 vocabulary instruction and thus learner production appears to remain unchanged.

Several studies with differing complexities of design have used tasks to tap the components of lexical richness as part of their target language performance measures. In one of such studies, Gebril and Plakans (2016) investigated the effects of borrowing words from source materials on lexical diversity and the difference in integrated writing score levels observed for lexical diversity. The participants, who all spoke Arabic as a native language, were undergraduate students at a university where English was the medium of instruction. In an integrated task, the participants were provided with two reading passages from which they were expected to integrate ideas or information into their writing. Using the D index (Malvern et al. 2004), lexical diversity of the original essays were compared with the versions where the source vocabulary was removed. Significantly higher lexical diversity mean scores were found for the essays with source-related vocabulary and differences were also found for lexical diversity across writing scores obtained, for essays with and without source vocabulary.

Ong and Zhang (2010) investigated fluency and lexical complexity in L2 writing from a task complexity perspective and with respect to three parameters: (1) time available for planning and writing (+/- planning time), (2) the kind of writing assistance given (+/- ideas and macro-structure), and (3) use of first drafts during the revision process (+/- draft availability). The participants were Chinese English as a Foreign Language (EFL) learners who were students at a comprehensive university. Lexical complexity scores were obtained by dividing the squared number of word types by the total number of words, which the authors argued would consider text length. An increase in task complexity concerning planning time produced a significant differen-

ce for lexical complexity, which was found to be significantly higher in free-writing (no planning time) condition than pre-task (10 min planning) and extended pre-task (20 min planning) conditions. Similarly, the pre-task condition (and even the control condition) revealed significantly higher lexical complexity values than the extended pre-task, but no significant differences were found in students' first drafts as the task complexity increased with regard to the type of writing assistance. For the second drafts, significant differences were found for the type of writing assistance condition, but not for planning time or draft availability. However, the conditions where the topic, ideas, and/or macrostructure were provided produced significantly higher lexical complexity than the condition which provided only the topic. Among the possibilities, Ong and Zhang pointed out to explain the better lexical complexity results found for the free writing condition were following: (1) the students in the planning conditions employed more online planning than those in the free writing condition and (2) the facilitating effect of the improved fluency observed in the free writing condition.

With a very large sample consisting of EFL learners at an institution who spoke Spanish as their first language, Johnson et al. (2012) examined the effects of pre-task planning sub-process on the fluency, lexical complexity, and grammatical complexity of L2 writing. The participants, whose mean age was reported lower than 21 years, were from their institution's advanced level classes; however, many of them were reported to be not as accomplished in terms of their writing proficiency. Five pre-task planning conditions (control, idea generation, organization, goal setting, and goal setting + organization) were created for the study. Lexical complexity of the essays was analyzed in relation to five measures. The results showed that pre-task planning did not have a significant effect on lexical complexity. In their discussion of the differences between the results of their study and those reported in others, Johnson et al. referred to issues such as a threshold proficiency level, a mediator which would have allowed room for pre-task planning to show its effects, and genre knowledge. The authors also noted that the lexical diversity measure used by Ong and Zhang (2010) was sensitive to text length and might have actually reflected the effect of pre-task planning on text length, as participants who were given more planning time were allowed less time to write their essays.

In addition to the studies investigating the effects of differing variables on lexical richness in L2 writing, development-oriented inquiries have focused on changes in lexical richness scores by comparing language production at two different points (e.g., at the beginning and at the end of a language program). In a relatively early study, Laufer (1991) compared essays written by two groups of university students, with the first essays written at an entrance exam and the second essays written at the end of either the first or second semester in the program. The participants, who spoke Hebrew or Arabic as their native language, were advanced level L2 learners in their first year of studies in a department of English language and literature. Lexical richness was opera-

tionalized in relation to lexical variation, lexical density, lexical originality, and lexical sophistication. Only lexical sophistication scores of the students who wrote their second essays at the end of the second semester improved significantly. Referring to the more progress achieved by students, the majority of whom started off with scores of lexical richness which was lower than their group averages, Laufer (1991) pointed out that it was the role of learner needs that worked, more strongly than comprehensible input, as a driving force behind the development of active vocabulary.

In another study focusing on language development over time, Mazgutova and Kormos (2015) examined the changes in lexical and syntactic properties of argumentative writings of undergraduate (lower proficiency) and postgraduate (higher proficiency) students at an English for Academic Purposes program. The essays were written at the beginning and at the end of the program. Significant differences were observed for all of the five lexical diversity measures for the lower proficiency group while only two of the measures were found to show significant improvement for the higher proficiency group. Mazgutova and Kormos concluded that the lower proficiency group's visible improvement in lexical diversity and sophistication was an indication of how gains in lexis may occur when no explicit teaching of vocabulary is provided.

Although the current study focuses on lexical richness in L2 writing, insight coming from research on L2 speech may prove useful for better positioning the characteristics as well as idiosyncrasies of the components of lexical richness and how they are used in explaining learner performance. With an extensive set of 26 measures used for analyses and Chinese learners' speech data coming from multiple groups, which were retrieved from a corpus and came from a test given to students majoring in English in 4-year colleges in China, Lu (2012) investigated how three of the components of lexical richness (i.e., lexical variation, density, and sophistication) were related to the ratings of L2 learners' oral narratives. The findings revealed that lexical variation, in comparison to lexical density and sophistication, likely had a larger impact on raters' judgments of the quality of L2 oral production. Given the impact of lexical sophistication, which was reported to be highly small, Lu underscored the necessity of an emphasis on variety, rather than on sophistication, for L2 vocabulary learning contexts.

By taking the focus from L2 output itself to how it is perceived by raters in terms of comprehensibility, Saito et al. (2016) analyzed oral narratives of a mixed proficiency group of participants in relation to 12 lexical variables including lexical variation (diversity) and sophistication (frequency, familiarity). The participants, native speakers of French from Quebec, Canada with a mean age of around 36 years, spoke English as an L2. Although lexical variation and word familiarity were both found to be significantly related to ratings of comprehensibility, it was lexical variation which significantly differed across the three proficiency groups, whereas word familiarity did so only between beginner and advanced speakers. Overall, the authors pointed out that accu-

racy (and probably fluency) with which L2 words were used was more central to better L2 oral production than lexical diversity and sophistication were. One of the possible explanations for the finding that word frequency, as part of lexical sophistication, did not have a significant relationship with participants' speech comprehensibility ratings had to do with task characteristics (i.e., a picture sequence description task).

In a longitudinal study focusing on L2 oral development in an uninstructed setting, Polat and Kim (2014) examined the speech of an advanced speaker of L2 English, an immigrant whose native language was Turkish, by using D for the calculation of lexical diversity scores of speech data coming from interviews held periodically with the participant for a year. When compared to accuracy and syntactic complexity, lexical diversity was found to show a more visible improvement and it improved with a constantly rising trend. The comparisons with the native speakers indicated that the lexical diversity values elicited from the participant were compatible with those produced by a comparison group of native speakers. The participant, who showed lexical diversity values close to the native speaker range during the initial phases of the study, outperformed the native speakers repeatedly through the end of the study.

To sum up, though cautiously due to the characteristics of written and oral modalities, it appears that lexical variation is positively linked to ratings of L2 output (Gebriel & Plakans, 2016; Lu, 2012; Saito et al., 2016) and is open to improvement even when no formal instruction is provided (Polat & Kim, 2014). Although it is affected by external source use (Gebriel & Plakans, 2016), how it is influenced by pre-task planning differs (Johnson et al., 2012; Ong & Zhang, 2010). When compared to lexical variation, lexical sophistication might improve more substantially in time (Laufer, 1991) and different proficiency levels may display different patterns of improvement in relation to lexical measures (Laufer, 1991; Mazutova & Kormos, 2015).

The Current Study

Many studies (e.g., East, 2007; Gebriel & Plakans, 2016) investigated the use of external sources in L2 writing with timed tasks. This study, however, attempts to trace learners' use of L2 vocabulary at a different end of the continuum, where the effects of instruction may well be observed: a take-home essay assignment and its revision. In other words, this study employs a twofold scaffold for learners with the use of (1) take-home essays which the learners had previously submitted and (2) a following in-class revision session where learners were supported with resources, namely a (monolingual) dictionary and thesaurus, to make their take-home essays lexically richer. That the first group of essays came from a take-home condition, where learners could have well-referred to any external resources at their disposal, is considered as a noteworthy aspect of this study which investigates learner performance informed by an explicit focus on lexical richness. Instead of taking a set of learner essays produced under exam conditions as the originating dataset, this study compares take-home es-

says with their revised versions produced with the help of external resources, thereby making the comparison between the two relatively flexible conditions more viable. This study, therefore, seeks answers to the following questions:

1. How do participants in the study revise their own texts?
2. Does revising take-home essays with the help of both a dictionary and a thesaurus lead to an increase in
 - 2a) lexical sophistication?
 - 2b) lexical variation?

Methodology

Research context and participants

The study adopts a within-subjects study design with intact classes. Twenty-seven L1 Turkish EFL students (17 female, 10 male) enrolled in an English Preparatory Program at a state university in İstanbul, Turkey participated in the present study. Following their placement in the undergraduate program in Foreign Language Education based on their performance in nationwide university entrance exams, the students are required to show proof of a certain level of English proficiency or pass an institution based proficiency exam in order to start their degree program. Those who fail the exam and those who do not take it, voluntarily or for other reasons, are required to take English courses in the preparatory program. The participants of the present study constituted a mixed-ability group whose English language proficiency was considered to be below C1 by the institution. Based on instructors' comments, the participants' language proficiency level ranged between zero beginner and B2. In the Fall semester, when the data were collected, the students' weekly schedule consisted of speaking, writing, reading, and listening courses (three hours each) as well as grammar and project-based learning courses (four hours each).

Materials

Take-home essays and self-check worksheets: By the time of the study the learners had practiced sentence and paragraph level conventions of academic writing as well as unity and coherence in paragraphs. Compare and contrast was the first essay type they learned in this course and they had had enough practice (i.e., two weeks). The participants were asked to write a 250-word compare and contrast essay and given the following instruction by their class instructor as a part of their regular course syllabus:

“Compare and contrast different aspects of living in a small town and living in a large city. You should provide supporting information for your ideas and give examples where necessary. The essay should be of at least 250 words.”

With the due date set two days before the implementation of the study, the comparison and contrast essays were submitted in an electronic format to an online platform used for the class. Printed copies of these essays worked as the base forms to which the participants would make their revisions. In addition to the individual essays, single-page self-check handouts (see Appendix A) were used in order to ensure that the participants understood the instructions correctly and to guide them through task completion.

Simple and rich text samples: The original text narrating the study abroad experience of a student was taken from a newspaper article (Carruthers, 2017). Only the first three paragraphs of the text were kept for the purposes of the study. The text created this way was the one that was presented to the students as “the text with more difficult vocabulary” (Rich text; see Appendix B). By replacing the selected words in this text with their more frequent counterparts (and making any grammatical changes where necessary), “the text with simpler vocabulary” (Simple text; see Appendix C) was created. The Corpus of Contemporary American English (COCA) (Davies, 2008-) was used to check the frequencies of the words that were considered more common in order to ensure an objective measure of rareness. Uninflected forms were searched for inflected verbs such as *felt*, *indulged*, *was*, and *did*. The selected words in the rich text and their more common counterparts used in the simple text, along with their frequency information, are given in Table 1.

Table 1. Word frequencies for selected words in the rich and simple text

Rich text words	COCA frequency	Simple text words	COCA frequency
tough	45446	difficult	87675
depart	2001	leave	104326
prove	26672	be	2594062
whirlwind	1313	busy	25295
constantly	18064	always	256615
anti-climax	23	disappointment	6789
feel	193225	be	2594062
indulge	2096	do	1866583
lounge	5295	sit	55330
tranquil	1183	quiet	37210
restless	3506	uncomfortable	11159

Vocabulary Resources: The online dictionary and thesaurus used in the study were (1) the Merriam-Webster Dictionary (<https://www.merriam-webster.com/>) and (2) Thesaurus.com (<http://www.thesaurus.com/>) respectively. The choice of a mo-

nolingual dictionary over a bilingual one was motivated by two reasons. First, the students were provided with their complete essays for revision, meaning that there was no need to produce new ideas (and thus write words) which would create the possibility of relying on Turkish words and of finding their English equivalents by using a bilingual dictionary. Second, monolingual dictionaries were considered to be better alternatives for providing more opportunities for learning and using L2 vocabulary better due in part to the number of examples provided.

Procedure

After a relatively shorter classroom meeting with their writing instructor on the day of implementation, the participants were directed to the computer lab, the setting of the study, by their instructor. The participants were then informed about the study and signed an informed consent form. They were further told that they would later revise their take-home midterm essays to make them lexically richer by using two online resources. Following this announcement which aimed a focus on the task, lexical sophistication was explained to the participants through the comparison of the Rich text and Simple text, in a single sitting and a presentational format. This comparison session was followed by a verbal information session on lexical diversity, which was aimed to guide the participants to use as many different words as possible in the revised versions of their essays. The participants were also asked to use words correctly but not to change their original points or ideas while revising.

After this brief instruction, print copies of the take-home essays and the single-page self-check handouts were given. The participants were given an hour to revise their essays by using the online dictionary and thesaurus, both of which were opened up on individual computers. It was expected that participants who first used the thesaurus to find more “difficult or less common” words would later search word usage, if not word meaning, in the dictionary before replacing those words with their original words in their revised essays. The participants were also told that blank sheets would be provided if they chose to revise their essays by writing on a separate sheet or that they could also get their previously submitted soft copies on which they could type their revised version. Only one of the participants submitted her revised essay electronically. The remaining of the participants submitted the print copies back as the revised versions of their essays. The complete session took an hour, as announced initially.

Analysis

The online version of the Lexical Complexity Analyzer (LCA) (Ai & Lu, 2010; Lu, 2012) was used for the analysis of the original and revised essays in the single mode. Overall, the program runs 25 measures of lexical richness. As one of the alternatives measuring lexical sophistication in the analyzer, Lu (2012) includes Laufer’s (1994) treatment of sophisticated words, i.e., those beyond 2,000 most frequent words. Laufer (1994) explains the categories following the first 2,000 words as (1) the University

Word List (Xue & Nation, 1984) and (2) those that are not found in any of the other lists. This study used only this alternative for the analysis and hence the definition of lexical sophistication.

To prepare the materials for analysis, the changes made in the revised print versions were transferred to the electronic copies of the original essays. Some changes were very legible and clear to spot, whereas others posed certain challenges and therefore necessitated certain decisions. Spelling errors (e.g., *seperate*) were corrected in the revised versions. To a large extent, this decision was motivated by the fact that this study focused more on word choice than on the accuracy of form. For cases when the participants used more than one word to replace the original words that they used (e.g., *astounding/devastating* for *overwhelming*), the first word was taken into consideration. When there was a question mark attached to the new word (e.g., *meager (?)* for *small*), the new word was still treated as the participant's ultimate decision and was inserted in the revised version. Changes made to other parts of the text (e.g., deletion of an incorrect preposition) and corrections made to the spelling of original words (e.g., *organically* for *organichly*) were not transferred to the revised versions. The newly written words were considered as missing when they were indecipherable and when the original words that they replaced were not marked. All information other than the essay texts were removed. In addition, incorrectly spaced punctuations (e.g., *life.Maybe*) which could have made two separate words count as if they were a single word, were corrected. Finally, as the analyzer required a choice of language variety (i.e., British English and American English), both groups of essays were checked to comply with the American English version with the help of MS Office speller utility.

Following the concerns raised against the TTR, multiple measures were targeted for the calculations of lexical variation. Given the sizeable differences in the text length produced in both conditions, such a preference indeed became even more inevitable. Among the alternatives provided for lexical variation in the analyzer, only the following were used:

1. Number of different words (NDW) for the first 50 words (NDW-50)
2. Mean NDW for randomly selected ten 50-word samples (NDW-ER50)
3. Mean TTR for every 50-word segment (MSTTR-50)

In a paper describing the LCA, Lu (2012) refers to strategies put forward by Malvern et al. (2004) for the NDW-based measures given above, and Johnson (1944) for MSTTR, in order to deal with the effect of text length on the calculations. Two of these three measures of lexical variation were also among those which functioned most successfully in Lu's (2012) analyses of L2 speech data. The LCA output for the essays was entered into SPSS. Wilcoxon signed-rank tests were performed to compare the

two groups of essays with regard to the target measures of lexical sophistication and variation described above.

Results

The first research question was descriptive in nature and concerned (a) the text length characteristics of the two groups of essays and (b) the amount of changes made for the revised versions. In line with the prompt given for the original take-home essays, which asked the participants to use at least 250 words, an initial analysis revealed that no essay included fewer words than the minimum amount. However, large differences were observable between the participants in terms of the number of words they used in their essays. The large difference between the shortest and the longest essay as well as the more reliable standard deviation of approximately 136 words meant that the participants were quite diverse in terms of elaboration (in quantitative terms) in their essays. In terms of the changes made to the vocabulary in the essays, differences were existent between the participants ($M = 18.07$, $SD = 11.30$). These changes were considered as the number of instances where a word (or a group of words) was replaced with another word (or a group of words). Overall, the number of changes made to the words showed a highly unequal distribution, with the minimum and maximum number of changes being 2 and 56 respectively (see Table 2).

Table 2. Text length comparisons and the number of changes in the revised essays

	Number of words		Number of changes
	Take-home essays	Revised essays	Revised essays
<i>M</i>	443.96	443.92	18.07
<i>SD</i>	136.44	136.33	11.30
<i>Min.</i>	272	274	2
<i>Max.</i>	808	801	56

The second research question concerned the difference between the two groups of essays in terms of the lexical sophistication and lexical variation scores they elicited. On average, with the help of a dictionary and thesaurus the participants used more less frequent words in their revised essays ($Mdn = 0.23$) than in their original take-home essays ($Mdn = 0.19$), $z = -4.30$, $p < .05$, $r = -0.59$. This indicated that when asked explicitly to reconstruct their essays in a lexically sophisticated manner, the participants used less frequent words in significantly more amounts than they did in their take-home essays.

As indicated, the second research question also explored the difference between the two groups of essays in terms of the lexical variation scores they elicited. Of the three measures targeted, significant improvement was found for only two of them, namely NDW-50 and MSTTR-50 (see Table 3).

Table 3. Results of Wilcoxon signed-rank tests

Measure	<i>Mdn</i> ^a	<i>Mdn</i> ^b	<i>z</i>	<i>p</i>	<i>r</i>
Lexical sophistication	0.19	0.23	-4.30	< .001	-0.59
Lexical variation					
NDW-50	34.00	35.00	-2.36	.018	-0.32
NDW-ER50	38.10	38.40	-1.27	.203	-0.17
MSTTR-50	0.74	0.74	-3.34	.001	-0.45

Note. ^atake-home; ^brevised

When the first 50 words of the essays were taken into consideration significant differences were found between the two conditions. In others words, with the help of a dictionary and thesaurus the participants used more diverse vocabulary in the first 50-word portion of their revised essays (*Mdn* = 35.00) than they did for their original take-home essay (*Mdn* = 34.00), $z = -2.36, p < .05, r = -0.32$. However, when the analysis was conducted on randomly selected ten 50-word samples from the essays, no significant differences were found in the number of diverse words between the essays revised with the help of a dictionary and thesaurus (*Mdn* = 38.40) and the take-home essays (*Mdn* = 38.10), $z = -1.27, p > .05, r = -0.17$. These findings, which are based on measures using the number of new words in specific locations or samples in the text, appear to be contradictory. This is further discussed in the following section. The final lexical variation measure, MSTTR-50, produced significant results and revealed that revising with the help of a dictionary and thesaurus led to lexical variation values that were larger (*Mdn* = 0.74) than those found in original take-home essays (*Mdn* = 0.74), $z = -3.34, p < .05, r = -0.45$. As a type of TTR, this measure paralleled the meaningful change found for NDW-50, with almost a large effect size following Cohen’s (1988) conventions for effect size (Field, 2009).

Discussion and Conclusion

This study aimed to investigate the revisions made to essays which were first written as part of a take-home assignment and later revised with the help of a dictionary and thesaurus to improve their lexical richness. It was observed that the participants, although descriptively, varied not only in the length of essays they produced but also

in the number of changes they made to create lexically richer essays. The fairly unchanged text length values after revision may be explained with a tendency on the part of the participants to replace their original words with others that comprised the same number of parts or constituents. This might have been caused by the brief introductory session on the lexical richness and the self-check worksheets which provided examples of changes primarily between words with identical numbers of constituents (e.g., *leaving for-departing to, different-distinct*). Treating words as alienated parts of sentences in this way, if overused, might work as a threat to the creativity of the writer and the writing process and may prove too mechanic and clumsy, especially for certain words. For this reason, learners need to be informed that lexical variation or sophistication is not necessarily achieved by the replacement of lexical units with others which have the same number of constituents, but changes between lexical units that have unequal numbers of constituents might be necessary, depending on the meaning targeted and the context.

As one of the lexical richness measures of this study, lexical sophistication was found to be an area of improvement. Revising with two online resources led to significantly higher lexical sophistication scores, meaning that the participants were able to use larger numbers of less frequent words. The observed improvement in lexical sophistication, a component of lexical richness involving “the use of technical terms and jargon as well as the kind of uncommon words that allow writers to express their meanings in a precise and sophisticated manner” (Read, 2000, 200), might be considered as a successful judgment of rare or less frequent words on the part of the participants. Although the thesaurus search provided the participants with a list of synonyms for the words they searched, it was the participants themselves who opted for a particular word.

Mixed results were elicited for lexical variation. The first 50 words of the revised essays were more varied than those of the original essays, whereas the randomly selected ten 50-word samples were not. One possible interpretation of these contradicting findings might be that the random samples of 50 words selected from the text were already at a certain level of lexical variation and thus did not change much after revision. As the development of ideas tends to emerge somewhat later than the first 50 words in a typical essay, it may be that the introductory parts (i.e., first 50 words) of the essays analyzed in this study already included more repetitive vocabulary or vocabulary that turned out to be more convenient for participants to revise. Because MSTTR is a more advanced measure than NDW, though with its limitations, (Malvern et al., 2004), it can be argued that revision with the help of a dictionary and thesaurus resulted in significantly more use of varied vocabulary. One point of caution is that the use of less frequent words might have positively contributed to the use of different vocabulary or vice versa. In other words, while replacing a less frequent word with a more frequent one, the participants might have used a new word which was both less frequent and unique

in the entire set of words used in a particular essay. It should be noted, however, that randomly selecting ten 50-word samples (NDW-ER50) and calculating the mean TTR for every 50-word segment (MSTTR-50) of essays had their unique weaknesses. The analyzer used in the study employs sampling with replacement for the former measure. This means that there is a chance for a word to be sampled in one of the ten samples more than once (Malvern et al., 2004). For essays exceeding 500 words, there was some waste of data as well. For the latter measure, as the texts could not be all divided into 50-word segments, there was again some loss of data (Malvern et al., 2004).

This study can be regarded as an attempt following the advice that “learners should be taught that effective lexical use implies not only the use of sophisticated vocabulary, but also effective variation of words” (Laufer, 1994, 32). Bearing in mind the features and limitations of the approaches to lexical richness discussed so far, it can be argued that rising trends were highly visible in lexical variation, although not always statistically significant. Overall, this study showed that using external resources for revising a take-home essay contributed positively to participants’ lexical sophistication. Different measures of lexical variation yielded differing results, but there was evidence for improvement. All the possibilities discussed so far in this section, however, require further analyses to make the results of this study more definitive.

Limitations and Future Research

Our results should be interpreted with caution due to some limitations. First, the participants’ use of vocabulary resources was not tracked. Therefore, it was not possible to ascertain whether all the changes in the lexical measures of the essays were as a result of participants’ use of the resources. Future studies should therefore quantify participants’ use of resources during task completion in order to tap the pure effects of resource use. Second, this study did not focus on the number of word errors, which is a component of lexical richness and “possible measure of writing quality” (Read, 2000, 201). Especially in studies of this kind, where the participants are not only explicitly asked to use sophisticated and diverse vocabulary but also are given thesauruses/dictionaries to do so, the rate and nature of errors may require even closer scrutiny. In the context of wrong word choices made by students as a result of using dictionaries emerges the term “dictionary howlers” (East, 2008, 8). East (2008) notes that the commonly encountered issue of such erroneous uses might be among the most compelling claims made against the use of dictionaries and that they may have deleterious effects on learners’ performance in exam situations. Although the revised essays written by the participants in this study were not scored as part of an exam, the same logic stays valid. It should be ensured that writers do not use sophisticated and diverse vocabulary at the expense of using correct words. Saito et al.’s (2016) study, which revealed that appropriateness was a more crucial element for comprehensible L2 oral production in comparison to lexical diversity and sophistication, was one empirical support for the importance of ensuring the correct use of words along with diversity and sophistica-

tion. After all, intelligibility is an important aspect of communication and it would be rather problematic to elicit learner output which is rich in rare and diverse vocabulary, but which makes little sense to the reader or hearer. Third, this study did not include language proficiency as a potential factor influencing participants' choice and use of different and less frequent words. Previous studies indicated that lower proficiency learners showed more improvement in lexical richness (e.g., Laufer, 1991) or in more features of it (e.g., Mazgutova & Kormos, 2015). Moreover, the discriminatory power of certain components of lexical richness for different proficiency levels (e.g., Saito et al., 2016) was also revealed previously. In light of these findings, it is expected that a proficiency-informed lens would stimulate a more comprehensive discussion of lexical richness in L2 writing, or learner production in general. Finally, the results are limited to this specific web-based analyzer and the specific selection of measures. Future research should explore the issues with different measures to ensure generalizability.

References

- AI, H., & Lu, X. (2010, June). *A web-based system for automatic measurement of lexical complexity*. Paper presented at the 27th Annual Symposium of the Computer-Assisted Language Consortium (CALICO-10), Amherst, MA.
- ARNAUD, P. J. L., & Savignon, S. J. (1997). Rare words, complex lexical units and the advanced learner. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 157-173). Cambridge: Cambridge University Press.
- CARRUTHERS, F. (2017, July 20). My reverse culture shock: returning from a year abroad is tough. *The Guardian*. Retrieved from <https://www.theguardian.com/uk-news>
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). HILLSIDE, NJ: Lawrence Erlbaum Associates.
- DALLER, H., van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24(2), 197-222.
- DAVIES, M. (2008-) *The Corpus of Contemporary American English (COCA)*. Available online at <https://www.english-corpora.org/coca/>.
- EAST, M. (2007). Bilingual dictionaries in tests of L2 writing proficiency: Do they make a difference? *Language Testing*, 24(3), 331-353.
- EAST, M. (2008). *Dictionary use in foreign language writing exams: Impact and implications*. Amsterdam: John Benjamins B.V.
- FIELD, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE Publications Ltd.
- GEBRIL, A., & Plakans, L. (2016). Source-based tasks in academic writing assessment: Lexical diversity, textual borrowing and proficiency. *Journal of English for Academic Purposes*, 24, 78-88.

The Effects of Vocabulary Resource Use on Lexical Richness in L2 Writing

- JOHNSON, W. (1944) Studies in language behavior: I. A program of research. *Psychological Monographs*, 56(2), pp. 1-15.
- JOHNSON, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, 21(3), 264-282.
- LAUFER, B. (1991). The development of L2 lexis in the expression of the advanced learner. *The Modern Language Journal*, 75(4), 440-448.
- LAUFER, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, 25(2), 21- 33.
- LAUFER, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- LEI, S., & Yang, R. (2020). Lexical richness in research articles: Corpus-based comparative study among advanced Chinese learners of English, English native beginner students and experts. *Journal of English for Academic Purposes*, 47, 1- 9.
- LINNARUD, M. (1983). On lexis: The Swedish learner and the native speaker compared. In K. Sajavaara (Ed.), *Cross-language analysis and second language acquisition 2*, (pp. 249-261). Jyväskylä: University of Jyväskylä.
- LU, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208.
- MALVERN, D. D., Richards, B. J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Houndmills: Palgrave Macmillan.
- MAZGUTOVA, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for academic purposes programme. *Journal of Second Language Writing*, 29, 3-15.
- MEARA, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16(3), 5-19.
- ONG, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, 19(4), 218-233.
- POLAT, B., & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics*, 35(2), 184-207.
- READ, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- SAITO, K., Webb, S., Trofimovich, P., & Isaacs, T. (2016). Lexical profiles of comprehensible second language speech: The role of appropriateness, fluency, variation, sophistication, abstractness, and sense relations. *Studies in Second Language Acquisition*, 38(4), 677-701.
- SCHMITT, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- VERMEER, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17(1), 65-83.
- XUE, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, 3(2), 215-229.

Appendix A. Self-check Handout

Self-check

Before you submit your essay, please put a check mark (✓) to the boxes (□) to show that you completed your task.

1. I used **difficult words** instead of simple words. □

Example:

Old sentence: Spending some time abroad is very important for learning about different cultures and lifestyles.

New sentence: Spending some time abroad is essential for learning about different cultures and lifestyles.

2. I used as many **different words** as possible. □

Example:

Old sentences: Spending some time abroad is very important for learning about dif-
ferent cultures and lifestyles. Moreover, meeting people from different
countries can be very useful for starting lifelong friendships. Different
travellers have shared such experiences on their blogs or social media;
however, what they suggest for their readers is quite different from
each other.

New sentences: Spending some time abroad is very important for learning about
different cultures and lifestyles. Moreover, meeting people from other
countries can be very useful for starting lifelong friendships. Various
travellers have shared such experiences on their blogs or social media;
however, what they suggest for their readers is quite distinct from
each other.

3. I did not change the meaning of the sentences. (I only used different and more difficult words.) □

Appendix B. Rich Text (“Text with more difficult vocabulary”)

My reverse culture shock: returning from a year abroad is tough

After a life-changing year studying in Canada, coming back home felt like a backward step.

Language barriers. Culture shock. Homesickness. These are the things you might worry about before departing to study abroad. But for me, returning home proved the

hardest of all. After the whirlwind experience of a year abroad – constantly meeting new people, having new experiences and gaining independence – coming home can feel like an anti-climax. Like many other exchange students I found myself living with my parents again after a year in Canada. It felt like I'd taken several steps backwards. At first, I indulged in everything I had always loved about being home: days spent lounging with cups of tea and books, tranquil walks and home-cooked dinners. But once the jetlag subsided, I found myself strangely lacking energy and motivation. I felt constantly restless and slept very little.

Appendix C. Simple Text (“Text with simpler vocabulary”)

My reverse culture shock: returning from a year abroad is difficult

After a life-changing year studying in Canada, coming back home felt like a backward step.

Language barriers. Culture shock. Homesickness. These are the things you might worry about before leaving for study abroad. But for me, returning home was the hardest of all.

After the busy experience of a year abroad – always meeting new people, having new experiences and gaining independence – coming home can feel like a disappointment. Like many other exchange students I found myself living with my parents again after a year in Canada. It was like I'd taken several steps backwards. At first, I did everything I had always loved about being home: days spent sitting with cups of tea and books, quiet walks and home-cooked dinners. But once the jetlag subsided, I found myself strangely lacking energy and motivation. I felt constantly uncomfortable and slept very little.