【Research Article】

# Training Learners of Writing to Use Online Vocabulary Tools to Increase Lexical Richness

**Tanya L. Erdelyi**

**Abstract**

Lexical richness is often used as an indicator of productive language proficiency. Researchers regularly utilize a variety of digital tools (i.e., VocabProfile, RANGE) to analyze the lexical richness of a language learner's written text. However, many studies do not provide the opportunity for learners to use the same tools to analyze their own writing. The following longitudinal mixed-methods study analyzed the written essays of 36 Japanese lower English proficiency first-year university students, before and after they received training in the usage of VocabProfile, Corpus of Contemporary American English, and Google's Ngram Viewer. The first and final drafts of the essays were analyzed with RANGE and compared with a paired-samples *t*-test. Next, 65 randomly selected sentences containing low-frequency words were analyzed with a 4-point scale for errors by six human raters. Finally, a self-efficacy questionnaire completed by the students was compared with the statistical results. The results indicated a significant increase in the error-free lexical richness of the language learners' essays between the first and final draft.

**Keywords:** *Academic writing, Lexical richness, VocabProfile*

Learning vocabulary is a key component for developing language proficiency. Once the most common high-frequency words in English have been obtained (i.e., the first 2000 most frequent words), the next logical step is for language learners to continue increasing their academic and low-frequency vocabulary knowledge, especially for language learners at the tertiary level of education. Increasing a learner's lexical richness can assist the learner in the reading comprehension of academic texts, as well as enable them to mirror the language used in their field of study while writing.

Lexical richness is often used to determine the academic quality of a language learner's productive language. In terms of writing, generally speaking, lexical richness is determined by analyzing a language learner's written work for the density and variety of low-frequency words. Furthermore, the words should be accurate in meaning, grammar, and word combinations or frequent collocates (Nation & Webb, 2011). To measure lexical richness, some tools that are often used are Lexical Frequency Profiles such as Tom Cobb's VocabProfile for measuring lexical frequency (Abbasian & Shiri, 2011; Cobb, 2002.; Laufer & Nation, 1995) or RANGE (Heatley *et al.*, 2002; Kyle, 2019). Additionally, concordances found in the Corpus of Contemporary American English (COCA) and Google's Ngram Viewer are often used to check collocations and multiword units. By using these tools to examine these factors in a written text, the lexical richness of said text can be determined.

Data-driven learning (DDL) and Corpus-based learning (CBL) are common approaches for increasing a language learners' lexical richness. The focus of DDL is the use of computers as a tool for language learners. CBL involves corpus-based tools, such as a concordancer, for language learning. By providing learners with examples of linguistic performance through tools such as a keyword-in-context concordance, the learners are encouraged to use their brains to decipher the correct way that language is used (Johns, 1991).

Although there are a multitude of studies conducted on lexical richness, DDL, and CBL, there is a need for more empirical data on CBL and DDL (Gaskell & Cobb, 2004; Gries, 2015). Moreover, there are a few issues that should be explored more fully. Many of the studies have focused primarily on intermediate to advanced language learners (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012; Henriksen & Danelund, 2015; Nesselhauf, 2005), with little focus on learners with lower language proficiency. Moreover, various studies have been performed in one sitting (Cobb, 2010; Gaskell & Cobb, 2004; Gilmore, 2009) as opposed to a more longitudinal design, such as across an entire semester. Additionally, most studies involve teacher-supplied resources or employ teacher-designed worksheet-based methods derived from some of the previously mentioned vocabulary tools (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012) instead of allowing the students to explore their own writing by personally using these tools. Finally, although Lexical Frequency Profiles can measure lexical richness in terms of word frequency (Abbasian & Shiri, 2011), they cannot measure the accuracy of word use in terms of grammar and semantics. Therefore, more research is required to address these issues.

As mentioned above, many researchers have made an effort to examine the vocabulary use of language learners to acquire information about the complexity of their language. They employ a vast array of tools designed to analyze and evaluate the vocabulary used by their participants. Nevertheless, equipping these learners with the very tools used by researchers might yield interesting results. By allowing the learners to analyze and evaluate their own writing through explicit instruction in utilizing some of these vocabulary tools and discussing the benefits of the data they receive, it is hoped that the lexical richness of the learners' writing may improve.

The following study attempts to contribute to the research that has been conducted on lexical richness, DDL, and CBL. It was designed to explore the effects of teaching learners to use vocabulary-related tools to increase the percentage of academic and low-frequency words in their academic research papers, thus improving the paper's lexical richness. An additional focus of the current study is to check the effectiveness of providing tools to help reduce the number of errors when the learners increase the number of academic and low-frequency words.

In order to address these issues, the following three research questions were created and answered through document and statistical analysis:

1. What, if any, is the increase in coverage of academic words after receiving explicit instruction in the use of vocabulary tools for analyzing written work?
2. What percentage of the academic and low-frequency words used to replace high-frequency words was both grammatically and semantically correct?
3. How do errors in writing before and after the treatment instruction compare in both quantity and quality?

## Literature Review

The following is a brief overview of literature related to the topic of study in this paper. A closer look will be given to studies conducted on lexical richness, DDL, and CBL.

### Lexical Richness

Several studies have provided results on the existence of lexical richness in language learners' writing. Some studies have shown that lexical richness in undergrad students reflects their

pre-existing knowledge of vocabulary (Ha, 2019). However, much of the existing literature on lexical richness shows that there is a paucity of lexical richness in second language writing. Some studies (Henriksen & Danelund, 2015) have shown that language learners with a higher level of English proficiency tend to rely heavily on comfortable and easier to use high-frequency vocabulary rather than attempting to incorporate their low-frequency vocabulary.

One criticism of how lexical richness is measured is that Lexical Frequency Profiles such as VocabProfile and RANGE only provide measures of vocabulary quantity and do not account for errors in grammatical and semantic use. Therefore, although these Lexical Frequency Profiles indicate lexical richness to some extent, they should be used in conjunction with other forms of measure (Abbasian & Shiri, 2011). Although some studies (Stæhr, 2008) use human raters to analyze student writing holistically for errors in grammar and vocabulary, many studies on lexical richness focus solely on the number and types of lexical items (Laufer & Nation, 1995; Lei & Yang, 2020).

### Data-Driven Learning (DDL) and Corpus-Based Learning (CBL)

Studies on DDL and CBL have uncovered many results concerning the use of online tools to help improve language learner writing and recognize errors. According to several researchers, DDL has successfully helped learners find their own solution to language problems using authentic resources and tools (Boulton, 2009; Cobb, 2010; Gilmore, 2009; Granger, 2012; Johns, 1991). Researchers have also discovered that CBL is useful for recognizing patterns in grammar and word use to help with error correction (Cobb, 2010; Gaskell & Cobb, 2004; Gilmore, 2009). Moreover, studies have shown that concordance information is useful for intermediate and advanced learners during writing activities (Cobb, 2010; Gaskell & Cobb, 2004).

Studies on DDL and CBL have employed a variety of methods. Given the inherent difficulty of using some of the available online tools (Lee & Lin, 2019), many of the studies have focused primarily on intermediate to advanced proficiency level language learners to lessen the cognitive load (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012; Henriksen & Danelund, 2015; Nesselhauf, 2005). Often in response to the difficulty in using some of the tools, various studies have resorted to employing teacher-supplied resources or worksheets designed by teachers that had been derived from the online vocabulary tools (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012) rather than asking the language learners to use the tools themselves. Additionally, regarding the implementation of many DDL and CBL studies, a common preference has been to conduct the research in one sitting (Cobb, 2010; Gaskell & Cobb, 2004; Gilmore, 2009) as opposed to a more longitudinal design, such as across an entire semester.

After reviewing the existing literature, several missing components were discovered. To begin with, more studies that evaluate lexical richness in terms of grammatical and semantic accuracy are needed. Additionally, there is a need for more studies on lexical richness with lower level proficiency language learners. Finally, there seems to be a need for more longitudinal studies conducted over a longer period of time that involve the use of authentic online vocabulary tools. Therefore, the current study has been designed to address these issues.

## Methodology

The following section contains an explanation of the methodology for the current study. Included are details about the context, research design, procedure, treatment, and questionnaire. It ends with

an explanation of how the data were analyzed.

## Context

Research was conducted on the written texts of 36 Japanese first-year university students, 25 women and 11 men between the ages of 18 and 20, learning academic English writing at a private Japanese university. This study was conducted in an academic reading and writing class that met thrice a week for 10 weeks during the third and final term of the learners' first year. At the beginning of the school year, the students wrote a TOEFL PBT for the purpose of placement. The scores of these learners range from 350 to 450, therefore the majority of the students in this study were in the low-intermediate level of English proficiency.

## Research Design

This mixed-methods longitudinal study was designed to analyze the lexical richness of academic papers written by lower-proficiency English learners over one entire university semester. It was designed to address some of the issues found in the literature by providing empirical evidence on lexical richness, DDL, and CBL.

## Procedure

The learners were asked to write an 8-paragraph secondary research paper comparing and contrasting a topic of their choice in bioethics. The paper was written in sections during the 10-week course. Two drafts of each section were written, with peer editing and revision conducted between each draft and section. At the beginning of the course, the students wrote two drafts of a single introduction paragraph. This was followed by two weeks devoted to the two paragraphs in Section 1. Another two weeks were spent focusing on the two paragraphs in Section 2. The final section, which comprised two main body paragraphs and the concluding paragraph, were written in two drafts over two weeks. Table 1 shows the timeline of the language learners' research paper.

**Table 1**
*Timeline of Research Paper and Treatment*

| Time | What Was Finished |
|---|---|
| Week 1 | Topic Selection and Outline |
| Week 2 | Introduction Draft 1 |
| Week 3 | Introduction Draft 2 |
| Week 4 | Section 1 Draft 1 |
| Week 5 | Section 1 Draft 2 |
| Week 6 | Section 2 Draft 1 |
| Week 7 | Treatment Instruction |
| Week 8 | Section 2 Draft 2 |
| Week 9 | Section 3 and Conclusion Draft 1 |
| Week 10 | Section 3 and Conclusion Draft 2 |

## Treatment

As shown in Table 1, the treatment occurred after the first draft of the second section, at the midway point of the paper. This ensured that the first drafts of the introduction, Section 1, and Section 2 were not affected by the treatment.

For the treatment, the participants were introduced to the Academic Word List (AWL) and asked to begin self-study of the words using various vocabulary-learning strategies such as word cards and practice exercises such as cloze sentences available on the Internet. Then, the participants were taught how to use various vocabulary-related tools to analyze their own writing. Following Nation's (2009) guidelines for training students in learning strategies, each tool, over two lessons in computer-equipped classrooms, was first modeled by the teacher. Next, the participants practiced the different steps alone, and then with partners, reporting back to the teacher when problems arose. Occasional feedback was provided by the teacher and fellow classmates in the peer editing sessions. Finally, further consultation was provided to individual learners during tutorial sessions in the teacher's office.

The learners were asked to analyze their papers with the tools and replace high-frequency words with words from the AWL or low-frequency words. The AWL (Coxhead, 2000) is a list of academic words derived from a corpus consisting of over a million words from academic texts. It contains the most frequent academic English words after West and West's (1953) General Service List (GSL) comprising the 2000 most high-frequency English headwords. The AWL encompasses 570 word families divided into 10 sublists according to frequency. Coxhead (2000) recommends that the AWL should be taught explicitly, allowing for opportunities for the vocabulary to be met in meaning-focused reading and listening texts, and used productively in speaking and writing. Meant purely as a goal to help with motivation, the students were asked to aim for 10% AWL words because according to Coxhead (2000), "The AWL accounts for 10% of the tokens in the Academic Corpus" (p. 222). Following a few tutorial sessions where some of the participants showed concern on finding enough AWL words to boost the percentage, the learners were encouraged not to ignore the low-frequency alternatives that were not present in the AWL as well.

The first tool taught to the participants was Cobb's (2002) VocabProfile, conveniently located on his website, Lextutor. The version they used categorized the first 1000 and second 1000 words from the GSL, and the AWL words in their research papers. All other words were marked as "Off-list" (i.e., low-frequency words, proper nouns, non-English words, spelling errors). Once they had learned to identify the high-frequency words, the participants were shown how to use the right-click functions of both Google Docs and Microsoft Word that provide possible synonyms for the highlighted word in question. The participants were warned that, even though a potential replacement might have been located, the word had to match the sentence in both grammar and meaning. To help check the replacement words in the sentences, the participants were taught how to use the Google Ngram Viewer (Michel *et al.*, 2010) for checking which word combinations are used most often and the color-coded keyword-in-context concordance on the COCA website (Davies, 2010; Johns, 1991) for checking possible collocations and recognizing possible patterns and how others use the language. With these tools, the participants set out to improve the lexical richness of their research papers.

## Questionnaire

On the final day of the course, when the final drafts of the research papers were due, the

participants were asked to complete a questionnaire about the tools they did or did not use while improving the lexical richness of their writing. The timing of the questionnaire completion was chosen for maximum effect, as the probability of the participants using the tools prior to the class in order to complete the assignment was high and therefore fresh in their minds. The questionnaire, as seen in Appendix A, consisted of seven multiple-choice questions and one open-ended question for qualitative purposes. The multiple-choice questions were asked to confirm which tools the students used and found most useful and easiest to use and to learn if the students used other tools that were not discussed in class. The open-ended question asked what was most difficult about changing the high-frequency words to more academic or low-frequency words. Thus, a complete picture of the tools that were used was obtained.

## Data Analysis

There were three sources of data used in this study. The first was the research papers written by the participants, 72 in total, with 36 first drafts and 36 final drafts. The next available data were from the questionnaire. The last source of data was individual sentences randomly chosen from the papers that contained an AWL or low-frequency replacement word.

Using the steps involved in measuring lexical richness from Nation and Webb (2011, p. 256), the following decisions were made:

***Steps Involved in Measuring Lexical Richness***

1. Decide on the text to be analyzed (research papers)
2. Decide on the unit of counting (word families)
3. Decide what to do with errors (compare the original and replacements words)
4. Decide on how to measure lexical richness (multiple human raters) (p. 256)

For analyzing data to help answer the first research question concerning the coverage of AWL words, the first and final drafts of all the participants were compared. To ensure that only the participants' words were being analyzed, all direct quotes and proper nouns were removed from each draft. Word family counts for the AWL words were obtained from the RANGE program (Heatley, Nation, & Coxhead, 2002). According to Durrant and Schmitt (2009), much can be learned by examining and comparing individual scores, as these results are often hidden when only entire corpora are compared. Therefore, the first drafts of each participant were compared with their final drafts using a paired-samples *t*-test.

With regard to the data analysis used to answer the second research question regarding the grammatical and semantic fit of the replacement words, 65 sentences from 20 participants were randomly chosen from the fourth paragraph of the first and final drafts that contained academic or low-frequency replacements of high-frequency words. The fourth paragraph was chosen as it came from Section 2 of the paper, where the treatment occurred, thus quite possibly representing the best effort by the participants in using the tools learned during the treatment. Each word replacement was presented in its original sentence and rated on a four-point scale as seen below. The capitalized word was the low-frequency replacement word. The word in brackets was the high-frequency word from the first draft.

Example:

No prospect of UTILIZATION (using) alternative ways has yet emerged.

- both grammar and word choice are correct
- correct grammar, problems with word choice
- correct word choice, problems with grammar
- both grammar and word choice are incorrect

Each of the 65 sentences was rated by at least three of six human raters, all members of the same applied linguistics doctoral cohort as the researcher. The raters were trained as a group to help ensure inter-rater reliability. Inter-rater reliability was calculated by dividing the number of agreed-upon items with the total number of items.

For the final research question concerning the quantity and quality of errors between the two drafts, the same human raters and 65 sentences were used. Following each sentence, the raters were asked to compare and assess each original and replacement word to determine which was more correct using the following four-point scale.

Example of rating scale:

For the previous sentence, which word was most correct?

- Both are equally correct
- Word in ALL CAPS
- Word in (brackets)
- Neither

## Results

The following are the results for this study. Discussion of the results is presented in the Conclusions section.

### Questionnaire

Table 2 presents the results from the questionnaire that was completed by the participants. As can be seen, nearly all participants used VocabProfile to analyze the word frequency of their papers, as opposed to those few who reported using Ngram or the concordance.

**Table 2**
*Data from Questionnaire*

| | Used | Most Useful | Easiest | More Instruction | Most Problematic |
|---|---|---|---|---|---|
| VocabProfile | 97.2% | 83.3% | 47.2% | 50.0% | Matching meaning |
| COCA | 11.1% | 2.8% | 2.8% | 38.8% | 41.6% |
| Ngram | 13.8% | 2.8% | 2.8% | 22.2% | Matching grammar |
| MSWord Syn | 66.7% | 27.7% | 41.6% | 16.6% | 16.6% |
| Google Syn | 36.1% | 13.8% | 5.5% | 2.8% | Collocations |
| Grammarly | 88.8% | 5.5% | 13.8% | 0.0% | 5.5% |
| Dictionary | 11.1% | 13.8% | 0.0% | 0.0% | |

*Note.* Grammarly and dictionary use were not taught by the teacher for this study.

### Texts Analyzed

Table 3 presents the descriptive data of the texts that were analyzed. The numbers reflect the number of words after removing the direct quotes and proper nouns.

**Table 3**
*Descriptive Data of Texts Analyzed (Adapted from Durrant & Schitt, 2009)*

| Description | Drafts | Number of Texts | Number of Writers | Total Words | Mean Words/ Text | Writers' L1 |
|---|---|---|---|---|---|---|
| Academic argumentative | First | 36 | 36 | 46,684 | 1,297 | Japanese |
| secondary research papers written by first-year Japanese university students with lower-intermediate English proficiency studying English academic reading and writing Paper topic - bioethics | First | 36 | 36 | 55,255 | 1,534 | Japanese |

## Individual Scores Between the First and Final Drafts

A paired-samples *t*-test was conducted to evaluate whether the number of AWL word families would increase between the first and final drafts after a treatment on the use of vocabulary tools. The results indicated that the mean AWL count for the final draft ($M$ = 61.86, $SD$ = 25.68) was significantly greater than the mean AWL count for the first draft ($M$ = 42.78, $SD$ = 20.11), $t(35) = -3.51$, $p = .0004$. The standardized effect size index, $d$, was .83. The 95% confidence interval for the mean difference between the two ratings was 35.97-70.55 (Green & Salkind, 2013).

## Percentage of Grammatically and Semantically Correct Replacements

Table 4 displays the percentages assigned to the replacements as determined by the human raters. The majority ruling for the rating of each replacement word was used in the analysis. The inter-rater reliability was 86%. It should be noted that the highest percentage was obtained by word choices that were both grammatically and semantically correct.

**Table 4**
*Analysis of the Replacement Words*

| Grading Criteria | Percentages |
|---|---|
| Both grammar and word choices are correct | 56.9% |
| Correct grammar, problems with word choice | 35.4% |
| Correct word choice, problems with grammar | 6.2% |
| Both grammar and word choice are incorrect | 1.5% |

## Comparison of Errors

Finally, a comparison between the original high-frequency word used in the first draft was compared with the replacement academic or low-frequency word to determine which word, if any, was more correct. Table 5 exhibits the percentages of the error judgments made by the raters. The scores used for the analysis were the ratings that received the majority vote from the raters for each

pair of words that were analyzed. The rater reliability for this was also 86%. Of particular interest is that 90.9% of the replacements were considered to be more correct by the human raters.

**Table 5**
*Analysis of the Errors*

| Grading Criteria | Percentages |
|---|---|
| Both are equally correct | 55.4% |
| The academic replacement is correct | 35.4% |
| The original high-frequency word is correct | 7.7% |
| Neither is correct | 1.5% |

## Conclusions

This discussion is presented to help analyze the results, situate the findings in the existing literature, and answer the following three research questions: 1) What, if any, is the increase in coverage of academic words after receiving explicit instruction in the use of vocabulary tools for analyzing written work? 2) What percentage of the academic and low-frequency words used to replace high-frequency words was both grammatically and semantically correct? 3) How do errors in the first and final drafts compare in both quantity and quality?

### Findings

Although many of the new tools, such as Ngram and the COCA, were not used by the majority of the participants, as indicated in Table 2, all but one participant used the VocabProfile. The fact that the more difficult Ngram and COCA were underutilized coincides with the findings from Lee and Lin (2019). In addition, many students also used the synonym functions in Microsoft Word and Google Docs. Therefore, as shown by the results of the *t*-test, there was a significant increase in the coverage of academic words from the first to the final draft. As there was no control group, it cannot be definitely said that the increase in coverage was due solely to the explicit instruction of the vocabulary tools. However, the learners appear to have gained awareness of the frequency level of their vocabulary just by using VocabProfile. Should they continue to use the tools, it could potentially increase their lexical richness in future papers.

Additionally, the second research question was created to address issues about the lack of attention given to the accuracy of vocabulary use especially regarding grammar and meaning, as discussed in studies such as that of Abbasian and Shiri (2011). Human raters analyzed the language learners' written work, similarly to Stæhr (2008). Ultimately, 56.9% of the low-frequency replacements were deemed both grammatically and semantically correct by the raters. Furthermore, the results from the raters coincided with the qualitative data collected from the questionnaire, as can be seen in Table 2. At 41.6%, the participants reported that matching the meaning of the replacement word to the original was the most difficult, whereas only 16.6% deemed grammar as the most difficult factor. When compared with the results in Table 4, the sample replacement words had 35.4% errors in word choice as opposed to 6.2% grammar errors. This might indicate a need for further instruction in the Ngram and COCA concordance tools, and it also coincides with Lee and Lin's (2019) findings.

Concerning the third research question, the raters determined that the replacement words, on average, were more correct than the original high-frequency words. This might indicate that the

tools were useful for replacing the high-frequency words, as they possibly provided the learners with opportunities to analyze more closely the grammar and meaning of the words being used than they normally would have. These results are consistent with some of the previous studies (Boulton, 2009; Cobb, 2010; Gilmore, 2009; Granger, 2012; Johns, 1991).

## Implications

This study provides more empirical data on lexical richness, CBL, and DDL that might help fill the gaps in the literature discussed earlier. It is also a longitudinal study across an entire semester on learners with lower English proficiency. However, the most important point about this study is that it could possibly show the potential involved in encouraging the learner to use some of the vocabulary tools generally used by teachers and researchers, especially VocabProfile. Because of the improvements in lexical richness that occurred through the use of these vocabulary tools, teachers might consider training their students to use vocabulary analysis tools more actively as part of their students' writing process.

## Limitations and Future Study Recommendations

There are a few limitations and possible recommendations for future research that arose during this study and should be addressed.

First, the learners were also asked to use low-frequency words that were off-list from the GSL and AWL as replacements. With the removal of proper nouns from the participants' research papers, the remaining off-list words are presumably all low-frequency words. However, both RANGE and VocabProfile do not report the word family count for these off-list words. Moreover, according to the raw data, all 36 participants' off-list word counts increased, at least as was indicated by the word types percentage in RANGE. Thus, processing the two drafts of the papers through the British National Corpus (BNC) version of RANGE to check for the frequency of the words used between the 1st -14th 1000 words in English was considered. However, it should be noted that the GSL and the BNC are not completely analogous, thus introducing a limitation to the analysis of this study. The GSL/AWL were chosen for their ease of comprehension and processing for the participants. Future researchers might consider using the BNC for obtaining a more accurate account of the increase in low-frequency word replacements.

Additionally, because of the low percentage of use with the Ngram and COCA for checking the appropriateness of the replacement words, and as indicated on the questionnaire, more instruction should be given in the use of these valuable tools.

Furthermore, there is the possibility of distorted results from participants who did not use all of the tools (Gilmore, 2009). However, a counter to this was attempted by comparing individual scores (Durrant & Schmitt, 2009; Granger, 2012).

Moreover, further confounding variables, such as the effects of peer editing, and other writing tools not taught in the treatment, such as Grammarly, should be isolated or controlled for in future studies.

In addition, as Gilmore (2009) explained, there is a need for a control group. This study did not have a control group. Potentially, a similar writing assignment from the previous year's cohort at the same university from the same academic reading and writing course could have been used as a control group for this study. However, due to time constraints and the need for obtaining the

permission of the previous year's students, this could not be accomplished.

Finally, the placement of the treatment was handled acceptably for this particular research paper. However, as this particular research paper was written in sections, the first drafts of Section 3 and the concluding paragraphs quite possibly were affected by the treatment. Therefore, future studies might elicit purer results if the treatment is placed between the first and second drafts of an entire paper or essay.

## References


Abbasian, G. R., & Shiri, P. M. (2011). Validation of lexical frequency profiles as a measure of lexical richness in written discourse. *Technology of Education Journal, 5*(2), 125-131.

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *Recall*, *21*(1), 37-54.

Cobb, T. (2010). Instructional uses of linguistic technologies. *Procedia-Social and Behavioral Sciences*, *3*, 14-23.

Cobb, T. (2002). Web VocabProfile (an adaptation of Heatley, Nation & Coxhead's (2002) Range). *Lextutor.* Retrieved on Feb 12, 2022 from http://www.lextutor.ca/vp/.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213-238.

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing, 25*(4), 447-464.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL-International Review of Applied Linguistics in Language Teaching*, *47*(2), 157-177.

Kyle, K. (2019). Measuring lexical richness. In S. Webb (ed.), *The Routledge handbook of vocabulary studies* (pp. 454-476). Routledge.

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, *32*(3), 301-319.

Gilmore, A. (2009). Using online corpora to develop students' writing skills. *ELT Journal*, *63*(4), 363-372.

Granger, S. (2012). Learner corpora. In C. A. Chapelle (ed.), *The Encyclopedia of Applied Linguistics* (pp. 3235–3242). Wiley-Blackwell.

Green, S. B., & Salkind, N. J. (2013). *Using SPSS for Windows and Macintosh* (7 edition). Pearson.

Gries, S. (2015). Statistical methods in learner corpus research. In G. Gaëtanelle, S. Granger, & F. Meunier (eds.), *The Cambridge handbook of learner corpus research* (pp. 59-181). Cambridge University Press.

Ha, H. S. (2019). Lexical richness in EFL undergraduate students' academic writing. *English Teaching*, *74*(3), 3-28. doi:10.15858/engtea.74.3.201909.3

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. From http://www.victoria.ac.nz/lals/staff/paul-nation.aspx.

Henriksen, B., & Danelund, L. (2015). Chapter two studies of Danish L2 learners' vocabulary knowledge and the lexical richness of their written production in English. *Lexical Issues in L2 Writing*, *29*.

Johns, T. (1991). *Should you be persuaded: Two samples of data-driven learning materials* (pp. 1-13). PDF file. https://lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*(3), 307-322.

Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, *81*, 14-25. https://doi.org/10.1016/j.system.2018.12.011

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. https://doi.org/10.1016/j.jeap.2020.100894

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Brockman, W., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2010). Quantitative analysis of culture using millions of digitized books. *Science, 331*(6014). doi:10.1126/science.1199644

Nation, I. S. P. (2009). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.

Nesselhauf, N. (2005). *Collocations in a learner corpus* (Vol. 14). John Benjamins Publishing.

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139-152. doi:10.1080/09571730802389975

West, M., & West, M. P. (Eds.). (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. Addison-Wesley Longman Limited.

# Appendix A
# Vocabulary Questionnaire

Please answer these questions HONESTLY about your research paper and vocabulary. You may check more than one answer if you need to:

1. Which vocabulary tools did you use?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms (right click and define)
   - ☐ Microsoft Word synonyms (right click and synonym)

2. Did you use any other tools I didn't teach you?
   - ☐ Dictionary – what kind? _____
   - ☐ Grammarly
   - ☐ Only my brain
   - ☐ Other(s) _____

3. Which tool(s) were the most useful?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms
   - ☐ Microsoft Word synonyms
   - ☐ Other(s) _____

4. Which tool(s) were the easiest to use?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms
   - ☐ Microsoft Word synonyms
   - ☐ Other(s) _____

5. Which tool(s) would you have liked more instructions/directions from the teacher?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms
   - ☐ Microsoft Word synonyms
   - ☐ Other(s) _____

6. Did your peer editors help with your academic words?

☐ Yes
☐ No
☐ I don't know

7. Did you help your writing group members with their academic words?

☐ Yes
☐ No

8. What was most difficult about changing the vocabulary to more academic words?