

### The Use of Corpus Technology as a Tool in Students' Word Choice in Academic Writing

Yuka Nohira<sup>1</sup>, and Aileen Bautista<sup>2,\*</sup> <sup>1</sup> FEU High School <sup>2</sup> De La Salle University - Manila \*Corresponding Author: aileen.bautista@dlsu.edu.ph

Abstract: Several scholars have examined how corpus tools could help students improve their writing, especially vocabulary. This action research examines whether Data-Driven Learning (DDL) corpus training and instruction improve senior high school students' vocabulary and word choice in their academic writing class. This was based on the needs assessment, which revealed that students struggle with word choice. Ten students were instructed to use the optional Corpus of Contemporary American English (COCA) during writing sessions to address the said issue, while the tool's impact on students' writing word choice was assessed during vocabulary self-evaluation. Also, participants' views on using the corpus for word choice correction were collected through post-project survey questionnaires and interviews. The results showed that there was no statistical significance in students' vocabulary performance after the use of COCA. However, there was an observed enhancement in word choice based on their writing activities and outputs. Additionally, participants held positive views about using corpus in academic writing, especially if it were developed further in terms of its user-friendly interface and more convenient navigation features. Thus, this study recommends employing corpus as a supplementary tool in writing classes that merits further research.

**Key Words:** corpus technology; COCA; data-driven learning; academic writing; action research

#### 1. INTRODUCTION

The Department of Education (2022), through DepEd Order No. 034, S. 2022, mandated that classes may be held online or in person, depending on the circumstances. This resulted in bolstering the online teaching and learning experience. Hence, online language classes must set realistic goals for gaining language skills. (Bailey & Lee, 2020). However, there are language skills that are more difficult to manage in an online learning environment than others, especially pronunciation and writing (Robinson et al., 2016). As a result, teaching and studying writing has become more challenging due to several complexities such as vocabulary, grammar, spelling, concept structure, and coherence (Ardianti et al., 2021).

Senior high school (SHS) students grapple with academic writing vocabulary and grammar (Urbando et al., 2021). Urbano et al. reported that students still struggle with academic vocabulary and error correction despite studying English courses in the new Enhanced Basic Education Curriculum (K-12) such as Reading and Writing Skills, EAPP, and Practical Research. It is consistent with the findings of Pablo and Lasaten (2018)



that SHS students' academic essays range from low to satisfactory, with errors and issues about lack of variety of ideas, connectives, organization, and language or word choice. This suggests that this area, along with the online teaching and learning experience, is worth exploring in action research

The needs assessment that we conducted revealed that students experience challenges in writing, especially in the accuracy of language. The Writing Needs Questionnaire (Urbano et al., 2021), interviews, diagnostic essay paragraph and sentence-level mistake evaluations, and classroom observations informed this. Additionally, the results indicated that academic writing's reliance on receptive vocabulary is the primary source of this difficulty. Meaning, that the issue of vocabulary impacts the written outputs. Thus, in the following parts of this action research, we will discuss the proposed implementation or instructional approach to address the issue: the use of corpus technology. In particular, the study examines the "Corpus of Contemporary American English" tool. This paper is outlined as follows: 1) background of corpus-based teaching approach in vocabulary and writing, 2) statement of the problem, 3) methodology, 4) results and discussion, and 5) conclusion.

# 1.1 Corpus-based teaching approach in Vocabulary

Corpora has become increasingly popular in studying vocabulary due to advances in Computer Assisted Language Learning (CALL) research and vocabulary teaching (Liu, 2013). The concordance tool allows teachers to discuss words, phrases, and collocations to help students uncover linguistic patterns because corpora offer various samples and linguistic variations of an item searched in its context (Reppen, 2011). According to some scholars, like Paker and Ergül Özcan (2017), corpus-based vocabulary activities outperform textbooks and dictionaries. Meanwhile, Binkai (2012) hypothesized that corpus technology in language learning increases vocabulary and promotes independent learning by giving students greater control of their own learning.

However, some experts, including Sinha (2021), indicate pupils hold varying perspectives on corpus' effectiveness in vocabulary learning. Some students expressed that corpus technology is time-consuming and difficult to analyze. According to Çalışkan and Kuru Gönen (2018), corpus technology is primarily beneficial to teachers. Thus, studies imply that corpus training is essential and is a crucial part of making the use of corpus technology in the classroom more effective.

#### 1.2 Statement of the Problem

This action research investigates whether corpus technology can benefit students with word choice in writing. The study also examined students' use of corpus technology in writing classes. Specifically, it seeks to answer these questions:

- 1. Is there a noticeable difference between the SHS students' self-evaluations of their vocabulary or word choice before and after the corpus-based technology was introduced as a tool in their during-writing phase?
- 2. Do the SHS students in the writing course find the corpus technology helpful in addressing their writing issues with word choice?

#### 1.3 Context

The research was conducted at a private school in Manila in the English for Academic and Professional Services (EAPS) class. A cooperating teacher oversaw it, while a pre-service teacher conducted her practicum and undergraduate thesis. Thus, the study was part of the practicum and the thesis writing of the student. The study was conducted when the pandemic was still in full swing in 2022, so the lessons were held online. The students were enrolled in the course, English for Academic and Professional Services (EAPS), which required weekly synchronous lecture-discussion, consultations, and draft writing. EAPS has a pre-requisite course: Reading and Writing Skills (RWS). The school also offers fourteen weeks of sessions; thus, the proposed technique was shorter than other schools' because the pre-service teacher had to prepare and defend her thesis in the final weeks.

Pilot testing was also conducted in a different class, which has one (1) corpus training session and two (2) separate discussions.

As for the students, who are the research participants, not all decided to participate in the action research. Only 10 students signed informed consent papers and finished all corpus training and investigation tasks. Importantly, some implementation sessions took place during asynchronous sessions, not Zoom class sessions.

## 2. METHODOLOGY

2.1. Intervention Plan

This study used the Corpus of Contemporary American English as its intervention tool. Davies (2010) said this is the first balanced and credible English monitor corpus. Today, COCA is the most comprehensive and equitable free corpus of English, with over one billion words in its database from 1990 to 2019 (Davies, 2020). By visiting the website address http://corpus.byu.edu/coca/, users may access the corpus.

The researchers, particularly the pre-service teacher, became familiar with COCA. User registration was required to learn how to use it. Since the school does not have an institutional subscription, the COCA used for this action research was a free account. Additionally, COCA features four main tabs: 1) SEARCH, 2) FREQUENCY, 3) CONTEXT, and 4) OVERVIEW. Therefore, we ensured that the pre-service teacher implementing the strategy was proficient in using COCA and its four main tabs.

Then, during the first training day, students were required to register for COCA's free account. In the early weeks of the study, students were trained on COCA's input page search tools (List, Chart, Collocates, Compare, and KWIC) before using this corpus in their work. Five (5) 30-to-40-minute corpus training sessions were undertaken for this investigation. These sessions happened during the cooperating teacher's asynchronous sessions, so the respondents were the only ones attending these corpus training sessions. These were also recorded in case they missed the meeting or could not understand the training sessions. Five sessions may not be enough to apply all COCA resources. Due to the cooperating teacher's asynchronous sessions, participants found it difficult to attend, therefore, this was an adequate amount of time for training. They may have also found it overwhelming to receive too much lecture-based direction on how and what to search for in the corpus, leaving them little freedom to develop their own query methods and skills to match their needs. As for vocabulary instruction, the study used an incidental and independent strategy development approach in which students did not receive direct instruction and discussion on vocabulary (explicit vocabulary learning). Rather, inferring, analyzing, and understanding word meanings from contexts through corpus use and linguistic activities like writing were utilized.

After immersing themselves in it, students were instructed to use the corpus throughout their writing: pre-writing, while-writing, and revision. Students' Google Document essay drafts were stored in a Google Drive for teachers. The researcher and teacher noted and commented on student draft word choice errors using Google Documents. Students checked them using COCA to fix issues on their own time before the

#### 2.2. Instruments

submission date.

This study employed questionnaires, interviews, observations, and other documentation. These were triangulated for reliability.

This study's 12-item self-assessment is based on the Nation's (2001) list of key vocabulary knowledge factors. This was adopted because it helps learners understand their learning process by focusing on performance rather than language learning. By identifying their own weaknesses in word choice in writing, they were able to know the areas they needed to improve (Mican & Medina, 2015) and became less anxious about performing the skill being assessed (de Saint Léger, 2009). To ensure item reliability, Cronbach's alpha of the pre-and post-test was assessed and yielded 0.81 and 0.91, respectively, demonstrating strong internal consistency.

Before and after the intervention, students received a Google Forms pre- and post-project questionnaire. The pre-project questionnaire asked eight questions on student demographics and corpora/corpus knowledge. The survey was modified from the instrument of Al-Qahtani (2021). However, the post-project questionnaire included 20 five-point Likert scale items from Chatpunnarangsee (2013). This questionnaire also asked students about using corpus consultation to increase their writing vocabulary (word choice).

After data collection, an interview was carried out. This study employed the interview method to understand teacher and student perspectives on corpus technology and web-concordancer use. Ten to fifteen questions were asked about their experiences throughout the intervention and how to improve their academic essay word choice.

This study collected two primary documents: the corpus search logs and students' written outputs. Students submitted their corpus search logs and revised writing assignments. To track student corpus entries while writing and revising, a search log was needed. Documents were collected to support or refute interview and survey results, justifying their use.

Field notes were employed during synchronous and asynchronous semi-structured observations. Students' replies during Zoom synchronous classes and Google document activity were observed since they could not be observed while composing their papers. Thus, the observation focused on how students corrected their writing with the corpus.

#### 2.3. Data Analysis

Quantitative data analysis was intended for the pre and post-project questionnaire and vocabulary self-assessment of this study. Descriptive statistics were conducted for the pre-project questionnaire to present the categorical variables before the intervention (e.g., age, English proficiency, the number of years spent learning English, knowledge of computers, and corpus). This involved the computation of the mean and standard deviations. After the implementation, a post-project questionnaire was conducted and evaluated using descriptive statistics. This questionnaire was designed to gather information on three key categories: the advantages of the corpus, issues with corpus use, and overall corpus use experience. In presenting these data, the researchers tallied the pre and post-survey results, and graphs and tables were used to present and analyze the data.

Descriptive statistics were also conducted for the vocabulary self-assessment, obtaining the sum, mean, and standard deviations for each level of each test. Moreover, it will be analyzed using the Wilcoxon Signed Rank Test. This test is the non-parametric counterpart of the Paired t-test. Since there were only ten (10) participants in the study and the data were drawn from similar samples, the Wilcoxon Signed Rank Test was utilized.

Semi-structured interview data was transcribed and analyzed using a similar content analysis of Erlingsson and Brysiewicz, 2017) that we also used in our needs analysis. The steps involved 1) familiarization, 2) condensing meaning units, 3) coding, and 4) generating categories and themes. To ensure inter-rater reliability, the pre-service teacher/researcher and another coder coded the interview data separately. Through discussion, the two raters compared, evaluated, and consolidated the codes into one set. Moreover, the students interviewed were anonymized.

Meanwhile, observation notes were used as baseline data analysis for the writing course and during intervention implementation. This was meant to confirm or refute evidence from past studies. It also involved reviewing participant records, including corpus search logs. This tool examined the frequency, search words, purpose, and results of corpus use. It demonstrated the linguistic aspects students struggled with in their writing and how they used corpora, emphasizing form, context, or both (Yoon, 2005).

#### 3. RESULTS AND DISCUSSION

#### 3.1 Story and outcomes

Students' view of corpus technology

The pre-test (M = 3.94, SD = 0.48) and post-test (M = 4.09, SD = 0.46) scores before and after the intervention program showed no significant difference since the critical value (CV = 8) is less than the test statistics (t = 9). This suggests that the respondents did not see any changes in their vocabulary knowledge or word choice after using the corpus to enhance their word choice in their academic writing. This is similar to Fang et al. (2021), who found no significant difference in error frequency between pre- and post-writing assignments and only slight improvements.

The small sample size (just 7 participants) and three corpus training sessions may explain this outcome. Researchers recommend corpus training sessions lasting 10–15 weeks or longer to guarantee students have completely mastered and navigated the corpus. The corpus training sessions in these investigations lasted one to two hours (Ashouri & Mashhadi Heidar, 2015; Li, 2017).

Students used COCA for six weeks in this study. This study limited corpus training sessions to 30–45 minutes. The shorter duration compared to previous studies may not guarantee full student immersion in the suggested instrument or intervention (Dolgova & Mueller, 2019), and not all students attended all corpus training sessions performed by the preservice teacher.

Students' experiences using corpus technology

However, despite the results having no significance, COCA caused minor word choice revisions. Students were observed to have altered several words or expressions because COCA declared them improper or rare in academic contexts. For instance, using COCA's word search function, a student altered the verb "move" in his paper to "transport". He replaced "move" with "transport," which is more typical in academic writing, after finding that it is only used in fiction and TV/movies.

Furthermore, another student used Cterms, whichOCA to update repeated phrases in the paper by finding their synonyms. She changed "opportunities" in a few parts of the article due to repetition. She found the most common synonyms for "opportunities" using the equal sign (=) in the List search function: "chances," "prospects," and "occasions."

Students are also more conscious of terms near the target word. In the first COCA self-editing task, one participant replaced "make strategic life choices" with "make \* life decisions." The asterisk sign (\*) indicates that the user is seeking a frequent term or the best collocates for the entry. Because COCA found no word combination or collocation for "strategic life choices," the student changed it to "make consequential life choices".

Some students searched COCA's Parts of Speech (POS / \_pos) function for word substitutions. The Parts of Speech feature is similar to the asterisk sign (\*) keyword but is more specified and limited to the chosen parts of speech. The difference between *POS* and \_*pos* is that the former searches for any word that matches this part of speech (e.g., rough NOUN) while the latter searches for words that have more than one part of speech (e.g., claim\_nn, claim as a noun; claim\_v, claim as a verb). Students generally used POS rather than \_pos in this study to find a better word in specific parts of speech paired with another word.

For instance, a participant's initial output was "utmost priorities," an adjective collocation (utmost adjective, priorities - noun). Using the List search tool with the terms "ADJ priorities" (adjective as the POS), it was observed that "top" is more suited before "priorities."

Some participants also searched for adjacent words using Collocates. The Collocates search function is similar to the POS functionality, except COCA users can choose their target parts of speech and where to place them. For example, the student was unsure if she used the right verb for "instrument," so she prompted the concordancer to check for the VERB collocates of "instrument" two-word strings before it. She revised this expression later on into "*developing* an instrument" because she based it on the frequency revealed to her by COCA

Overall, after navigating the corpus, participants experienced fewer word choice and combination challenges. They also altered some terms, which COCA found unsuitable for academic contexts, making their views more formal and academic.

### Comparing The Use of a Corpus to Other Resource Tools

The data suggests that SHS students would likely compare concordancer (COCA) characteristics and capabilities to dictionaries and thesaurus. The interviewed students expressed that dictionaries gave them "direct meaning while the corpus provides more examples." They quickly find the definition of a vocabulary term in a dictionary but with few examples. However, COCA, as a concordancer, provides many examples from varied contexts and self-analysis. COCA is considered a superior resource compared to dictionaries due to its extensive search options, concordance lines with examples, and the ability to connect with external websites or resources (Mueller & Jacobsen, 2016; Quin, 2015).

However, other participants had diverging views on the comparison of the corpus to other resource tools. Essentially, students acknowledge the significance of referring to a corpus, but they prefer using a dictionary that offers immediate and specific information. This is true regardless of whether these terms are correctly paired or suitably employed in a particular register, such as an academic environment. Data also showed that a corpus is a tool that may be used by highly proficient students because it requires a prior understanding of particular words. However, students who are less proficient in English may find the corpus impractical.

However, several recent studies found that "substantial scaffolding and assistance" can make a corpus-based approach effective for low to intermediate-level learners (Vyatkina, 2016; Saeedakhtar et al., 2020), even though high-proficient learners benefit more. Thus, participants may have struggled due to a lack of corpus-research abilities and training rather than linguistic competency.

Based on the interview data, a comparison was made between the corpus and other resource aids such as dictionaries, thesaurus, and Google extensions. However, the survey results showed that six participants found corpus to be more helpful than dictionaries. The other four (4) participants disagreed or were neutral in the statement. As illustrated, this item gained a mean score (M) of 3.7, while the degree of scatteredness (SD)of the scores is 1.42. The standard deviation of 1.42 indicates that the individual replies deviated slightly more than one point from the mean. This suggests that the data points are distributed across several scales. This indicates that students hold different perspectives on the comparative usefulness of corpora and dictionaries. The variability of the findings may be attributed to the frequency of their attendance at corpus training sessions and their proficiency in comprehending the outcomes in COCA.

One explanation for the disagreements is that corpus data does not immediately reveal phrase meanings. Unlike dictionaries, users can use a corpus or concordancer to find the meaning of an unfamiliar phrase by looking at examples in the concordance lines (Xiaoli & Altunel, 2018). Since some students failed to recognize the meaning of a word by reading the complete text from the concordance lines, context hints may have helped.

The corpus can disclose students' shortcomings, but internalization of corpus-derived standards may depend on several elements for successful corpus data processing. Park (2012) stated that "learner's achievements depend both on their ability to interpret and exploit the search results" (p. 361). Thus, learners profit from corpus use only when they "devote sufficient time and effort to evaluate search results based on careful analysis" (p. 381). Non-self-analysis and appraisal rarely yield positive results.

Students observed that corpora, dictionaries, and other resource tools were comparable and different. They see the potential of the corpus as being newly introduced to them, but they find it difficult to navigate at first because of its complexities in typing in keywords and different search functions, unlike other resource tools that are direct. In terms of the difficulties of using corpus, recent research has found that using corpus can be challenging for some students (Hirata et al., 2013; Leńko-Szymańska, 2015; Oktavianti et al., 2022). According to the current study, corpus use is difficult due to unfamiliarity compared to long-term dictionary use.

#### 3.1 Self-reflection and learning

The COVID-19 pandemic has significantly influenced educational environments, compelling Filipino teachers to adapt to different learning management systems and utilize the Internet to access a wide range of information and tools.

Our experience, both as cooperating and pre-service teachers, allowed us to see the various advantages and challenges of integrating data-driven learning (DDL) into a computer-assisted language learning (CALL) classroom that used to be contained by Microsoft software only, such as PowerPoint. Thus, exposing corpus technology as part of the writing classes was generally overwhelming. It was exciting to see how much COCA works more effectively than a dictionary, but it was also frustrating when we experienced the limitations of a free account in COCA.

We also realized how the use of corpus technology tries to change the perspective of the learners, as using COCA, they become less dependent on the teachers' feedback, which leads to spoonfeeding. However, the learning experience allowed us to appreciate that a computer-assisted language learning environment facilitates learner independence. It may not be the current preference of students, especially in writing, but DDL seems to be a promising avenue to explore, as hybrid classrooms are meant to stay.

#### 4. CONCLUSIONS

The present study aimed to investigate and explore whether SHS students could improve their vocabulary or word choice in their academic writing outputs after completing corpus training and instruction under the DDL framework. Results showed that students perceived that there were no major differences in their vocabulary after being exposed to the corpus for six (6) weeks. This result might be compounded by the fact that only five (5) corpus training sessions were held with ten (10) student participants, unlike previous studies with large sample sizes and long hours and sessions of exposure to the corpus. However, as observed in participants' writing outputs and corpus activities, most of them have made

DLSU Research Congress 2024 De La Salle University, Manila, Philippines June 20 to 22, 2024

minor improvements in various important aspects of vocabulary. They became aware of 1) the appropriate use of a word in context, especially in the academic setting; 2) synonyms of the word to avoid redundancy, and 3) the word combinations or words beside / nearby (before and after) the target word, also known as collocates.

The findings indicated that students generally have positive responses to utilizing a corpus or a concordancer when self-editing their academic papers. This suggests that artificial intelligence in a language class may improve students' vocabulary if teachers help them learn and use a technological tool's online interface. As a challenge, teachers ought to model the usage of AI and basic technologies like online corpora with pupils. Through this, the students will see that using these alone does not suffice, as they are just tools in aiding the learning experience.

#### 6. REFERENCES

Al-Qahtani, A. A. (2021). The Integration of Corpus-Based Approach with EFL Academic Writing in the Saudi Context. International Journal of English Language and Literature Studies, 10(1), 22–45.

https://doi.org/10.18488/journal.23.2021.101.22.45 Ardianti, N. P. E., Nitiasih, P. K., Budiarta, L. G. R., &

- Adnyayanti, N. L. P. E. (2021). Enhancing Students' Writing Skill with Problem-based Learning Video on Pandemic Situation. *Jurnal Ilmiah Pendidikan Profesi Guru*, 4(2). https://doi.org/10.23887/jippg.v4i2.34815
- Ashouri, S., & Mashhadi Heidar, D. (2015). The impact of teaching corpus-based collocation on EFL learners' writing ability. *International Journal of Foreign Language Teaching and Research*, 3(10), 53-62. http://jfl.iaun.ac.ir/article\_563502\_7fc3999a90dc0ed1e7 2159a4bb655627.pdf
- Bailey, D., & Lee, A. (2020). Learning from experience in the midst of Covid-19: benefits, challenges, and strategies in online teaching. *Computer-Assisted Language Learning Electronic Journal*, 21(2), 178-198. <u>http://callej.org/journal/21-2/Bailey-Lee2020.pdf</u>
- Binkai, J. (2012). An Empirical Study on Corpus-Driven English Vocabulary Learning in China. *English Language Teaching*, 5 (4), 131-137. <u>https://files.eric.ed.gov/fulltext/EJ1079027.pdf</u>
- Bridle, M. (2019). Learner use of a corpus as a reference tool in error correction: Factors influencing consultation and success. *Journal of English for*

Academic Purposes, 37, 52-69. https://doi.org/10.1016/j.jeap.2018.11.003

Çalışkan, G. & Kuru Gönen, S. İ. (2018). Training teachers on corpus-based language pedagogy: Perceptions on using concordance lines in vocabulary instruction. *Journal of Language and Linguistic Studies*, 14 (4), 190-210.

https://dergipark.org.tr/en/pub/jlls/issue/43366/528108

- Chang, J.-Y. (2014) The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL*, 26(2): 243–259. doi:10.1017/S0958344014000056
- 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. doi:10.1093/llc/fqq018
- Davies, M. (2020). The COCA corpus. The Corpus of Contemporary American English. https://www.english-corpora.org/coca/help/coca2020\_o verview.pdf.
- de Saint Léger, D. 2009. "Self-assessment of Speaking Skills and Participation in a Foreign Language Class." *Foreign Language Annals* 42 (1): 158–178. doi:10.1111/j.1944-9720.2009.01013.x.

Dolgova, & Mueller, C. (2019). How useful are corpus tools for error correction? Insights from learner data. *Journal of English for Academic Purposes*, 39, 97–108. https://doi.org/10.1016/j.jeap.2019.03.007

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, 7(3), 93-99.

https://doi.org/10.1016/j.afjem.2017.08.001

- Fang, L., Ma, Q. & Yan, J. (2021). The effectiveness of corpus-based training on collocation use in L2 writing for Chinese senior secondary school students. *Journal* of China Computer-Assisted Language Learning, 1(1), 80-109. https://doi.org/10.1515/jccall-2021-2004
- Hirata, Y., Hirata, Y., & Thompson, P. (2013). Two different types of corpora: Japanese students' perceptions. In J. Lam, K. C. Li, S. K. S. Cheung, & F. L. Wang (Eds.), *Knowledge Sharing through Technology* (Vol. 407, pp. 1–15). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-45272-7\_1
- Kim, H. (2019). The perception of teachers and learners towards an exploratory corpus-based grammar instruction in a Korean EFL primary school context. *Primary English Education*, 25(1), 123-152. <u>http://dx.doi.org/10.25231/pee.2019.25.1.123</u>

- Koo, K. (2006). Effects of using corpora and online reference tools on foreign language writing: A study of Korean learners of English as a second language (Order No. 3211793). Available from ProQuest Dissertations & Theses A&I. (305335210). https://www.proquest.com/dissertations-theses/effectsusing-corpora-online-reference-tools-on/docview/30533 5210/se-2?accountid=190474
- Leńko-Szymańska, A. (2015). A teacher-training course on the use of corpora in language education: Perspectives of the students. *Insights into Technology Enhanced Language Pedagogy*. Peter Lang.
- https://www.researchgate.net/publication/282355022
  Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning and Technology*, 21(3), 153–171.
- http://llt.msu.edu/issues/october2017/li.pdf Liu, D. (2013). Using Corpora to Help Teach
- Difficult-to-distinguish English Words. English Teaching (영어교육), 68(3), 27–50. https://doi.org/10.15858/engtea.68.3.201309.27
- Mican, A. D., & Medina, L. C. (2015). Boosting vocabulary learning through self-assessment in an English language teaching context. Assessment and Evaluation in Higher Education, 42(3), 398–414. https://doi.org/10.1080/02602938.2015.1118433
- Mueller, C., & Jacobsen, N. (2016). A comparison of the effectiveness of EFL students' use of dictionaries and an online corpus for the enhancement of revision skills. *ReCALL*, 28(1), 3-21. doi:10.1017/S0958344015000142
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. West Nyack: Cambridge University Press. https://doi.org/10.1017/CBO9781139524759
- Oktavianti, I. N., Triyoga, A., & Prayogi, I. (2022). Corpus for Language Teaching: Students' Perceptions and Difficulties. *Project (Professional Journal of English Education)*, 5(2), 441-455. https://pdfs.semanticscholar.org/6ff1/3595413ba39ff116 cf48fe3ccfe8f4ae0e34.pdf
- Pablo, J. C. I., & Lasaten, R. C. S. (2018). Writing difficulties and quality of academic essays of senior high school students. *Asia Pacific Journal of Multidisciplinary Research*, 6(4), 46-57.

http://www.apjmr.com/wp-content/uploads/2018/08/AP JMR-2018-6.4.06.pdf

Paker, T., & Ergül Özcan, Y. (2017). The effectiveness of using corpus-based materials in vocabulary teaching. International Journal of Language Academy. http://hdl.handle.net/11499/3743

- Quinn, C. (2015). Training L2 writers to reference corpora as a self-correction tool. *ELT journal*, 69(2), 165-177. https://doi.org/10.1080/09588221.2019.1667832
- Reppen, R. (2011). Using corpora in the language classroom. In *B. Tomlinson* (Ed.), Cambridge UP.
- Republic of the Philippines (2022). Department of Education Order No. 034 Series of 2022.
- Robinson, D. E., & Wizer, D. R. (2016). Universal Design for Learning and the Quality Matters guidelines for the design and implementation of online learning events. *International journal of Technology in Teaching and Learning*, 12(1), 17-32.

https://files.eric.ed.gov/fulltext/EJ1213328.pdf

Saeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of hands-on and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. *System*, 91, 102268.

https://doi.org/10.1016/j.system.2020.102268

- Sinha, T. S. (2021). EFL Learners' Perception of and Attitude to Corpus as a Vocabulary Learning Tool. *The Reading Matrix: An International Online Journal*, 21(2). http://mail.readingmatrix.com/files/25-3ds10cs0.pdf
- Urbano, C. M., Gumangan, M. A., Gustilo, L., & Capacete, M.
  P. A. (2021). Reading and Writing Needs of Senior High School Students: The Case of Filipino Students in the Philippines. *Modern Journal of Studies in English Language Teaching and Literature*, 3(1), 140-166. <u>https://connect.academics.education/index.php/mjselt/</u> <u>article/view/54</u>
- Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology*, 20(3), 159–179. http://hdl.handle.net/1808/26310

Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology*, 21(3), 1–8.

https://hal.archives-ouvertes.fr/hal-01237582

Xiaoli, Y. U., & Altunel, V. (2018). Second language vocabulary learning from context clues: A review of research in the past decade and implementation in digital environment. *Journal of Educational Technology and Online Learning*, 1(1), 1-12. https://doi.org/10.31681/jetol.375803