

## THE APPLICATION OF CORPUS TECHNOLOGY TO IMPROVE STUDENTS' READING SKILL

Karnaeni\*<sup>1</sup>, Muhammad Zuhri Dj<sup>2</sup>, Musfirah<sup>3</sup>  
<sup>1,2,3</sup>IAIN Bone

Author corresponding email: [Karnaeni790@gmail.com](mailto:Karnaeni790@gmail.com)

### Abstract

This research aims to determine the application of corpus technology to improve students' reading skills at SMAN 26 Bone. The participants in this research were students of class XI IPS 2 SMAN 26 Bone and the sample selected was cluster random sampling. In this research, the researchers used quantitative method with pre-experimental design. In addition, the instrument used was a test where the test consisted of pre-test and post-test. In the data analysis technique, the researchers used the SPSS version 23 program. The results of this research showed that the average value of class students in the pre-test (34.75) and post-test (65.00) after being given Corpus Technology treatment. The t-test value showed that the sig. (2-tailed) of 0.00 which means  $0.00 < 0.05$ . Based on the description above, the Alternative Hypothesis ( $H_a$ ) of this research is accepted and the Null Hypothesis ( $H_0$ ) of this research is rejected. Based on this, it can be concluded that Corpus Technology is effectively used to improving students' reading skills at SMAN 26 Bone.

**Keywords:** Corpus, Reading Skill, Technology

### INTRODUCTION

Reading is one of the four main skills that students of foreign languages need to master in order to successfully learn languages need to master in order to successfully learn the language (Rraku, 2013). Reading is one of the main activities of our everyday life as every one of us is engaged in reading in a way or another, be it just reading a newspaper article or taking a glimpse at the main headlines, reading a book or the label of a bottle of medicine. We consider developmental aspects of reading skills and strategies and suggest how these two related but different aspects of reading can be reconciled in a productive manner. We then describe implications of the distinction between skills and strategies for reading instruction and reading assessment (Afflerbach et al., 2008).

Reading is strategy exercise calls on the reader to effectively locate information in the reading text (Ni Kadek Yuli Cahyani et al., 2022). The reading activities conducted by the English teacher is already fulfill the criteria in teaching reading, where all the students instructed to participate during learning process. Today, like many teachers and researchers, we use the terms skills and strategies, both formally and informally, to describe features of children's reading development as well as features of teachers' reading instruction. One of them is media technology. Technology has affected the way we read and write. Along with the development of technology, reading sources have shifted from print media to electronic media such as e-books, online articles and news that can be accessed through mobile devices. In addition, information technology has also affected people's interest in reading, especially among students. Technology allows access to information anywhere, and anytime, without space and time constraints.

Technology media is very interesting to implement because it can integrate the use of information technology in education. The integration of technology in education has benefits in the development of education for the better. The use of learning media is also in accordance

with the times that continue to develop. One of the technology-based learning media that is often used in the learning process is website-based learning. Web-based learning allows students to connect with each other at the same time through the internet during the learning process.

With the development of technology, more technologies were introduced to the classroom to help facilitate the learning process and the integration of corpus technology in language classroom also starts to get attention from EFL (English as a Foreign Language) teachers. Corpus means electronic language database. Corpus analysis adopts computer assisted method to process electronic language database to reveal language operational models with statistic methods which usually centered on phenomenon of probability (Xia, 2022). Corpus analysis is often used in language research filed to determine the significant differences between two corpora or language features relation within one corpora with statistical methods like chi-square test, log-likelihood rate and correlation analysis (Oktavianti et al., 2022).

In English language learning, one of the technology-based media that is often used both inside and outside the classroom is the Corpus of Contemporary American English (COCA Corpus). This website is very familiar among learners and even experts of various levels in the field of English throughout the world. It is an official website that provides more than one million words of Contemporary American English, which includes various features to learn various aspects of learning in English. COCA Corpus is used to help find out the use of an English sentence, starting from meaning, pronunciation, grammar, word in context from various types of sources such as news, academic, spoken, blogs, websites, fiction, TV, and so on. COCA Corpus has very complete features and authentic sources so it is very related to its use in everyday life (Istifadah, 2023).

Based on the pre-observation conducted by the researchers with the English teacher at SMAN 26 Bone, the problem found was the low reading skill of students where the English teacher said that students had difficulty in conveying ideas in English which resulted in low reading comprehension. Students have difficulty understanding what they read so that the learning process is not optimal, students are afraid of being wrong in reading English texts For example, when given a reading text, students only read without knowing the meaning of the reading and the pronunciation is still wrong so that students tend to have difficulty in pronouncing words, understanding written words and naming things quickly.

To improve students' reading skills, researchers used the corpus technology method. Because as explained in the previous paragraph that corpus technology is a collection of natural texts, both spoken and written language, which are arranged systematically, and are easy for students to understand. So that students can learn to read well because the material is easy to understand and fun.

The British National Corpus has been the source for many corpus-related studies since its release in the early 1990s perhaps more than any other corpus of English during this same period (Davies, 2009). A corpus is intended to be "a collection of naturally occurring language text, chosen to characterize a state or variety of a language". As such, very few of the so-called corpora used in current natural language processing and speech recognition work deserve the name. For English, the only true corpora that are widely available are the Brown Corpus and the British National Corpus (Hawtin, 2018). Although it has been extensively used for natural language processing work, the million words of the Brown Corpus are not sufficient for today's largescale applications. For example, for tasks such as word sense disambiguation, many word senses are not represented, or they are represented so sparsely that meaningful statistics cannot be compiled. Similarly, many syntactic structures occur too infrequently to be significant.

Corpus means electronic language database. Corpus analysis adopts computer assisted method to process electronic language database to reveal language operational models with

statistic methods which usually centered on phenomenon of probability (Xia, 2022). In addition Islamiyah & Fajri, (2019) A corpus generally refers to a collection of naturally-occurring text from diverse context either spoken or written. corpus is intended to represent a whole language or part of a language. Corpus consultation can decide the materials included in ESL/EFL classes because of the limited teaching and learning activity time. Bringing this issue to Indonesia context in which English is a foreign language, it is of prominence to notice the role of a corpus in the ELT practice. However, it is not surprising to find relatively poor recognition of corpus use among English teachers or ELT practitioners in Indonesia. Apart from that, some studies on corpus in Indonesia context have been carried out in terms of the direct use of corpus. It informed materials have several strengths, such as describing the register-specific and frequency information, improving language learning awareness, learning motivation, critical thinking, and many more. However, corpus-informed materials should not be the only thing to rely on, so lecturers or instructors should provide adequate explanations for the materials and corpus research. These results show that it is necessary to promote the use of corpus-informed materials in EFL classrooms, including grammar classes. (Oktavianti et al., 2022)

Since the corpus was introduced throughout these decades, scholars have begun analyzing, evaluating, and annotating students' articulations. (Adolph, 2016) presented a corpus-based pronunciation training method based on learner corpora, as opposed to most corpora used in previous corpus-based pronunciation studies, and proved its effectiveness. The development of corpus linguistics has brought breakthrough to such deadlock in vocabulary instruction. It thrives on data to analyze and discover what language speakers do. Large bodies of text reveal these patterns in words, grammar and discourse. When computer aid this process, those texts can be handled in seconds, especially if they are tagged for parts of speech or specific information. (Yusu, 2014) claims that "corpus linguistics can make a difference for language learning and teaching and that it has an immense potential to improve pedagogical practice".

The realization of interesting media can be done by integrating the use of information technology in education. The integration of technology in education has benefits in the development of education for the better. The use of learning media is also in accordance with the times that continue to develop. One of the technology-based learning media that is often used in the learning process is website-based learning. Learning using the web allows students to connect with each other at the same time via the internet during the learning process. Website learning media can be accessed at any time by students without any time restrictions both in the classroom and outside the classroom.

In English language learning, one of the technology-based media that is often used both inside and outside the classroom is the Corpus of Contemporary American English (COCA Corpus). This website is very familiar among learners and even experts from various levels in the field of English throughout the world. This website is an official website that provides more than one million words of Contemporary American English, in which there are various features for learning various aspects of learning in English. Leech describes the term "corpus" as the plural of "corpora" which refers to a context-based collection of materials that has a specific purpose in the process. COCA Corpus is used to help find out the use of an English sentence, starting from meaning, pronunciation, grammar, word in context from various types of sources such as news, academic, spoken, blogs, websites, fiction, TV, and so on. COCA Corpus has very complete features and authentic sources so it is very related to its use in everyday life (Istifadah, 2023).

The basic skill of reading supports all aspects of education and intellectual development. It enhances vocabulary, comprehension, and critical thinking—all of which are essential for effective communication and lifetime learning. In order to provide a useful tool

For the use of corpus in vocabulary teaching, an appropriate corpus must be selected, particularly one that enables study of the metalinguistic awareness and is user-friendly and accessible (Yusu, 2014). According to Davies, (2010) an introduction to COCA, in terms of its balance from year to year. Let us now provide a number of more in-depth examples of how the COCA data can be used to answer interesting questions about changes in English morphology, syntax, semantics, and discourse in ways that are not possible with any other existing corpus, text archive, or linguistic resource.

The Corpus of Contemporary American English (hereafter COCA) a new corpus that has recently been placed online, and which is intended to compensate for the limitations of the two corpora just mentioned (Mark, 2009). COCA uses the same architecture and web interface that we have created for other large corpora that we have placed online. Therefore, although researchers referred to the “COCA architecture and interface”, this is simply a shorthand abbreviation for the architecture and interface used by COCA, yet which are also used for a wide range of corpora that we have created and placed online. Accordingly, this study seeks to answer the following research question: Does the application of corpus technology improve student reading skill?

## **METHOD**

### **Design and Sample**

The research adopts a quantitative approach using an experimental design to examine the effects of corpus technology to improve students' reading skill. The sample in this research used cluster random sampling take students of class XI IPS 2 SMAN 26 Bone, consist of 20 students who were not divided and remained as one intact group receiving treatment. This class was selected because it was accessible throughout the research process time, it was representative of the target group, and it was possible to coordinate with the school. Additionally, the class had a significant number of students and a somewhat steady academic performance, which made it appropriate for successfully applying the experimental design. In this strategy, variables are measured both prior to and following a particular intervention or treatment. The researchers assessed the reading proficiency of the pupils both before and after corpus-based technology was implemented.

### **Data Collection**

The Researchers used many strategies to collect data, which were referred to as data collection methods (Sugiyono, 2003). To collect data, the Researchers used reading tests in the form of pre-test and post-test. before treatment started, Researchers conducted a pre-test at the beginning of the research, and then after the pre-test the researchers conducted treatment, in this treatment the researchers gave students several readings of different level instructions at each meeting. After completion the researchers conducted a post-test. The effectiveness of the corpus technology was then assessed by comparing the pre-test and post-test results.

### **Data Analysis**

In this research, the technique was used descriptive statistical analysis technique, namely a research method that describes facts according to the data obtained with the aim of knowing the learning outcomes achieved by students as well as obtaining student responses to learning activities and student activities during the learning process.

Scoring the student's reading skill test answer

$$N = \frac{\text{student's correct answer}}{\text{total number}} \times 100$$

N: Students Score

1. Classification the students score based on the following classification:
- 2.

**Tabel 1. Score Classification**

Score	Classification
91-100	Very Good
75-90	Good
61-74	Fair
51-60	Poor
Less than 50	Very Poor

(Ugwuanyi & Okeke, 2020)

For the technique of data analysis in this research uses SPSS 23. The researchers used the SPSS 23 program to assist researchers in analyzing data. Pair sample t-test on SPSS 23 is used to analyze the significance of the data. Pair sample t-test was one of the test methods used to assess the effectiveness of treatment which is characterized by a difference in mean before and after treatment was given.

The basis for making a decision to accept or reject Ho in this test is as follows.

1. If the significant value is > 0.05 then Ho is accepted and H1 is rejected (the performance difference is not significant).
2. If the significant value < 0.05 then Ho is rejected or H1 is accepted

**3.3.1. Student Motivation on English Reading by Using Readable Apps**

To know the average value of students, researchers apply the following:

$$\bar{x} = \frac{\sum x}{N}$$

Notes:  $\bar{x}$  = The means of the students

$\sum x$  = The total score

N = The number of sampel

(Sudijono, 2018)

**FINDINGS AND DISCUSSION**

**Findings**

The findings of this research are related to the classification obtained by students' pre-test and post-test. To answer the research questions in the previous chapter, the researchers gave a test which was given twice, namely the pre-test and post-test. In this research, the researchers used one group pre-test and post-test. At the initial stage, the researchers gave a pre-test, then treatment, where the treatment used corpus technology and then a post-test. The results of the post-test can answer the question of whether the application of corpus technology can improve students' reading skills at SMAN 26 Bone.

This research finding was related to student classifications for the pre- and post-tests, student frequency and percentage, average, and standard deviations for the pre- and post-tests. All of this is explained as follows:

1. The results of the pre-test and post-test scores obtained through the reading skills test
  - a. Pre-Test

**Tabel 2. Percentage of Students' Pre-test**

No	Score	Classification	Frequency	Percentage (%)
1	91-100	Very Good	0	0%
2	76-90	Good	0	0%
3	61-75	Fair	0	0%
4	51-60	Poor	2	10%
5	< 50	Very Poor	18	90%
Total			20	100%

Based on table 2, the table shows that during the pre-test there were no students who got the classification of very good, good, and fair scores. It can also be seen that there are 2 students by getting 10% who are classified as Poor. And there were 18 other students with 90% who were classified as Very Poor. It can be concluded that students' reading skill before being given the learning application of corpus technology was low. After the pre-test, students were given treatment. The researchers took six meetings for treatment. Reading materials created using information from the Corpus of Contemporary American English (COCA) were presented to students in the experimental group. Every session concentrated on particular reading passages that included corpus-extracted high-frequency and contextually relevant vocabulary. The students participated in a variety of exercises during this treatment, including recognizing collocations, examining word frequency, and comprehending how words are used in real-world situations. The instructor assisted the students in connecting vocabulary to meaning by guiding them as they investigated the roles that words play in various genres. To strengthen understanding, interactive exercises such as vocabulary matching, comprehension tests, and brief conversations were employed. By exposing students to real-life language use taken directly from the corpus, this corpus sought to improve both their vocabulary acquisition and general reading skill.

- b. Post-Test

**Tabel 3. Percentage of Students' Post-test**

No	Score	Classification	Frequency	Percentage (%)
1	91-100	Very Good	0	0%
2	76-90	Good	10	50%
3	61-75	Fair	4	20%
4	51-60	Poor	0	0%
5	< 50	Very Poor	6	30%
Total			20	100%

Based on table 3, the table shows that during the post-test there were no students who received Very Good and Poor score classifications. And it can also be seen that there are 10 students who get 50% who are classified as Good, there are four students who get 20% who are classified as Fair, and six students who get 30% who are classified as Very Poor. It can be

concluded that students' abilities after being given learning the application of corpus technology increased compared to before treatment.

### 2. The Students Mean Score And Standard Deviation

After calculating the students' pre-test and post-test results, the differences in the mean scores and standard deviations of students' English reading skills that apply the use of corpus technology can be seen in the following table :

**Tabel 4. The Students Mean Score And Standard Deviation**

		Mean	N	Std. Deviation
Pair 1	Pretest	34,75	20	11,973
	Posttest	65,00	20	20,326

Based on table 4 which this table shows the mean value and standard deviation of students in both the pre-test and post-test using the SPSS application. The mean value of the students' pre-test was 34.75 and the post-test was 65.00. The standard deviation of the pre-test was 11.973 and the post-test was 20.326. After obtaining the average value and standard deviation of students, it can be concluded that the average value of the post-test is higher than the pre-test (34.75 > 65.00). Based on this, it can be concluded that the application of corpus technology is effective to improve the language reading skill of students of SMAN 26 BONE.

### 3. Normality Test

In this research, the normality test used IBM SPSS version 23. The results obtained are as follows:

**Table 5. Test of Normality**

	Shapiro-Wilk		
	Statistic	Df	Sig.
Pre-test	.961	20	.559
Post-Test	.895	20	.033

Based on table 5, the results of the normality test using Shapiro-Wilk with a significance value of the pre-test of 0.559 > 0.05 and at the post-test of 0.33 > 0.05. Based on this, it can be said that the data obtained is normal data.

### 4. Test of Significance

To see whether the t-test is statistically greater than the t-table at the 5% significance level, it can be seen in the following table:

**Tabel 5. Test of Significance**

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	Df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Pre test- Post test	30,250	19,767	4,420	-39,501	-20,999	6,844	19	,000

Based on table 5 above, it can be seen that the sig value. (2-tailed) of  $0.000 < 0.05$ , then as the basis for decision making on the significant t test, it can be concluded that  $H_0$  is rejected and  $H_a$  is accepted. Furthermore, from the table above, it can be seen that the mean difference (-30.250) this value shows the difference between the average student in the pre-test and the average student in the post-test ( $65.00 - 34.75 = -30.25$ ) and the difference is -39.501 to -20.999 (95% Difference Confidence Interval).

The negative t value (-6.844) is not an error. However, the calculated t has a negative value because the pre-test mean is lower than the post-test mean.

## 5. Hypothesis Testing

The research formulation in this research is the Null Hypothesis ( $H_0$ ), The application of corpus technology cannot improve students' reading skills at SMAN 26 BONE.  $H_0$  criteria is not accepted if the sig value. (2-tailed)  $> 0.05$ . While the Alternative Hypothesis ( $H_a$ ), The application of corpus technology can improve students' reading skill at SMAN 26 BONE. The Alternative Hypothesis is accepted if sig. (2-tailed)  $< 0.05$ . And in this research it was found that sig. (2-tailed) is 0.00 which means  $0.00 < 0.05$ . After calculating the t-test, it is then compared with the t-table value with a significance level of  $p = 0.05$  with degrees of freedom ( $df = 20$ ) because the number of students is 20 people where  $N-1 = 19$  then  $df$  (19). So the t-test value is greater than the t-table value.

Then to test the hypothesis by comparing the t-test value with the t-table value can be seen in the following explanation: Based on the previous explanation, the Alternative Hypothesis ( $H_a$ ) of this study is accepted and the Null Hypothesis ( $H_0$ ) of this study is rejected. Which means that the application of corpus technology can improve students' reading skill at SMAN 26 BONE.

## Discussion

The application of the Corpus of Contemporary American English (COCA Corpus) website gave students access to authentic language use from a variety of genres and situations. Students developed their understanding of how language works in authentic communication by examining texts from a corpus, which introduces them to patterns of vocabulary, grammar, and collocations. Instead of depending only on textbook examples, corpus-based materials in reading instruction let students understand how words and phrases are used in real-world contexts. Reading fluency, comprehension, and vocabulary development are all enhanced by this exposure. Based on the results of quantitative data, it is proven that the application of corpus technology can improve students' reading skill at SMAN 26 BONE. At the time of the post-test there were no students who got the classification of Very Good and Poor scores. And it can also be seen that there are 10 students who get 50% which is classified as Good, there are four students who get 20% which is classified as Fair, and six students who get 30% which is classified as Very Poor. Based on this, when compared to before treatment, it means that students after being given learning the application of corpus technology have increased compared to before treatment.

At the time of treatment, treatment was carried out for six meetings, namely on the eleventh of October to the sixth of November. The first meeting of treatment was held on Friday the eleventh of October, where the activities of the researchers explained the material and gave examples of reading texts about Corpus Technology. The second meeting was on Wednesday the sixteenth of October, where the researchers explained the material and gave examples of descriptive text. The third meeting on Friday the eighteenth of October, the activities of the researchers explained the material and gave examples, about recount text. The fourth meeting

on Wednesday twenty-third October, the activities of the researchers explained the material and gave examples of Narrative Text. The fifth meeting on Friday 1 November, the activities of the researchers explained the material and gave examples of review text. And the sixth meeting on Wednesday the sixth of November, the activities of the researchers explained the material and gave examples of Expository text.

After calculating the results of students' pre-test and post-test, the difference in the mean value and standard deviation of students' English reading skill applying the use of corpus technology using SPSS application. The mean score of the students' pre-test was 34.75 and the post-test was 65.00. The standard deviation of the pre-test was 11.973 and the post-test was 20.326. After getting the students' mean score and standard deviation, it can be concluded that the post-test mean score is higher than the pre-test ( $34.75 < 65.00$ ). Based on this, it can be concluded that the application of corpus technology is effective to improve the language reading skill of students of SMAN 26 BONE.

Next, the researchers conducted a normality test, the purpose of this test is to determine whether the data obtained is normal or not. This normality test uses Shapiro-Wilk with a pre-test significance value of  $0.559 > 0.05$  and in the post-test of  $0.33 > 0.05$ . And it can be said that the data obtained is normal data.

The researchers then conducted a significance test, the researchers conducted a significance test to see whether the t-test was statistically greater than the t-table at the 5% significance level. it can be seen that the sig. (2-tailed) of  $0.000 < 0.05$ , then as a basis for decision making on the t significance test, it is concluded that  $H_0$  is rejected and  $H_a$  is accepted. Furthermore, from the table above, it can be seen that the mean difference (-30.250) this value shows the difference between the average student in the pre-test and the average student in the post-test ( $65.00 - 34.75 = -30.25$ ) and the difference is -39.501 to -20.999 (95% Difference Confidence Interval).

It was found that the sig value. (2-tailed) is 0.00 which means  $0.00 < 0.05$ . After calculating the t-test, it is then compared with the t-table value with a significance level of  $p = 0.05$  with degrees of freedom ( $df = 20$ ) because the number of students is 20 people where  $N - 1 = 19$  then  $df (19)$ . So the t-test value is greater than the t-table value. Based on the previous explanation, the Alternative Hypothesis ( $H_a$ ) of this study is accepted and the Null Hypothesis ( $H_0$ ) of this study is rejected. This means that the application of corpus technology can improve students' reading skill at SMAN 26 BONE.

The application of the Corpus of Contemporary American English (COCA Corpus) website improved students' reading skills at SMAN 26 Bone. In learning students have low reading comprehension, and students have difficulty understanding what they read. COCA CORPUS is one of the websites created to be a solution to the problem. The application of CORPUS technology can improve students' reading skills at SMAN 26 BONE.

This research is in line with previous research by Xia (2022) stated that corpus-based teaching activates students' autonomous learning skill and independent thinking with the guidance of teachers and provides students with native language materials with which they can conduct language exploration and discovery. Corpus analysis uses computer-assisted methods to process electronic language databases to reveal the operational model of language by statistical methods, which usually center on probability phenomena. The use of corpus technology in college English improved college students' vocabulary learning outcomes. The integration of corpus technology in language learning aligns well with the flexible and interactive nature of hybrid learning. Corpus—defined as a collection of naturally occurring texts both spoken and written—provides authentic, real-life language input which significantly enhances students' reading skills through exposure to contextual vocabulary, grammar patterns, and collocations. This is especially effective in hybrid learning environments, where students

can access digital corpus tools like the COCA Corpus both synchronously and asynchronously. In line with this, (Musfirah & Nur 2024) found that learners perceive hybrid learning as flexible, allowing access to digital resources that accommodate different learning styles and needs. By combining the benefits of corpus-informed materials with the digital adaptability of hybrid learning, educators can foster deeper personal, academic, intellectual, and social engagement among EFL learners.

Another researches in line with previous research, (Kuehnast et al., 2024) benefits from corpus-based teaching strategies and presents MachinaCallida, a psycholinguistic ally-informed e-tutor suitable for supporting Latin vocabulary acquisition and reading comprehension at beginner and intermediate levels. Using a digital corpus of original Latin texts, the application semi-automatically generates contextual vocabulary exercises tailored to the needs of different groups of learners.

## CONCLUSION

The conclusion is that the application of corpus technology to improves students' reading skill at SMAN 26 BONE. This is proven by looking at the results of data analysis based on the results of the one-sample t-test showing that sig. (2-tailed) is 0.00 which means  $0.00 < 0.05$ . Based on the description above, the Alternative Hypothesis ( $H_a$ ) of this study is accepted and the Null Hypothesis ( $H_0$ ) of this study is rejected. This means that the application of corpus technology can improve students' reading skill at SMAN 26 BONE. This application is a useful tool for analyzing vocabulary acceptance for comprehending how language is used in different contexts. Teachers can help students in improving their reading strategies and gaining a better understanding of language through integrating corpus data into their classroom instruction.

## REFERENCES

- Adolph, R. (2016). *A Corpus-based Pronunciation Teaching Model: A Conceptual Paper*. 14(1), 1–23.
- Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). *Clarifying Differences Between Reading Skills and Reading Strategies*. *The Reading Teacher*, 61(5), 364–373. <https://doi.org/10.1598/rt.61.5.1>
- Agustin, L., Wisudaningsih, E. T., & Fatmawati, R. (2023). *Exploring How Skimming and Scanning Fosters EFL Students' Reading Comprehension at an English Club Senior High School in Indonesia*. *Tamaddun*, 22(1), 20–27. <https://doi.org/10.33096/tamaddun.v22i1.309>
- Ahmad, A. (2017). *Penerapan Permainan Bahasa (Katarsis) Untuk Meningkatkan Keterampilan Membaca Siswa Kelas Iva Sd Negeri 01 Metro Pusat*. *EduHumaniora / Jurnal Pendidikan Dasar Kampus Cibiru*, 9(2), 75. <https://doi.org/10.17509/eh.v9i2.7024>
- Banditvilai, C. (2020). *The Effectiveness of Reading Strategies on Reading Comprehension*. *International Journal of Social Science and Humanity*, 10(2), 46–50. <https://doi.org/10.18178/ijssh.2020.v10.1012>
- Bojie Qian, F. diana deris. (2016). *A Corpus-based Pronunciation Teaching Model: A Conceptual Paper*. 4(1), 1–23.
- Breyer, Y. (2009). *Learning and teaching with corpora: reflections by student teachers*. *Computer Assisted Language*, 22(2), 153–172. <https://doi.org/https://doi.org/10.1080/09588220902778328>
- Cheng, W., & Lam, P. (2010). *What can a corpus tell us about language teaching? The Routledge Handbook of Corpus Linguistics, Second Edition, January 2010*, 299–312.

- <https://doi.org/10.4324/9780367076399-21>
- Daniel, B. K., & Harland, T. (2017). *Higher Education Research Methodology*. <https://doi.org/10.4324/9781315149783>
- Davies, M. (2009). *The 385+ million word Corpus of Contemporary American English (1990–2008+)*. *International Journal of Corpus Linguistics*, 14(2), 159–190. <https://doi.org/10.1075/ijcl.14.2.02dav>
- 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. <https://doi.org/10.1093/lc/fqq018>
- Fajriadi, D. (2022). *Peningkatan Kemampuan Siswa Pada Reading Comprehension Kelas IX Melalui Metode Group Discussion Mata Pelajaran Bahasa Inggris di SMP Negeri 3 Setia Bakti*. *Cendikia : Media Jurnal Ilmiah Pendidikan*, 13(1), 102–111.
- Flowerdew, L. (2009). *Applying corpus linguistics to pedagogy*. *International Journal of Corpus Linguistics*, 14(3), 393–417. <https://doi.org/10.1075/ijcl.14.3.05flo>
- Hawtin, A. (2018). *The Written British National Corpus 2014: Design, Compilation and Analysis*. September.
- Ide, N., & Macleod, C. (2014). *The American National Corpus : A standardized resource for American English \* Department of Computer Science Vassar College Computer Science Department New York University New York , New York 10003-6806 USA Why we need a corpus of American English*. 274–280.
- Islamiyah, M., & Fajri, M. S. Al. (2019). *Skinny, Slim, dan Thin: Analisis Berbasis Korpus Kata Sifat Identik dan Implikasinya pada Pengajaran Bahasa Inggris*. *Ranah: Jurnal Kajian Bahasa*, 8(1), 19. <https://doi.org/10.26499/rmh.v8i1.894>
- Istifadah, queen salsabila. (2023). *Corpus+Grammar+Media+Pembelajaran+Mandiri+Berbasis+Teknologi+Untuk+Meningkatkan+Kemampuan+Penguasaan+Tata+Bahasa+Inggris+Pembelajar+Bahasa+Asing+Di+Era+5.0+(1)*. 15(1), 8–19.
- Musfirah, M., & Nur, H. (2024). *Efl Learners' Perspectives on Their Engagement Through Hybrid Learning*. *Ekspose: Jurnal Penelitian Hukum Dan Pendidikan*, 23(1), 66–77. <https://doi.org/10.30863/ekspose.v23i1.6371>
- Ni Kadek Yuli Cahyani, Artini, L. P., & Wahyuni, L. G. E. (2022). *Reading-Related Activities in English Textbook and How the Texts are Exploited in Junior High Schools*. *Journal of Education Research and Evaluation*, 7(1), 115–120. <https://doi.org/10.23887/jere.v7i1.48784>
- Rraku, V. (2013). *Vasilika Rraku*. *Social and Natural Sciences Journal*, 7(2), 2013.
- Xia, S. (2022). *The Effectiveness of Using Corpus Technology in College English Teaching*. Atlantis Press International BV. [https://doi.org/10.2991/978-94-6463-044-2\\_68](https://doi.org/10.2991/978-94-6463-044-2_68)
- Yusu, X. (2014). *On the application of Corpus of Contemporary American English in vocabulary instruction*. *International Education Studies*, 7(8), 68–73. <https://doi.org/10.5539/ies.v7n8p68>