Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

Effectiveness of Sketch Engine for Language Learning as an Educational Tool for Improvement of Paraphrasing Skills of ESL/EFL Undergraduate Students

Summaira Sarfraz
Director & Professor, Humanities & Sciences Department, National University of Computer and Emerging Sciences, Lahore Campus

Nokhaiz Zahra (Corresponding Author)
Lecturer, Humanities & Sciences Department, National University of Computer and Emerging Sciences, Lahore Campus

Abstract
This study aims to investigate the effectiveness of SkELL - Sketch Engine for Language Learning in terms of improving students' paraphrasing skill using the three features of SkELL web interface, i.e., examples, word sketch and similar words. Paraphrasing skill is an under explored area in language learning which highly affects ESL/EFL students' academic performance however, Pakistani students lack this skill as they encounter deficiency in vocabulary, comprehension, semantic and syntactical expertise. SkELL based activities utilizing word collocations, thesaurus and concordance, helped students overcome these problems. For the evaluation of the intervention of SkELL in teaching paraphrasing skill, exploratory sequential method was followed; a mixed method approach using a quasi-experimental approach for quantitative data collection and focus group interviews were conducted for qualitative data collection. Keck's taxonomy was adopted as a conceptual framework moreover, it was utilized as the rubric for evaluating paraphrasing skill, encompassing four types of paraphrasing: near copy, minimal revision, moderate revision, and substantial revision. For the quantitative data analysis IBM SPSS- 25 and for the qualitative data analysis MAXQDA-2022 were used. The Wilcoxon test result indicated a significant difference in the paraphrasing skill of students between the pre-test and post-test.

Keywords: Collocation, concordance, Keck's taxonomy, SkELL, types of paraphrases

1. Introduction

For the past two decades, linguists have increasingly utilized the Web as a major source of language data due to its free, instantly available, versatile, and extensive nature (Fukushima et al., 2012). Sketch Engine, selected for this study on language acquisition, stands out among numerous automated corpora tools, each offering distinct features for acquiring various language skills (Boulton & Cobb, 2017). Sketch Engine, a web-based program developed by British lexicographer and research scientist Adam Kilgarriff, offers basic annotation,

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

exemplary sentences, and thesaurus alongside advanced functions like the Corpus Query System (CQS) (Biasi, 2018). While the web interface includes built-in corpora in six languages, registered users can create and analyze corpora in any language, accessing features such as examples, word sketches, and thesaurus options.

1.1. Examples

'*Examples'* presents 40 most frequently used valid exemplary sentences of the given input. Concordancer is a computer program that analyzes texts and generates a list of all the words in a corpus along with their frequency and context, making it a useful tool for linguistic and textual analysis.

1.2. Word Sketches

Word Sketch, per Kilgarriff et al. (2014), is a succinct summary revealing grammatical and collocational behaviors, leveraging statistical data for larger corpora effectiveness. Sketch Engine's lempos ensures rapid generation of collocate lists, aiding users in selecting functions and enabling language experts to compare word patterns cross-linguistically and cross-culturally.

1.3. Similar Words (Thesaurus)

Similar words present the cloud of the similar words hence stating the near and far synonyms (Jakubíˇcek & Rychlý, 2019). The backbone of the thesaurus is the similarity score of the shared collocates of the words within certain corpus. The job of finding similarities between the collocations of the words is done instantly in billion words corpus with much efficiency and accuracy though Sketch Engine.

1.4. Problem Statement

Silviana (2022), Tran and Nguyen (2022) emphasize the crucial role of paraphrasing in academic writing, a significant challenge for non-native speakers. Pakistani students, particularly in undergraduate programs, face hurdles due to limited vocabulary and syntactical expertise in their second language. The absence of ICT tools exacerbates the difficulty in acquiring effective paraphrasing skills.

1.4 Purpose of the Research

The aim of the research is to determine the effectiveness of SkELL in order to improve the paraphrasing skill of the English language learners of undergraduate students in a private university in Pakistan.

1.5 Research Objectives

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

- To determine the implementation of SkELL for the improvement of paraphrasing skills.
- To investigate the effectiveness of SkELL for the improvement of paraphrasing skills of undergrad students in Pakistani university.

1.6 Research Questions

i. How SkELL can be implemented for the improvement of paraphrasing skills.
ii. How effective is SkELL for the improvement of paraphrasing skill of undergrad students in Pakistani university?

1.7 Significance of the Study

In the context of Pakistani universities, undergraduate students encounter significant challenges in producing academic assignments, such as lengthy essays, research papers, dissertations, and thesis. The academic writing demands systematic presentation of facts, requiring a profound understanding and extensive vocabulary. Paraphrasing skills become crucial when students need to express ideas in their own words after conducting in-depth analyses of research and academic articles. The research aims to assess SkELL's effectiveness in enhancing the paraphrasing skills of undergraduate students in this academic context.

1.8 Research Scope and Constraints

SkELL's efficacy may vary due to unexplored factors like user preferences and learning styles. The study solely focuses on paraphrasing, limiting generalization, and overlooks potential technical issues. The study confines itself to a 3-week period, assessing SkELL's impact on paraphrasing skills with 50 participants. Findings, while crucial, are limited to the standardized test, excluding other interventions or factors affecting language proficiency.

2. Literature Review

Kilgarriff et al. (2010) evaluate SkELL's word sketch summary, analyzing collocates across languages. Hirata and Hirata (2019) support Sketch Engine for collocation learning. Kilgarriff et al. (2014) contrast 'clever' and 'intelligent' in word sketch difference. Stewart et al.'s (2018) study challenges absolute synonyms. Koeva (2015) discusses paraphrasing, while Yuliawati and Indira (2019) analyze 'courageous' and 'brave.' Barrs (2016) underscores Sketch Engine's importance. Aldhubayi and Alyahya (2014) affirm its utility in exploring semantic similarities. Taran et al.'s (2021) stylometric study reveals low-frequency words. Kovar et al. (2016) commend its significance for scholars. Hu and Yang (2015) compare 'raise' and 'increase' usage patterns. Fantini (2018) guides on Sketch Engine's capabilities. Choy and Lee (2012) reveal students' paraphrasing perception using SkELL. Irmadamayanti (2018) assesses self-perception using SkELL. Dung (2010) advocates an empowering learning environment. Tran and Nguyen (2022) find paraphrasing productive. Chen's (2019) study evaluates PREFER's effectiveness, which is a tool to measure student’s paraphrasing skill. Tsou (2015) notes a surge in online corpora demand. Kocincova et al. (2015) use large text corpora for evidence. Herman

ISSN Online: 3007-3154
ISSN Print: 3007-3146

et al. (2019) highlight word sketches in the 'semiautomatic dictionary.' James (2016) describes SkELL's value as a language analysis tool.

In summary, these studies collectively underscore Sketch Engine's versatility, ranging from collocation learning to stylometric analysis, and its pivotal role in enhancing language exploration and comprehension across diverse linguistic contexts.

Multiple studies suggest that students find the paraphrasing skill really challenging in their academic tenure advocating for improved vocabulary and practice. Research done on eight graders by Ruslan et al (2020) and Hidayat (2022) highlight the significance of student-teacher relation on improving paraphrasing skill of students and their level of understanding. Another research conducted by Pratama at al. (2022) examined the strategies and challenges of doing paraphrasing using the mixed method approach on EFL students. Providing the rubric for evaluating the paraphrasing performance of EFL students, Keck (2006, 2014) emphasized the strategies of avoiding plagiarism. Also, Sanjay (2021) addresses the importance of paraphrasing for EFL students in their academic career and teaching the strategies of paraphrasing skill to help them improve. Linguistically, Ji's (2018) study correlates paraphrasing wit lexical competence, and Sarair et al. (20190 underscore the need for teaching paraphrasing strategies to Indonesian university students. In the realm of technology, Choi (2012) Kim (2020), and Syahnaz and Fithriani (2023) explore paraphrasing tools, with Choi highlighting the positive effects of teacher-led and web-led approached, Kim introducing a paraphrasing test, and Syahnaz and Fithriani suggesting AI-based tools for improved writing skills.

In conclusion, these studies collectively emphasize the multifaceted importance of paraphrasing in education, ranging from academic integrity to linguistics proficiency and technology in education, ranging from academic integrity to linguistic proficiency and technological interventions. As educators navigate the challenges presented by these studies, incorporating effective teaching strategies, technological tools, and linguistics insights will contribute frosting students' paraphrasing skills for comprehensive academic success.

3. Methodology

The efficacy of Sketch Engine for Language Learning in terms of improving paraphrasing skill has been examined using the mixed method approach. The participants selected through proposed sampling were EFL students of Computer Science department. The study employs both quantitative and qualitative approach. For quantitative data, the results of the quasi experiment with pre and post test results were analyzed using the SPSS 25, and for qualitative data, the focus group interviews of 8 students were analyzed using MAXQDA 2022, which helped generate thematic clouds, interactive word tree and semantic analysis.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

3.1 Population and Sample Size

The study enlisted 50 purposively chosen Computer Science students to ensure a diverse and representative sample in the second semester of the bachelor's program.

3.2 Quantitative Data

The primary source of the data collection was the students' write ups from the pre-test and post-tests. Two separate tests based on a paragraph, with academic content were composed. The test contained quite simple instructions indicating that the given paragraph is to be paraphrased. The pre-test was conducted before any kind of intervention. 50 students of the Computer Science department from two sections were pre-tested on the identical tests based on the concept of paraphrasing. The participants of the research were given the pre-test in their English language lab sessions, the test contained a paragraph to be paraphrase. All the students had the same paragraph to be paraphrased. The excerpt was taken from an academic article and contained words of low to medium difficulty words. The time of 10 minutes was given the students, in order to perform this activity. No instructions regarding the solution of the task were given to the students. The pre-tests were then collected by the researcher for the purpose of analysis.

3.3 Theoretical Framework

Keck's (2006) "Taxonomy of Paraphrasing" played a role of theoretical framework for this research and was used as the rubric to evaluate students' paraphrasing write ups. According to Keck's (2006) taxonomy of paraphrasing, the categorical division of paraphrasing contains four types: 1) near copy, 2) minimal revision, 3) moderate revision and 4) substantial revision. Students' pre-test-post-test paraphrase types will firstly be identified. Pre-test and post-test will be evaluated based on the frequency of paraphrased words and percentage bracket relative to the preprophase type. This will help identity the paraphrase type at sentence level.

Table 1: ***Taxonomy of Paraphrase***

Student	Frequency	Percentage	Paraphrase Identifications			
			Near Copy	Minimal Revision	Moderate Revision	Substantial Revision
Student 1						
Student						

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

2
Student
3
Student
4
Student
5

3.4 SkELL Based Activities

Once the students had attempted and done the paraphrasing's of the paragraphs given to each one of them, pre-test based on their prior knowledge, the students were given a brief introduction of this web-based corpus tool, Sketch Engine for Language Learning. In total 6 activities were designed in order to make the students able to use SkELL and all its features conveniently. These activities were focused on making students familiar with SkELL and help them develop paraphrasing skill step by step. Students will practice using SkELL for the purpose of paraphrasing. The students experienced using this e-tool in the language labs on their systems individually. The basic three features of the tool, 1) examples, 2) word sketch 3) similar words along with their use was elaborated in a precise manner to the students.

3.4.1. Activity 1 – Use of "Examples" The students used the feature of *examples.* They were given 10 words and were asked to use this feature of examples to find the meanings of the given words. Students, entered the words one by one in the search box of SkELL. For each word, 40 authentic sentences were presented to them. After reading multiple sentences, they implied the meaning of that words in multiple contexts. This activity of reaching out to the meaning of the word through sentences was conducted within 10 minutes.

3.4.2 Activity 2 Use of 'Similar words" In this activity, the students made use of another feature of SkELL, which is 'similar words.' The participants were given sentences with few words written in bold and underlined. The words underlined and written in bold were supposed to be replaced with their synonyms using this particular feature of similar words. Once the words were entered in the search bar, this feature presented all the near and far synonyms in form of a cloud. Based on the context of the sentence, the students were supposed to replace the words with their synonyms. They were also allowed to make structural changes while replacing the key word.

3.4.3. Activity 5 Using "Word Sketch" Students were instructed to use this feature in order to understand the multiple functions of the same words. After writing the word in the search box, the tool represents the behaviour of that word X as a noun/verb/adjective etc. Also, it showed the behaviour of searched word at the place of subject and object. Moreover, the collocations of the searched word, helped students get ideas about what word can be appearing as the neighbor of searched word. Students then made sentences of their own using the words as instructed.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

3.4.4. Activity 6 "Unique Links and General Links" In paraphrasing practice, students located replaceable and irreplaceable words, using SkELL's similar word feature to find alternatives within the sentence context, enhancing their paraphrasing skills.

3.5 Qualitative Data

The semi-structured interview consisted of a total of 12 questions, carefully designed to gather comprehensive insights from the participants. The interview aimed to delve into various aspects of the subject matter and encourage open-ended responses. The questions were strategically crafted to elicit rich and nuanced information, enabling a deeper understanding of the topic under investigation.In the focus group, a total of 10 students actively participated in the interview process. To ensure a diverse representation, five students were selected from each section.

4. Findings & Analysis

The combination of descriptive analysis with graphs and tables, and inferential analysis with statistical tests, enables a comprehensive exploration and interpretation of the data. This two-part approach enhances the rigor and depth of the analysis, providing a robust foundation for the subsequent discussion and implications of the research findings.

4.1. Descriptive Analysis

The table below carries the sample IDs of the 50 participants in sequence in the first Column from left. Next to which, the frequency of the copied string of words have been calculated using the rubric of taxonomy of paraphrase by Casey Keck and is mentioned in the second column. In the third column, the percentages of the copied words from the original text have been mentioned, proceeding to which there are 4 categories of the paraphrasing types has been mentioned such as near copy, minimal revision, moderate revision and substantial revision. Based on the percentages predetermined earlier that is near copy: 50% to 100% minimal revision: 20% to >50%, moderate revision 1% to >20% and substantial revision 0%, the category of each paraphrasing done by each participant has been marked.

Table 2: Individual Student wise Pre-Test Results

			Paraphrase Identifications of Pre-Test			
Sample ID	**Frequency**	**Percentage**	**Near Copy**	**Minimal Revision**	**Moderate Revision**	**Substantial Revision**
1	**13/57**	**24.07 %**		□		
2	**14/60**	**23.33%**		□		
3	**23/56**	**50%**	□			
4	**20/45**	**44.44 %**		□		
5	**10/36**	**27.77%**		□		

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

6	13/42	30.95 %		□	
7	23/47	48.93 %		□	
8	11/48	22.91 %		□	
9	11/69	15.94 %			□
10	17/63	26.98%		□	
11	16/61	22.22%		□	
12	14/77	18.18%			□
13	9/53	16.98%			□
14	4/51	7.84%			□
15	22/49	44.89%		□	
16	13/61	21.31%		□	
17	16/31	51.61%	□		
18	14/61	22.95%		□	
19	7/59	11.86%			□
20	36/50	72 %	□		
21	2/51	3.92%			□
22	33/56	58.92%	□		
23	5/58	8.7%			□
24	14/54	25.92%		□	
25	16/56	28.57		□	
26	24/49	50%	□		
27	8/45	17.77%			□
28	20/59	33.89%		□	
29	11/55	20%		□	
30	8/85	4.4%			□
31	8/60	13.33%			□
32	16/56	28.57%		□	
33	19/54	35.18%		□	
34	17/65	26.15%		□	
35	26.44	59.09%	□		
36	14/53	26.41%		□	
37	22/58	37.93%		□	
38	13/66	19.69%			□
39	14/72	22.58%		□	
40	20/60	33.33%		□	
41	38/56	67.85%	□		
42	10/63	15.87%			□
43	11/63	17.46%			□
44	11/52	21.15%		□	
45	8/60	13.33%			□
46	16/59	27.11%		□	
47	12/57	21.05%		□	
48	14/49	28.57%		□	
49	4/56	7.14%			□

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

50	**4/56**	**7.14%**	□

The pre-test and post-test results were analyzed using the rubric adopted by Kech (2006), called as taxonomy of paraphrasing. The results of the pre-test showed that the out of total students 50, 7 were classified in near copy, indicating a high degree of similarity to the original test. Whereas majority of the students, that is 27 students' work were classified as the minimal copies, indicating the minimum revisions in the original text. In contrast 16 students' work was classified as the moderate revision, indicating the moderate level of similarity with the source text through incorporation restructuring and rewording of the original text. On the other hand, no student's work was categorized as the substantial revision, which is the highest and the most appropriate level of paraphrased version. These findings highlight the need for targeted instruction on paraphrasing strategies to enhance students' ability to restate information effectively in their own words.

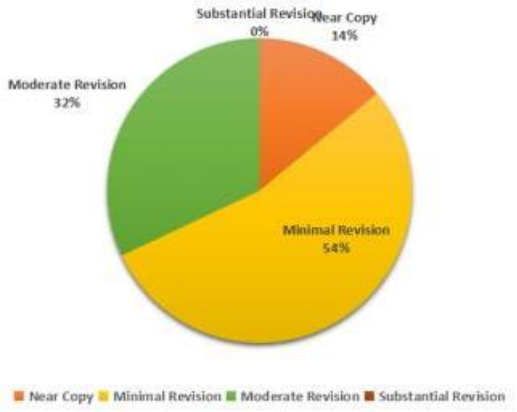


Figure 1
Pie Chart depiction of Pre-Test Results

Table 3: Individual Student wise Post Test Result

		Paraphrase Identifications of Post Test				
Sample ID	**Frequency**	**Percentage**	Near Copy	Minimal Revision	Moderate Revision	Substantial Revision
1	9/96	9.37 %			□	
2	16/117	13.67 %			□	
3	33//85	38.82 %		□		
4	0/80	0 %				□
5	6/113	5.30 %			□	
6	0/94	0 %				□
7	6/106	5.66 %			□	

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

8	2/78	2.56 %		□	
9	8/103	7.76 %		□	
10	0/135	0 %			□
11	0/74	0 %			□
12	0/119	0 %			□
13	12/123	9.75 %		□	
14	0/124	0 %			□
15	19/51	37.25 %	□		
16	4/104	3.84 %		□	
17	12/71	16.9 %		□	
18	0/123	0 %			□
19	0/109	0 %			□
20	19/88	21.59 %	□		
21	0/74	0 %			□
22	17/86	19.76 %		□	
23	0/134	0 %			□
24	4/78	5.12 %		□	
25	9/112	8.03 %		□	
26	17/86	19.76 %		□	
27	2/73	2.73 %		□	
28	15/121	12.39 %		□	
29	0/100	0 %			□

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

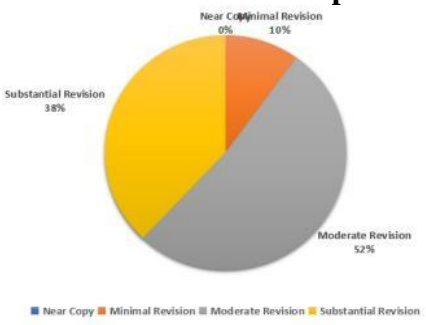
Vol. 3 No. 6 (June) (2025)

30	4/111	3.60 %			□	
31	0/92	0 %				□
32	8/94	8.51 %			□	
33	3/98	3.06 %			□	
34	0/31	0 %				□
35	38/91	41.75 %	□			
36	0/104	0 %				□
37	30/118	25.42 %	□			
38	1/85	1.17 %			□	
39	7/95	7.36 %			□	
40	5/95	5.26 %			□	
41	15/86	17.44 %			□	
42	0/85	0 %				□
43	0/118	0 %				□
44	0/92	0 %				□
45	0/93	0 %				□
46	5/148	3.37 %			□	
47	4/111	3.60 %			□	
48	13/111	11.71 %			□	
49	0/109	0 %				□
50	5/82	6.09 %			□	

The results of the post-test, which were evaluated based on Keck's (2006) paraphrasing taxonomy, revealed that none of the paraphrases produced by the participants were categorized as near copy, indicating significant departure from the original text. On the other hand, five paraphrases were classified as minimal copies, indicating a moderate level of similarity to the source text. A large proportion of paraphrases 26, were categorized as moderate revision, suggesting that the participants made considerable efforts to rephrase and restructure the original text. Interestingly, 19 paraphrased were classified as substantial revision, indicating that the participants were able to make substantial changes to the original text while maintaining its original meaning. These findings highlight the importance of teaching paraphrasing strategies to students to enable them to convey information effectively in their own words.

able to make substantial changes to the original text while maintaining its original meaning. These findings highlight the importance of teaching paraphrasing strategies to students to enable them to convey information effectively in their own words.

Figure 2
Pie Chart depiction of Post Test Results



Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

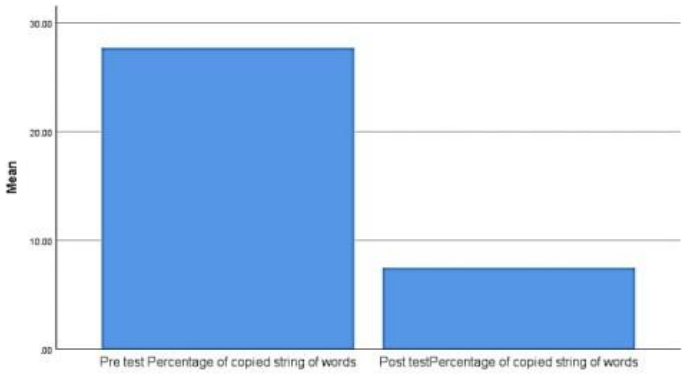
Vol. 3 No. 6 (June) (2025)

4.2. Data Analysis Procedure

In order to analyze the data collected from the quasi-experimental study on the effectiveness of teaching paraphrasing strategies to students, the researchers utilized the SPSS software, version 25, for sequential data analysis To ensure the validity of the findings, multiple tests were applied, including tests of normality and Wilcoxon signed-rank test. These tests were applied to the data that was analyzed using Casey Keck's taxonomy of paraphrasing. By using a quasi-experimental design and a systematic approach to evaluate the quality of paraphrasing, the researchers were able to draw reliable conclusions about the effectiveness of teaching paraphrasing strategies.

Figure 7 shows the differences in the mean of the total copied words used n the pre-test and the posttest, regardless of the different categories. The mean of the pretest percentage of the copied words in approximately 28 per cent where that of posttest is 8 percent. This suggests a significant decrease in the use of copied words following targeted instruction of paraphrasing strategies.

Figure 3
Average Difference of Pre test and Post Test results



The chart provided displays the performance of each student before and after a targeted intervention aimed at improving their paraphrasing skills. The blue bars represent the percentage of copied words used by each student in the pre-test, while the red bards represent the percentage of copied words used in the post test. The results reveal a noticeable decrease n the use of copied words among the students after the treatment. This suggests that the intervention was effective in improving their ability to paraphrase effectively.

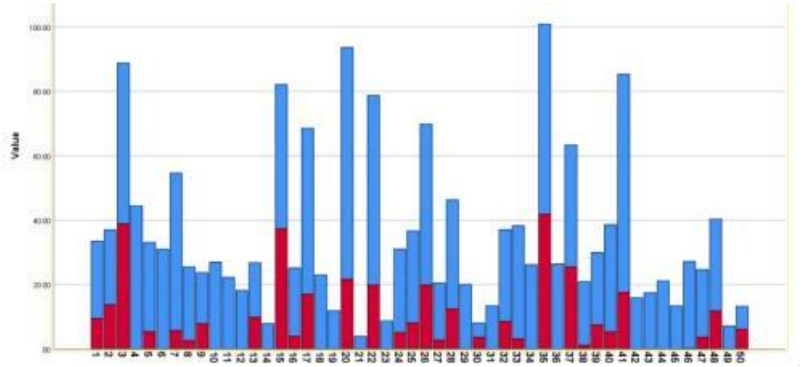
Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 6 (June) (2025)

Figure 4
Individual Performance Difference

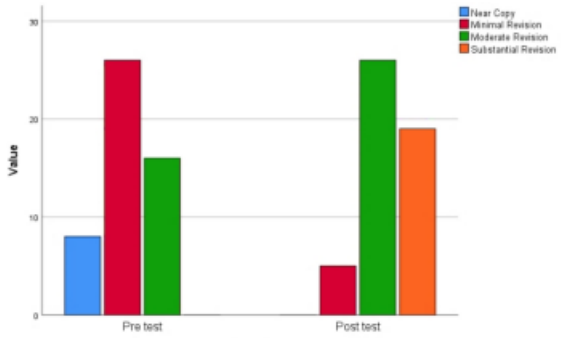


4.3. Mean Difference in Pre and Post-Test Category Wise

The blue bar in the pre-test section indicated the mean score of the near copy in the pre-test, which was completely eliminated in the post-test results. This suggests a significant improvement in the students' ability to paraphrase effectively. In the pre-test results, the red bar displaying minimal revision had a mean percentage of approximately 26 percent which reduced significantly to 4 percent in the post test result. The green bar displaying moderate revision had a mean percentage of around 16 percent in the pretest, which increased significantly to 25 percent in the post test results.

Figure 5

Mean Difference in Pre and Post-Test Category Wise



It was found that the substantial revision as not present in the pre-test results, However, the mean of the substantial revision was almost 19 percent in the post-test results. The results demonstrate that the intervention was successful in improving the students' ability to identify and execute good paraphrasing techniques. The mean percentage of bad paraphrasing type in the pre-test has decreased while the mean percentage of good percentage types has increased in the post-tests results.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 6 (June) (2025)

4.4. Inferential Analysis

In this study, we aimed to investigate the relationship between results of pre-tests and post- tests variables. Before conducting any statistical analyses, we checked the normality of our data using SPSS. Our hypothesis was that the data followed a normal distribution. We used both the Kolmogorov-Smirnov to test this hypothesis.

4.5. Test of Normality

To check whether the data follows the normal distribution or not, Kolmogorov Smirnov Test was applied.

Following assumption was made in order to check the normality:

Null Hypothesis (Ho): The data is normally distributed. Alternative Hypothesis (Ha): The data is not normally distributed.

On the basis of the results of pre-tests and post-tests of the 50 participants, the data set was created in SPSS, where the variables were: pre-test and post test results of paraphrasing. Following were the stats of test of Normality:

Table 4: Kolmogorov-Smirnov

	Statistic	Df	Sig.
Difference	.127	50	.042

The normality of the data was tested using Kolmogorov-Smirnov test. The null hypothesis for these tests is that the data follows a normal distribution. If the p-value associated with the test is less than the significance level (0.05), then we reject the null hypothesis and conclude that the data does not follow a normal distribution. For the Kolmogorov-Smirnov test, the test statistic is 0.127 with 50 degrees of freedom, and the associated p-value is 0.042. Since the p- value is less than 0.05, we reject the null hypothesis and conclude that the data is not the normally distributed according to this test.

4.6. Wilcoxon Signed-Rank Test

The Wilcoxon signed-rank test was used to analyze the paired data in this study since the assumption of normality was violated according to the results of the normality tests.

Table 5: Wilcoxon Signed Rank Test

Null Hypothesis	Sig.	Decision

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 6 (June) (2025)

The median of differences between Pre-test Percentage of copied string of words and Post test Percentage of copied string is insignificant.	.000	Reject the null hypothesis

The significance level is 0.05

On the basis of the results of the two sets of data compared i.e., pre-test and post-test paraphrasing, it was concluded through the implementation of the Wilcoxon signed rank test that the there is a statistically significant difference between the results of the paired samples being compared.

4.7. Qualitative Data Analysis

Interviews of the focus group were conducted in a systematic way. 5 participants from each section were selected in order to conduct a focus group interview of total 10 students. 12 open ended questions were asked to the participants which encouraged students to participate and provide detailed and responses. The questions were designed to explore student's perception, experiences, attitudes and opinions related to the used of SkELL for teaching paraphrasing skill to the students. The interviews conducted within the focus group are transcribed using MAXQDA. These transcriptions are first colour coded in the software and then based on the colour coding, the themes are generated.

4.8. Sentiment Analysis

The sentiment analysis results indicate an overall positive sentiment towards the subject being analyzed. With 8 instances of positive sentiment and 16 instances of slightly positive sentiment,it is clear that the majority of the sentiment expressed is optimistic and favorable. This suggests that the subject has garnered a generally positive response from the users. The presence of only one instance of slightly negative sentiment indicates that there is very minimal negativity associated with the subject, further reinforcing the predominantly positive sentiment observed. Overall, the sentiment analysis portrays a positive perception of the subject.

Figure 6 Sentiment Analysis

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 6 (June) (2025)

4.9. Thematic Analysis

In the context of this study, MAXQDA was utilized to generate three word clouds, aiding in the interpretation of three main themes and facilitating the derivation of subthemes, aligning with the research objective of investigating the effectiveness of an e-tool for teaching paraphrasing skills.

Theme 1: Learning and
Improvement Theme 2:
Support and Resources

Theme 3: User Experience and Evaluation

Figure 7

Word Cloud for Thematic Analysis

4.9.1. Theme 1: Learning and Improvement

1. Learning and Improvement
2. Classroom Application
3. Exploring and Discovering
4. Effort and Change
5. Paraphrasing Skills
6. Beginner and Different

Figure 8
Word Cloud 2 for Thematic Analysis

4.9.2. Theme 2: Support and Resources

1. Tools and Resources
2. Friendly and Supportive
3. Context and Application
4. Explanation and Examples
5. User-friendly and Easy-to-Use
6. Vocabulary and Synonyms

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 6 (June) (2025)

Figure 9
Word Cloud 3 for Thematic Analysis

Theme 3: User Experience and Evaluation

1. Support and Assistance
2. Contextual Understanding
3. Context and Classroom
4. Dimension and Bank
5. Positive Evaluation
6. Problem-solving and Alternatives

4.10. Interactive Word Tree Figure 10

Interactive Word Tree

In an interactive word tree, the root node is labelled as ‘the’. This signifies that “the” serves as the starting point or the main focus of the ord tree. From this central node, several branches extend each representing different aspects related to the word “the” we come across the leaf nodes of the interactive word tree. These leaf nodes consist of specific notes assosiated with “the”. They include descriptors such as “prominent”, effective, feature, synonyms and meaning. Each leaf note provides concise information or key points that contribute to a more nuanced understandng of the word “the”. The leaf nodes also emcompass phrase like “ very useful”, indicating that the information contained within the word tree can be highly benefiial for individuals seeking to enhance their language skills, expand their vocabulary, or deepen their understanding. These leaf notes serve as valuable takeaways, highlighting the importance and practicality of the interactive word tree’s content.

5. Conclusion & Discussion

The research contributes to a deeper understanding of using SkELL in language learning classroom and its impact on enhancing the paraphrasing skill of

ISSN Online: 3007-3154
ISSN Print: 3007-3146

students. They highlight the importance of using Sketch Engine for language learning in influencing academic writing improvement and providing valuable information for practitioners, policymakers, and researchers in the field of linguistics and language learning classroom specially in Pakistan.

The results of this study also have practical implications. They can be utilized to make students independent in terms of doing paraphrasing skill without being dependent on AI powered online tools. By implementing the findings, students and teachers both can be benefitted such as teachers work load and responsibility of explaining each term can be lessen and students can fully take the responsibility of carrying out their academic tasks all by themselves. Moreover, the study provides a foundation for further research and exploration in the field of creating a prototype for making a tool that specifically follows all the aspects of paraphrasing that can assist the teachers in teaching paraphrasing and help students practice and drilling the intext citation formats as well which will boost paraphrasing skill offering opportunities for future investigations to build upon and expand the current knowledge.

It is important to acknowledge the limitations of this study. Such as this experimental study was carried out on merely 50 participants and that on the students of Pakistani private university. These limitations provide opportunities for future research to address and overcome these challenges, enhancing the validity and generalizability of the findings.

In conclusion, this thesis has made a significant contribution to the field of ELT. It has advanced our understanding of effectiveness of SkELL for teaching paraphrasing skill through rigorous analysis and interpretation of the results drawn before and after the treatment using the Casey Keck's taxonomy of paraphrase. The findings have implications for language learning and offer valuable insights for educationists, linguists and all the languages learners and facilitators.

Reference

Atif, A., Richards, D., & Bilgin, A. (2013). A student retention model: empirical, theoretical and pragmatic considerations. Auji, H. (2018). Marketing Views of Modernity, Evangelism and Print Specialization in the American Mission Press Catalogs (1884–1896). *Middle East Journal of Culture and Communication*, *11*(3), 316-354.

Bailey, R. (2006). Physical education and sport in schools: A review of benefits and outcomes. *Journal of school health*, *76*(8), 397-401.

Barrs, K. (2016). Using the sketch engine corpus query tool for language teaching. LANGUAGE TEACHER, 40, 23. Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language learning*, *67*(2), 348-393. Banegas, D. L., Poole, P. M., & Corrales, K. A. (2020). Content and language integrated learning in Latin America 2008- 2018: Ten years of research and practice. *Studies in Second*

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Language Learning and Teaching, *10*(2), 283-305.

Choi, Y. H. (2012). Paraphrase practices for using sources in L2 academic writing. *ENGLISH TEACHING 67*(2), 51-79. Cho, W. I., Kim, J. I., Moon, Y. K., & Kim, N. S. (2020, May). Discourse component to sentence (DC2S): An efficient human- aided construction of paraphrase and sentence similarity dataset. In *Proceedings of The 12th Language Resources and Evaluation Conference* (pp. 6819-6826).

Dung, T. T. (2010). An investigation in paraphrasing experienced by Vietnamese students of English in academic writing. *Da Nang: Da Nang University*.

Fantini, A. (2018). Sketch Engine: breve guida all'uso con esempi da ItTenTen, BNC e EnTenTen. Rassegna Italiana di Linguistica Applicata, 50(1-2), 141-146.

Herman, O., Jakubíček, M., Rychlý, P., & Kovář, V. (2019, October). Word Sense

Ernawati, E., Sinambela, E. A., Cici, C., Silviana, R. J., Azizah, R. N., & Naudalia, S. (2022). The Effect of Social Support and Extraversion Personality on Entrepreneurial Interest in Students. *Journal of Social Science Studies (JOS3)*, *2*(2), 39-44. Gunawan, F., Nababan, M. R., Syukri, H., & Burdah, I. (2022). Revisiting interpretive translation method: A case study of Muhammad Thalib's Quranic translation. *Jordan Journal of Modern Languages and Literatures*, *14*(1), 111-127.

Hidayati, D. (2022). The Effect of Paraphrasing Strategy on the Students' Ability in Comprehending Narrative Text at the Eighth Grade of SMP Negeri 1 Mandrehe. *The Explora*, *8*(1), 25-42.

Injai, R. (2015). *An analysis of paraphrasing strategies employed by Thai EFL students: case study of Burapha University* (Doctoral dissertation, Doctoral dissertation), Burapha University, Chon Buri. Retrieved from http://digital_collect. lib. buu. ac. th/dcms/files/54910112. pdf).

Induction Using Word Sketches. In International Conference on Statistical Language and Speech Processing (pp. 83-91). Springer, Cham.

Jakubícek, M., & Rychlý, P. (2019, December). A Distributional Multi-word Thesaurus in Sketch Engine. In *RASLAN* (pp. 143-147).

James, C. (2016). Discovering English with Sketch. Laurence King Publishing.

Ji, N. Y. (2018). Investigation into validity of paraphrasing task as a writing performance test item for EFL learners., 19(2), 20-29.

Jordan, R. (2003). Academic Writing Course (3rd ed., Vol. 1). Charlbury, Oxfordshire: Bluestone Press.

Kawabata, H., Takai, K., Kojima, M., Nakamura, N., Aoki, S., Nakamura, S., & Masaki, Y. (2013). Castleman-Kojima disease (TAFRO syndrome): a novel systemic inflammatory disease characterized by a constellation of symptoms, namely, thrombocytopenia, ascites (anasarca), microcytic anemia, myelofibrosis, renal dysfunction, and organomegaly: a status report and summary of Fukushima (6 June, 2012) and Nagoya meetings (22 September, 2012). *Journal of clinical and experimental hematopathology*, *53*(1), 57-61.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, *1*(1), 7-36.

Kilgarriff, A., Reddy, S., Pomikálek, J., & Avinesh, P. V. S. (2010, May). A

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Corpus Factory for Many Languages. In LREC. Kasai, N., Fukushima, K., Omori, K., Sugaya, A., & Ojima, T. (2012). Effects of early identification and intervention on language development in Japanese children with prelingual severe to profound hearing impairment. Annals of Otology, Rhinology & Laryngology, 121(4), 16-20.

Kim, M. K. (2020). Investigating the reliability and validity of scores from a paraphrasing test for Korean EFL learners. Unpublished MA thesis, Seoul National University, Seoul, Korea.

Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of second language writing*, *15*(4), 261-278.

Kocincová, L., & Baisa, V. (2015, May). Interactive visualizations of corpus data in sketch engine. In *Proceedings of the Workshop on Innovative Corpus Query and Visualization Tools at NODALIDA 2015, May 11-13, 2015, Vilnius, Lithuania* (No. 111, pp. 17-22). Linköping University Electronic Press.

Kunilovskaya,M. & Koviazina,M.(2017).Sketch Engine: A Toolbox for Linguistic Discovery. Journal of Linguistics/Jazykovedný casopis,68(3) 503-507. https://doi.org/10.2478/jazcas-2018-0006

Lindsay, J. B. (2018). A practical guide to academic writing and publishing. *European Business Review*, *18*(6), 479-490. León-Araúz, P., Martín, A. S., & Reimerink, A. (2018). The EcoLexicon English corpus as an open corpus in Sketch Engine. *arXiv preprint arXiv:1807.05797*.

Pratama, Y., & Prawesti, A. (2022). An analysis of students' writing skills in paraphrasing: a case study of the 5th-semester diploma students of english language and culture department of darma persada university. *Getsempena English Education Journal*, *9*(1), 13-28.

Pertiwi, I. (2019). *Students' Paraphrasing Techniques in Writing (A Study at the Senior Students of English Program of IAIN Parepare)* (Doctoral dissertation, IAIN Parepare).

Pinto, A. (2022). Morceau de Concours (2007): Recycling, Transcription and the Concert-Paraphrase in the Music of Roger Smalley (1943–2015). *Musicology Australia*, 1-25.

Poole, S., Kidd, S. P., & Saeed, K. (2018). A review of novel technologies and techniques associated with identification of bloodstream infection etiologies and rapid antimicrobial genotypic and quantitative phenotypic determination. *Expert Review of Molecular Diagnostics*, *18*(6), 543-555.

Rafiqa, R., Pabbajah, M. T. H., Asfihana, R., Naazaruddin, R., & Latifah, L. (2023). The Effect of Paraphrasing On EFL Students' writing at Indonesian Higher Education. *LLT Journal: A Journal on Language and Language Teaching*, *26*(1), 335- 345.

Rahman, A. B. S., Ta, H. T., Najjar, L., & Gelbukh, A. (2022). Paraphrase identification: Lightweight effective methods based features from pre-trained models. In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022)*.

Sarair, S., Astila, I., & Nurviani, R. (2019). Acehnese students' skills in paraphrasing English texts: a case study at two universities in banda aceh. *Proceedings of EEIC*, *2*, 151-159.

Sanjay, N. I. (2021). *Students' paraphrasing skills in academic writing: a case*

ISSN Online: 3007-3154
ISSN Print: 3007-3146

study at English literature department of Muslim university of Indonesia (Doctoral dissertation, Universitas Hasanuddin).

Sulistyaningrum, S. D. (2021). Employing Online Paraphrasing Tools to Overcome Students' Difficulties in Paraphrasing. *Stairs: English Language Education Journal*, *2*(1), 52-59.

Stewart, D., Biasi, P., & Binelli, A. Using sketch engine: an analysis of five adverbs.

Tran, T. T. T., & Nguyen, H. B. (2022). The effects of paraphrasing on EFL students' academic writing. *Journal of Language and Linguistic Studies*, *18*(1), 976-987.

Thomas, J. A. (2014). *Meaning in interaction: An introduction to pragmatics*. Routledge.

Thienthong, A. (2018). Inferential Reasoning as a Means of Avoiding Plagiarism in Paraphrasing. *PASAA: Journal of Language Teaching and Learning in Thailand*, *56*, 9-32.

Vannestål, M. E., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, *19*(3), 329-350.

Virisi, F., Miliha, L., & Cahyadin, W. (2022). Paraphrasing Techniques Used by Students at English Department of Halu Oleo University. *Journal of Teaching of English*, *7*(2).

Yahia, I., & Egbert, J. L. (2023). Supporting Non-Native-English Speaking Graduate Students with Academic Writing Skills: A Case Study of the Explicit Instructional Use of Paraphrasing Guidelines. *Journal of Writing Research*, *14*(3), 305–341. https://doi.org/10.17239/jowr-2023.14.03.01

Zou, B., Smith, S., & Hoey, M. (Eds.). (2015). *Corpus linguistics in Chinese contexts*. Houndsmill, Basingstoke, Hampshire, New York: Palgrave Macmillan.