



The Impact of Learning-Oriented Assessment on EFL Learners' Vocabulary Learning and Retention in Online Classes

Seyed, Farzaneh Sadat and Tavassoli, Kobra

Department of ELT, Karaj Branch, Islamic Azad University, Karaj, Iran

[10.22034/jelt.2023.13448.1043](https://doi.org/10.22034/jelt.2023.13448.1043)

Abstract

Different techniques can be used to enhance EFL students' vocabulary learning. One such technique can be learning-oriented assessment (LOA) which may have a positive impact on vocabulary learning. To investigate the effect of LOA on EFL learners' vocabulary learning and retention, 60 intermediate EFL learners were selected. After checking the homogeneity of the participants by using an Oxford Placement Test (OPT), 60 participants were selected and they took a pretest of vocabulary. Afterward, they were divided into 2 groups, the experimental group received the LOA treatment and the control group received the placebo. Based on Carless framework (2007), LOA was applied in the experimental group for 40 hours. Then, both groups took the post-test and the delayed post-test after 2 weeks. The reliability of the pretest, post-test, and delayed post-test was considered acceptable. Two repeated-measures ANOVAs were used to determine whether there were statistical differences between the pretest, post-test, and delayed post-test of each group. Also, two independent-samples t-tests were used to determine the differences between the post-test of the two groups and their delayed post-tests. The results showed that LOA significantly affected the experimental group's vocabulary learning. On the other hand, LOA did not have any significant impact on the learners' retention of vocabulary in the delayed post-test. The most important implication of the study is that for better retention of words, more hours of LOA instruction are needed.

Keywords: EFL learners, learning-oriented assessment (LOA), online class, vocabulary learning, vocabulary retention

Introduction

Learning-oriented assessment is characterized as an assessment where an essential center is on the potential to create profitable understudy learning forms. In specific, the 'right kind' of summative assessment can be productive in fortifying fitting understudy learning miens and behaviors. Summative evaluation can be learning-oriented when, for illustration, it empowers profound instead of surface approaches to learning and when it advances a high level of cognitive engagement reliably over the term of a module. The forms of working toward well-designed summative assessment can moreover bear openings for formative assessment techniques, such as peer criticism, understudy self-evaluation, and related instructor criticism (Carless, 2015).

The evaluation strategy and the open information system were connected to a project-based course on web building. This involvement has given positive confirmations since the review estimation was sponsored up with appraisal confirmations and calculated with less exertion (Traverso-Ribon et al., 2016). The term learning arranged assessment (LOA) has been broadly utilized to encapsulate the thought that all shapes of evaluation ought to advance learning (Green, 2016). In arranging for evaluation to design enduring learning, Carless (2007) distinguished three LOA standards: (1) evaluation errands must fortify sound learning hones; (2) learners must effectively lock in with the appraisal exercises; and (3) the evaluation must give fitting and convenient input which learners can eventually feedforward (Carless, 2007). LOA moves past the conventional double situating of evaluation's developmental and summative purposes to a systemic and energetic relationship between teaching, learning, and appraisal (Jones & Saville, 2016; Turner & Purpura, 2016).

Literature Review

Based on Carless (2015) who introduced a framework for learning-oriented assessment, the model centers on three interrelated forms: the evaluation assignments which the learners give it a try; students' advancement of self-evaluative capacities; and understudy engagement with criticism. These three strands are investigated through the investigation of evaluation honed in the setting. The inquiry about strategy includes in-depth classroom perceptions of five beneficiaries of grants for instructing greatness over different disciplines; semi-structured interviews with these

instructors and a test of their understudies. The results clarify the assessment tasks expected and practiced in teaching based on the basic detailed examinations to create understudy understandings of quality work, and 'same day feedback' to advance convenient discoursed with understudies.

Carless (2007) claimed that practitioners have to create and after that systematize appraisal feedback to maximize its potential for students' activity in line with the LOA system.

Learning-Oriented Assessment

Learning-oriented assessment speaks to an endeavor to accommodate developmental and summative evaluation and center all evaluation on the advancement of profitable understudy learning. Learning-oriented assessment comprises three interlocking measurements: assessment errands as learning errands; understudy inclusion in appraisal; and the closing of criticism circles. The learning-oriented evaluation extension was organization's endeavor to advance and spread valuable appraisal hones: its qualities were in its worldwide profile and distributions, while its confinements included a need for supported effect on appraisal at the course and program levels. Boundaries such as responsibility and doubt are considered and a few conceivable ways forward for learning-oriented evaluation are proposed (Carless, 2009).

In arrange to meet the learning needs of understudies and prerequisites of instructive policies, researchers and instructors have made awesome endeavors to change the appraisal approaches. Learning-oriented appraisal has been created against the scenery of social change and instructive change emphasizing a learning society (Colantonio, 2005) and classroom assessment (Antoniou & James, 2014).

Based on Carless (2007) framework, there are some steps for the implementation of learning-oriented assessment. The primary and most vital strand of LOA is represented by the term appraisal assignments as learning errands. This conceptualization holds that when assessment errands optimize the required learning results, understudies are prepared for profound learning encounters by advancing toward these results. The second component of LOA is understudying inclusion in evaluation so that they create better improvement; a much better; a higher; a stronger; and an improved understanding of learning objectives with criteria and guidelines. Thirdly, for evaluation to advance learning, understudies have to get suitable feedback that they can utilize to

'feedforward' into future work. Feedback is likely to be more successful when understudies are cognizant of criteria and are observing their advance toward the expressed benchmarks.

There are some guidelines that were introduced by Carles (2007, p. 59) related to the learning-oriented principles:

Guideline 1: Assessment tasks ought to be outlined to stimulate sound learning practices among understudies.

Guideline 2: Evaluation ought to include understudies effectively in locks in with criteria, quality, their possess and/or peers' performance.

Guideline 3: Feedback ought to be convenient and forward-looking so as to bolster current and future understudy learning.

Vocabulary Learning and Retention

To influence the learning time for vocabulary learning, there needs to be wealthy instruction with a sensible sum of time going through each word. At least this is often likely someplace between 3 to 5 minutes per word. Some distributed ponders of lexicon education appear that this could be an awfully time-consuming process (Nation, 2021).

One of the most extremely difficult perspectives of dialect learning is vocabulary. A wide vocabulary is basic for successful and valuable communication. Subsequently, understanding the methodologies that dialect learners utilize to memorize lexicon may be a matter of extraordinary significance (Al-Khresheh & Al-Ruwaili, 2020). There are lots of problems with learning new vocabulary such as knowing the implications of new words, articulating new words, utilizing new words accurately, and memorizing and spelling new lexicon (Afzal, 2019).

Khoshsima and Khosravi (2021) worked on quasi-experimental research with three groups of WhatsApp-based lexicon instruction, Anki-based lexicon instruction, and the traditional one, to measure vocabulary retention through the mentioned applications. Anki-based and WhatsApp-based lexicon instruction were viable in upgrading their lexicon retention. The outcomes of MANOVA appeared that there were noteworthy contrasts among the three bunches in their lexicon

retention and using applications significantly affected the vocabulary retention of the learners in comparison with the traditional one.

Baleghizadeh and Ashoori (2010) searched the effect of word lists and keyword methods on vocabulary retention of learners, results appeared that the keyword strategy delivered better recall compared to the word list strategy, recommending promising educational value for its utility.

Also, web-based language teaching has a significant impact on learners' vocabulary retention (Hajebi, et al., 2018), and using self-developed mobile applications has a significant impact on learners' retention (Ma & Yodkamlue, 2019).

In this study, the impact of learning-oriented assessment on the learners' vocabulary learning and retention has been examined. The following two research questions have been posed:

RQ1. Does the learning-oriented assessment have any effects on EFL learners' vocabulary learning?

RQ2. Does the learning-oriented assessment have any effects on EFL learners' vocabulary retention?

Methodology

Participants

The number of participants in the current study was 60 EFL learners who consisted of male and female learners. They were selected as a result of their performance based on an English language proficiency test. The participants took the Oxford Placement Test (OPT) to ensure their homogeneity in terms of proficiency level and the intermediate learners were selected. Therefore, 60 students who scored between one standard deviation above and below the mean were selected as final participants. They consisted of female and male learners who voluntarily took part in the study with the age range of 12-45 years old. Concerning the sampling of the participants, the researcher benefited from convenience sampling to select those potential candidates who were available at the time of the research (Dornyei, 2007). Afterward, the participants took a vocabulary test (<https://www.oxfordonlineenglish.com/english-level-test/vocabulary>), then they were divided into two groups. After 40 hours of the learning-oriented assessment process and involving tasks in

class and after the class, collaboration, and cooperation, the teacher's feedback, the learner's feedback, summative and formative assessments, they took another test as posttest and a delayed posttest as the retention test. The differences between these two tests showed whether the learning-oriented assessment had a significant effect on the learners' vocabulary learning and retention or not.

Materials and Measures

In order to find out the effect of learning-oriented assessment on learners' vocabulary learning and retention, the participants were asked to take two different vocabulary tests with the same difficulty, meanwhile, they were supposed to attend the summative and formative assessments that the teacher provided. The teacher has observed them by giving feedbacks related to their progress, and they were supposed to leave comments on the other students' exams to help their peers to develop. The materials that were used can be listed as:

Oxford placement test

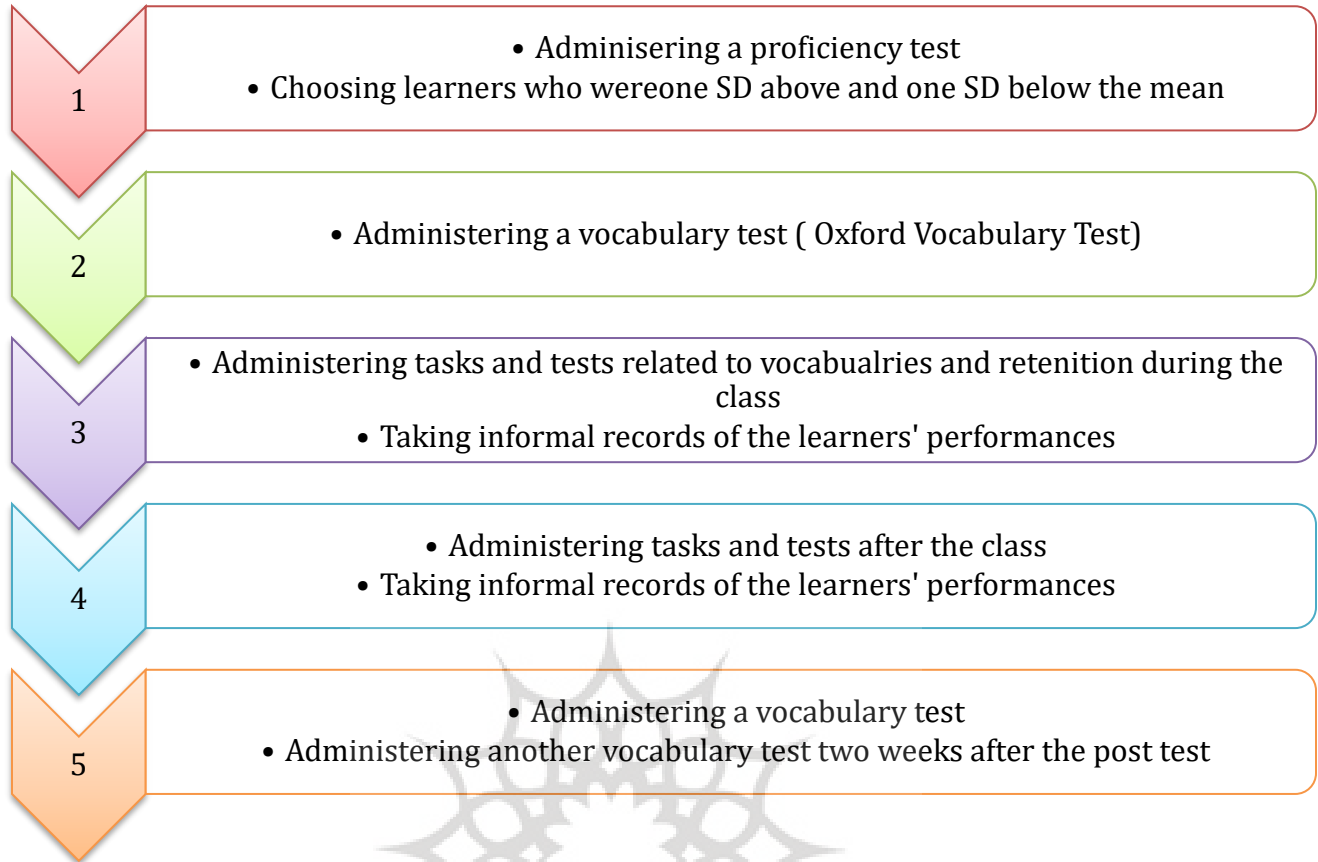
Oxford vocabulary test as pretest

Oxford vocabulary test as posttest

Procedure

The order of the major steps of the study which were done for collecting data is clarified using the following figure:

شعبه پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی



The participants were asked to do the Oxford Placement Test (OPT). After choosing students who were one SD above and below the mean based on their scores on the OPT, they were divided into two groups. For calculating the homogeneity of the two groups, an independent samples t-test was used and the results showed the homogeneity of variances between the two groups. One group received learning-oriented assessment (experimental group) and another group was the control group. The teacher of the two groups was the same to avoid the effect of the teacher on the results. The first group was asked to participate in the assessments related to vocabulary tests, and after each assessment, the teacher gave feedback to help them to improve and learners were asked to leave comments for their classmates as well. In some cases, they were asked to do the same test twice. For measuring the retention, a delayed posttest was used.

Design

The design assumes that a unique plan has been used to answer the questions. In this study, quasi-experimental design was used since there were two groups (experimental and control) of

students who were selected non-randomly (Dornyei, 2007). The functions and measurement scales of the variables in this study are presented in Table 1.

Table 1.

Types and Measurement Scales of the Variables

Variable	Function	Measurement and scale
Learning Oriented Assessment	Independent	
Vocabulary Learning	Dependent	Interval
Vocabulary Retention	Dependent	Interval

Results

Checking the Reliability of the Pretest and Posttest

Table 2.

Reliability Statistics of the Pretest and Post-test

	Cronbach's Alpha	N of Items
Pretest	.743	40
Post-test	.767	50

Table 2 elaborates on the results of the reliability statistics of the pretest and post-test. As it is mentioned, Cronbach's alpha value for 40 items related to the vocabulary test in the pretest is 0.74 which is considered a reliable test. Also, Cronbach's alpha value for 50 items related to the vocabulary test for the post-test is 0.76 which is also considered a reliable test. (Bland & Altman, (1997).

Investigating Research Question 1:

Table 3.

Descriptive Statistics of the Vocabulary Tests of the Pretest and Post-Test of the Experimental Group, the Pretest and Posttest of the Control Group, and the Delayed Posttest of Experimental and Control Groups

	N	Mean	Std. Deviation
Pretest Control group	30	21.5000	5.43774
Posttest Control group	30	21.5000	5.23746
Pretest Experimental group	30	21.4000	5.67268
Post-test Experimental group	30	27.4667	5.39306
Delayed post-test Experimental group	30	22.7000	4.91409
Delayed post-test Control group	30	19.9667	4.92344
Valid N (List-wise)	30		

Table 3, which reports the descriptive statistics of the vocabulary tests of the pretest and posttest of the experimental group, the pretest and posttest of the control group, and the delayed posttest of the experimental group and control group, points to a difference between the analyzed sets of scores. In fact, the mean of the pretest in the experimental group is smaller than the mean of the posttest. Also, the mean of the delayed post-test is higher in comparison with the pretest. On the other hand, the mean of the posttest and the delayed posttest in the experimental group is higher than the mean of the posttest and delayed posttest in the control group ($\bar{X}_{\text{pretest}} = 21.40$, $\bar{X}_{\text{posttest}} = 27.46$, $\bar{X}_{\text{delayed posttest}} = 22.70$; $SD_{\text{pretest}} = 5.67$, $SD_{\text{posttest}} = 5.39$, $SD_{\text{delayed posttest}} = 4.91$).

An independent sample T-test was used to identify if the learning oriented assessment had any significant impact on the learners' vocabulary learning.

Table 4.

Independent Samples T-Tests between the Pretest of the Control and Experimental Groups and the Posttest of the Control and Experimental Groups

		Levene's Test for Equality of Variances		t-test for Equality of Means		Sig. (2- tailed)	Mean Difference	Std. Error Difference
		F	Sig.	t	df			
Pretest vocabulary	Equal variances assumed	.022	.883	.070	58	.945	.10000	1.43467
	Equal variances not assumed			.070	57.897	.945	.10000	1.43467
Posttest vocabulary	Equal variances assumed	1.183	.281	- 4.347	58	.000	-5.96667	1.37254
	Equal variances not assumed			- 4.347	57.950	.000	-5.96667	1.37254

Table 4 which is dedicated to the results of the independent samples t-test to measure the significance of the observed difference statistically proves the improvement caused by the treatment which was the learning-oriented assessment of vocabulary. Based on the independent samples t-test, the sig level in comparison between the post-test of the experimental group and the control group is lower than the research confidence interval which is 0.05, so the null hypothesis is rejected, and there is a significant difference between the posttests of the two groups.

A paired samples t-test was run for identifying the significance of the difference between the pretest and the posttest of the experimental group.

Table 5.

Paired Samples T-Test of Pretest and Posttest of the Experimental Group

Paired Differences					
Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2- tailed)

Pretest vocabulary -	-6.06667	8.03412	1.46682	-4.136	29	.000
Posttest vocabulary						

Table 5 which is dedicated to the results of the paired samples t-test of the significance of the observed difference statistically proves that the improvement caused by the treatment which was the learning-oriented assessment, improved the EFL learners' vocabulary. Based on the paired samples t-test, the sig level in comparison between the pretest and posttest of the experimental group is lower than the research confidence interval which is 0.05, so the null hypothesis is rejected, and there is a significant difference between the pretest and the post-test of the experimental group. ($t_{29}=4.13$; $\alpha=0.05$, $\rho=0.000$, $\rho<\alpha$).

Investigating Research Question 2

Table 6.

Repeated Measures ANOVA between the Pretest, Posttest, and Delayed Posttest of the Experimental Group

Tests of Within-Subjects Effects

Measure: Vocabulary

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Vocabulary test	Sphericity Assumed	47.022	2	23.511	1.067	.351
	Greenhouse-Geisser	47.022	1.449	32.450	1.067	.334
	Huynh-Feldt	47.022	1.505	31.238	1.067	.336
	Lower-bound	47.022	1.000	47.022	1.067	.310
Error(vocabulary test)	Sphericity Assumed	1277.644	58	22.028		
	Greenhouse-Geisser	1277.644	42.023	30.403		
	Huynh-Feldt	1277.644	43.654	29.268		
	Lower-bound	1277.644	29.000	44.057		

Table 6 is elaborating on the repeated measures ANOVA to measure whether the difference between the pretest, posttest, and delayed posttest was significant or not. The sig level is 0.35 which is more than 0.05 so the null hypothesis is not rejected. In conclusion, there is no significant relationship between the pretest, posttest, and delayed post-test of the experimental group.

To be sure about the effect of the learning-oriented assessment on the results, an ANOVA test was also used for comparing the pretest, post-test, and delayed post-test of the control and experimental groups.

Table 7.

One Way ANOVA of Pretest, Posttest, and Delayed Posttest of Control and Experimental Groups

		Sum of Squares	df	Mean Square	F	Sig.
Pretest of the Control group and experimental group	Between Groups	.150	1	.150	.005	.945
	Within Groups	1790.700	58	30.874		
	Total	1790.850	59			
Post-test of the Control group and Experimental group	Between Groups	534.017	1	534.017	18.898	.000
	Within Groups	1638.967	58	28.258		
	Total	2172.983	59			
Delayed post-test of the Control group and Experimental group	Between Groups	112.067	1	112.067	4.632	.036
	Within Groups	1403.267	58	24.194		
	Total	1515.333	59			

Table 7 clarifies the one-way ANOVA between the pretest, post-test, and delayed post-test of both the control group and the experimental group. As a matter of fact, the sig level in the comparison between the posttest of the control group and the experimental group is less than 0.05 and as a result, the null hypothesis is rejected. So, there is a significant difference between the means of the experimental and the control group, and the learners have done better in the experimental group because of the effect of learning-oriented assessment. Otherwise, there is no significant difference between the pretest and the delayed posttest of the experimental group and control group.

Discussion

Organizing includes making sure that the conditions that support lexicon learning have a chance to occur. These conditions incorporate reiteration, taking notes, recovery, assembly, and utilizing words in varied contexts, elaboration, and thinking consideration (Nation, 2021). Based on this research, learning-oriented assessment has a significant impact on the learners' vocabulary learning. They could enhance their ability to learn vocabulary by following the rules related to the learning-oriented assessment.

On the other hand, the learning-oriented assessment did not have any significant impact on the learners' retention ability for memorizing the words. Nation (2020) claimed that the amount of vocabulary learning is important. Regularly, on immediate posttests of responsive information taken after lexicon instruction, less than 50% of the words instructed are really recalled. A rising approach to classroom-based evaluation that has picked up increasing consideration within the past decade is learning-oriented assessment, conceptualized as an appraisal that manages power to the learning introduction of classroom evaluation, putting specific accentuation on the dynamic engagement of learners in evaluation and input (Carless, 2015; Turner & Purpura, 2016). Furthermore, other studies related to vocabulary retention suggested that advanced and intermediate learners could be better in comparison with beginner ones (Kohnke et al., 2021). Moreover, vocabulary retention can be enhanced by contextual clues as well (Huyen, 2019). Learning vocabulary in English as a second or foreign language is crucial and it can be increased by using software and mobile assistance (Yang et al., 2021).

Conclusion

Vocabulary is the building block of language learning which acts as the kernel for all dialect abilities (Arjmandi & Aladini, 2020). The present study endeavors to examine the impact of applying learning-oriented assessment on the learners' vocabulary ability and the retention of vocabulary. Based on the findings, learning-oriented assessment can enhance the number of vocabulary that students have to memorize or learn in 40 hours of learning-oriented assessment which were held in the classroom and out of the classroom. The implementation of learning-oriented assessment was based on Carless' (2007) framework with students' and teachers' feedback, peer assessments, and self-assessments. On the other hand, a learning-oriented assessment cannot increase the score of retention in a 40 hours program. Learners might need more than 40 hours of treatment. It is believed that the students can memorize approximately 50 percent of the vocabulary which is in input. By giving more time to the experiment, the learners might be able to learn more vocabulary. The most important implication of the study is that for better retention of words, more hours of LOA instruction are needed. Limitations and delimitations of the study are related to the following factors: the first delimitation of the study was selecting merely upper-intermediate level learners. Second, the participants of the study were delimited to the context of private language institutes, to make the study more manageable. Third, this study was delimited to a quantitative data collection methodology. Fourth, the current research did not benefit from true experimental research. Last but not least, this study did not take into account EFL learners' gender.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
رتال جامع علوم انسانی

References

- Afzal, N. (2019). A study on vocabulary-learning problems encountered by BA English majors at the university level of education. *Arab World English Journal (AWEJ)*, 10.
- Al-Khresheh, M. H., & Al-Ruwaili, S. F. (2020). An exploratory study into vocabulary learning strategies used by Saudi EFL learners. *Journal of History Culture and Art Research*, 9(2), 288-302.
- Antoniou, P., & James, M. (2014). Exploring formative assessment in primary school classrooms: Developing a framework of actions and strategies. *Educational Assessment, Evaluation and Accountability*, 26(2), 153-176.
- Arjmandi, M., & Aladini, F. (2020). Improving EFL learners' vocabulary learning through Short Story Oriented Strategy (SSOS). *Theory and Practice in Language Studies*, 10(7), 833-841.
- Baleghizadeh, S., & Ashoori, A. (2010). The effect of keyword and word list methods on immediate vocabulary retention of EFL learners. *Pakistan Journal of Social Sciences (PJSS)*, 30(2), 723-754.
- Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *BMJ*, 314(7080), 572.
- Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57-66.
- Carless, D. (2009). Learning-oriented assessment: Principles, practice, and a project. In L. Meyer, S. Davidson, H. Anderson, R. Fletcher, P. Johnston, & M. Rees (Eds.), *Tertiary assessment and higher education student outcomes: Policy, practice, and research* (pp. 79-90). Ako Aotearoa.
- Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education*, 69(6), 963-976.
- Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Hajebi, M., Taheri, S., Fahandezh, F., & Salari, H. (2018). The role of web-based language teaching on vocabulary retention of adult pre-intermediate EFL learners. *Journal of Language Teaching and Research*, 9(2), 372-378.
- Huyen, V. T. (2019). Effects of using contextual clues on English vocabulary retention and reading comprehension. *International Journal of English Literature and Social Sciences (IJELS)*, 4(5).

- Khoshsiman, H., & Khosravi, M. (2021). Vocabulary retention of EFL learners through the application of ANKI, WhatsApp, and traditional methods. *Journal of Foreign Language Teaching and Translation Studies*, 6(4), 77-98.
- Kohnke, L., Zou, D., & Zhang, R. (2021). Exploring discipline-specific vocabulary retention in L2 through app design: Implications for higher education students. *RELC Journal*, 52(3), 539-556.
- Ma, X., & Yodkamlue, B. (2019). The effects of using a self-developed mobile app on vocabulary learning and retention among EFL learners. *PASAA: Journal of Language Teaching and Learning in Thailand*, 58, 166-205.
- Nation, P. (2021). Is it worth teaching vocabulary? *TESOL Journal*, 12(4), e564.
- Ortlieb, E., & Cheek, E. H., Jr. (Eds.). (2012). *Using informative assessments towards effective literacy instruction*. Emerald Group Publishing.
- Traverso-Ribon, I., Balderas-Alberico, A., Doderio, J. M., Ruiz-Rube, I., & Palomo-Duarte, M. (2016). Open data framework for sustainable assessment of project-based learning experiences.
- Turner, C. E., & Purpura, J. E. (2016). Learning-oriented assessment in second and foreign language classrooms. In *Handbook of Second Language Assessment* (pp. 255-274). De Gruyter Mouton.
- Vocabulary quiz. (2019). Retrieved from <https://www.oxfordonlineenglish.com/english-level-test/vocabulary>
- Yang, X., Kuo, L. J., Eslami, Z. R., & Moody, S. M. (2021). Theoretical trends of research on technology and L2 vocabulary learning: A systematic review. *Journal of Computers in Education*, 8(4), 465-483.