ReCALL (2020), 32: 2, 195–212
doi:10.1017/S0958344020000038

ARTICLE

Using image-to-text recognition technology to facilitate vocabulary acquisition in authentic contexts

Rustam Shadiev
Nanjing Normal University, China (rustamsh@gmail.com)

Ting-Ting Wu
National Yunlin University of Science and Technology, Taiwan (danytingting@gmail.com)

Yueh-Min Huang*
National Cheng Kung University, Taiwan (huang@mail.ncku.edu.tw)

Abstract

A vocabulary acquisition learning activity was designed and a learning system featuring image-to-text recognition technology to support the activity was developed. The effectiveness of the system with regard to facilitating vocabulary acquisition was tested. The perceptions of learners toward this tool and the affordances of the system for vocabulary acquisition were also explored. To this end, we designed an experiment in which 40 native speakers of Russian learning English as a foreign language from an elementary school participated. They were assigned to either a control condition or an experimental condition. All learners learned new vocabulary in class and then applied their new knowledge to contexts with a realistic simulation of the real world by completing a learning task. The learners in the control group used a traditional approach (e.g. the learners learned vocabulary from corresponding pictures in a textbook), whereas the learners in the experimental group used the proposed learning system (e.g. the learners learned vocabulary using the system). A pre-test–post-test/delayed post-test design was employed to test the effectiveness of the treatment on vocabulary acquisition. Learner perceptions and perceived affordances of the system for vocabulary acquisition were explored through a questionnaire survey and interviews. The quantitative results showed that the learners in the experimental group outperformed their counterparts on both the vocabulary post-test and delayed post-test. The qualitative results revealed that most learners in the experimental group had positive perceptions of the system. In addition, the qualitative results showed the three main categories of affordances. Based on these results, several suggestions and implications are provided for the teaching and research community.

Keywords: image-to-text recognition; vocabulary acquisition; authentic contexts; elementary education; EFL

1. Introduction

Vocabulary is a vital component of language competence (Duman, Orhon & Gedik, 2015; Nation, 2001; Wu, 2018). As stated by Ansari and Sabouri (2016: 89), "Words are the building blocks of a language since they label objects, actions, ideas without which people cannot convey the intended meaning." Following this notion, Nation (2001) has suggested that vocabulary carries the

Cite this article: Shadiev, R., Wu, T.-T. & Huang, Y.-M. (2020). Using image-to-text recognition technology to facilitate vocabulary acquisition in authentic contexts. *ReCALL* 32(2): 195–212. https://doi.org/10.1017/S0958344020000038

*All correspondence regarding this publication should be addressed to Dr Huang (Email: huang@mail.ncku.edu.tw).

© The Author(s), 2020. Published by Cambridge University Press on behalf of European Association for Computer Assisted Language Learning. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

basic information learners want to understand and express. Therefore, enriching acquisition of vocabulary is a priority for any language learning program (Tanaka, 2017).

Recent evidence suggests that memorization of vocabulary has long been a common way for many learners to learn lexical items (Pan, 2017). Learners usually learn vocabulary words by rote from a list along with their equivalents in a native language (Alemi, Sarab & Lari, 2012). Nation (2013) suggested that rote memorization takes place when a learner memorizes material by repeating it over and over again until it is learned. According to Tanaka (2017), rote learning is more about memory or habit, rather than understanding (Pan, 2017). Wu (2018) argued that memorization is a shallow educational practice that does not lead to deep learning.

In order to understand a word and remember it for a long time, one must use it for the purpose of communication. This type of practice will ensure that language learners both remember and understand vocabulary (Duman *et al.*, 2015). This notion has been especially highlighted because of the input and output hypotheses (Krashen, 1989; Swain, 1985). According to these hypotheses, the language learning process includes two functions: (1) language input, where a learner receives learning information, and (2) language output, where a learner applies learned information through producing language output. Harmer (2007) argued that both functions are essential for language learning and that a balance should be kept between them. Therefore, the cognitive involvement of learners in vocabulary acquisition should be increased from mere memorization to language production (Abdous, Camarena & Facer, 2009). Learners should initiate learning, actively search for helpful resources, and have a chance to apply newly learned knowledge in a new context other than in the context where it was learned initially (Nation, 2007). This can be achieved by learners incorporating learning activities into their daily routines.

One of the most common vocabulary acquisition strategies is the use of pictures. Scholars have suggested that vocabulary acquisition with both labels and pictures is beneficial and is more effective compared to vocabulary acquisition with labels only (Lin & Yu, 2017). This is because providing pictures enables various parts of the brain to operate simultaneously and thereby improves cognitive ability. This notion is supported by the dual-coding theory (Paivio, 1986). According to this theory, visual and verbal information are processed in different parts of the brain. The visual channel processes visual information and produces pictorial representations, whereas the verbal channel processes verbal information and produces verbal representations. Both visual and verbal information are selected and held in both visual and working memory. A learner then mentally builds connections that organize information into cause-and-effect chains. Finally, the visual model, verbal mental model, and prior knowledge are merged through constructing referential connections among them. Peker, Regalla and Cox (2018) suggested that remembering information can be greatly enhanced after such associations are created. For this reason, educators frequently use pictures as a visual aid for vocabulary acquisition.

The importance of making context and learning inseparable has been emphasized by many scholars (Herrington & Herrington, 2006). Related studies suggest that classroom learning is abstract and disconnected from real-life scenarios when schools ignore the interdependence of context, situation, and cognition (Collins, 1988). However, learning in authentic contexts helps achieve effective, meaningful learning (Kiernan & Aizawa, 2004; Kukulska-Hulme & Viberg, 2018).

The theory of authentic learning asserts that learning is context related, so the importance of learning in a specific context should always be emphasized. Herrington and Herrington (2006) discussed several core features of an authentic environment: (1) It provides authentic contexts that reflect the way the knowledge will be used by learners in real life; (2) it provides authentic activities that have real-world relevance, ideally ones that present complex tasks to be completed over a sustained period of time; and (3) it promotes reflection and enables authentic learning assessment within the tasks. Therefore, authentic contexts are very meaningful and relevant to learners during their learning process (Collins, 1988).

Following this notion, the importance of "seamless learning" was emphasized by Wong (2013). The scholar suggested that seamless learning takes place when in-class and out-of-class learning

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

experiences are linked. Therefore, Kuh (1996) encouraged learners to make the best of learning resources that exist both inside and outside of the classroom during the seamless learning process. Consequently, learners can learn whenever and wherever they are keen to learn using different available resources in a variety of scenarios and situations (Shadiev & Yang, 2020).

New technologies are being widely employed in the language learning process, including vocabulary acquisition (Ros i Solé, Calic & Neijmann, 2010; Shadiev, Hwang & Huang, 2017; Shadiev, Liu & Hwang, in press). This is because technologies enable language learners to learn seamlessly – that is, link classroom learning with learning in the real world (Shadiev, Hwang & Liu, 2018; Wong, 2013). Technologies also help learners learn vocabulary with multimedia texts, sound, and video rather than only with printed texts. Furthermore, related studies (Kurt & Bensen, 2017; Lin & Yu, 2017) have indicated that mobile technology enables vocabulary acquisition to become portable, socially interactive, and context sensitive. Moreover, mobile technology multimedia tools allow learners to create their own content (Comas-Quinn, Mardomingo & Valentine, 2009). That is, learners can practice their language skills in the real world by creating their own textual content, taking photos, and recording audio and video files using multimedia tools. For example, language learners can study a language in class and apply their new knowledge to the real world. When language learners engage with the real world, they can create their own learning content instead of using that created or provided by their instructors. Thus, instructors only have to guide learners.

Scholars reveal the great potential for mobile learning systems to support vocabulary acquisition (Nah, White & Sussex, 2008). For example, Kurt and Bensen (2017) employed Vine, a mobile application on which learners recorded and shared video clips with a maximum length of six seconds. The results showed that practice with Vine is effective and improves vocabulary acquisition. Lin and Yu (2017) carried out a mobile-assisted vocabulary acquisition program featuring multimedia presentations. During the program, the participants received MMS messages with four sets of target words presented in text mode, text-picture mode, text-sound mode, and text-picture-sound mode. The results revealed that audio input helped the participants recall the meanings of new words and reduced their cognitive load related to learning new words more effectively than was the case under the other conditions. Wu (2018) proposed a mobile English vocabulary practice system. It imitated the popular block elimination game and combined article, difficulty, and teacher model test items in accordance with curriculum objectives and demands. The results showed that the learners who used the system learned more effectively and that they were more motivated to learn than their counterparts without the system.

2. Research motivation and questions

A review of related studies has shown several important issues in vocabulary acquisition programs that could potentially be addressed in the present study. For example, in earlier studies, although learning activities were designed around the learners, learning content for vocabulary acquisition was provided by the instructor. That is, learning activities were designed based on an instructor-centered approach (Kurt & Bensen, 2017; Lin & Yu, 2017; Wu, 2018). In the present study, a learner-centered approach to vocabulary acquisition was used. We asked learners to learn vocabulary they were interested in that was relevant and meaningful to them. Furthermore, many earlier related studies focused on language input only and neglected language output.

For the present study, a learning activity was designed based on both language input and language output. That is, the learners learned new vocabulary in class, and then they applied their newly learned knowledge to contexts with a realistic simulation of the real world by producing language output. To this end, we developed a learning system featuring image-to-text recognition (ITR) technology. Learners took a picture of an object they were interested in, and the ITR tool simultaneously generated the word's label in the target language. This is a novel approach to

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

Table 1. Profile of participating learners

	Control group		Experimental group	
Category	Frequency	Percentage	Frequency	Percentage
Gender				
Male	8	40	9	45
Female	12	60	11	55
Age (years)				
10	17	85	18	90
11	3	15	2	10

facilitate vocabulary acquisition and, to the best of our knowledge, this approach has not been previously described in any mobile-assisted language learning research.

In this study, we aimed to test the feasibility of our novel approach; specifically, we explored the effectiveness of ITR technology application on vocabulary acquisition. Our experiment used a pre-test–post-test/delayed post-test design and compared the learning performance of the learners under both control and intervention conditions. Furthermore, the perceptions of the learners in the intervention condition towards the ITR technology and their satisfaction with learning with it were measured. Affordances of the ITR tool were also investigated.

The following research questions were addressed: (1) Do learners who study using ITR technology learn vocabulary better compared to those who study using traditional methods? (2) How do learners perceive their learning experiences using ITR technology? (3) What are the affordances of ITR technology for vocabulary acquisition?

3. Methods

Forty native speakers of Russian learning English as a foreign language (EFL) from an elementary school participated in this study. The profile of the participating learners is shown in Table 1. The learners were assigned to either a control ($n = 20$) or an experimental ($n = 20$) group. Creswell (2014) suggested that, as a rough estimate, an educational researcher needs approximately 15 participants in each group in an experiment.

All learners had three years of EFL learning experience; that is, they were still beginning-level learners. The elementary school EFL curriculum emphasizes both reading and writing skills.

The research procedure used in this study is shown in Figure 1. Before the experiment, we assigned the participants to either the control or experimental group. After that, we carried out a pre-test, collected demographic information, and informed the participants of the details of the learning activity. EFL classes were conducted, and all the learners worked on a learning task after class. In the last class, we carried out a post-test, distributed the questionnaire, and conducted interviews. Two weeks after the last class, a delayed post-test was carried out. Details of all tests and statistical analyses, along with other relevant information, are provided later in this and following sections.

The language course covered several topics; however, we focused only on the "at the market" topic for the present study. The instructor, who was not one of the members of the research group, carried out three classes a week, each for an hour, for two weeks. In class, the instructor taught grammar, new English vocabulary, and sentence patterns related to shopping (e.g. fruits, vegetables, weighing, calculating the price, etc.) and then assigned the learners to practice their language skills by completing textbook exercises (e.g. a role-play dialogue). The instructor corrected learners when necessary. After class, the instructor assigned a learning task in which

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

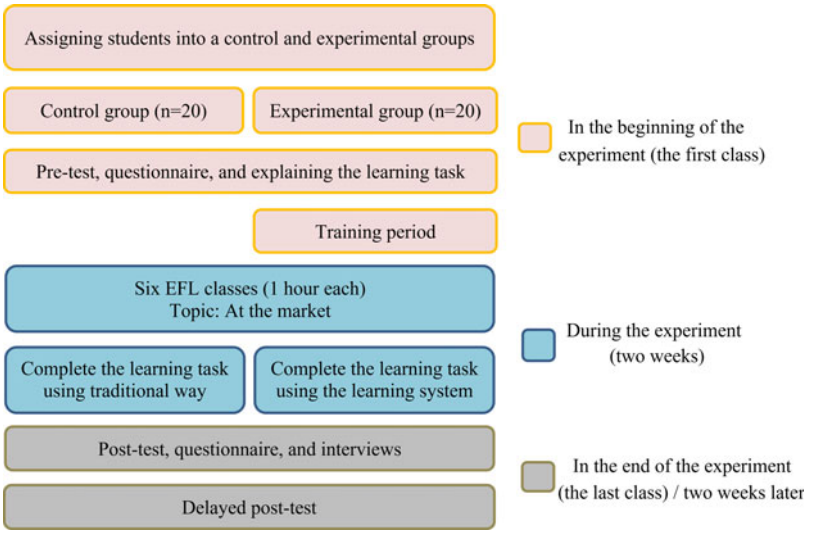


Figure 1. Research procedure

the participants were asked to apply the newly learned information to contexts with a realistic simulation of the real world. In the task, they were asked to describe their last shopping experience, such as where they shopped, what they bought, and how much they spent on each item.

The difference between the two groups was in the aid used to support task completion. The learners in the control group used the traditional method; that is, they learned vocabulary, viewed the corresponding pictures in their textbook, and described their shopping experience in their paper-based notebooks. The learners in the experimental group used the proposed learning system; that is, they took photos of target objects in a shop/market, used ITR technology to create labels for the captured objects, and described their shopping experience using the proposed learning system.

An Android-based mobile learning system was developed for this study and installed on tablet PCs (see Figure 2). The system included several main functions from which the following were used in this study: (1) Camera: the learners took photos of objects of interest (e.g. fruits or vegetables in the supermarket); (2) ITR: the learners took photos of objects and had their labels generated in English; (3) Notes: the learners created content in which they described their shopping experiences by typing and inserting photos that they had taken; (4) Textbook: the learners could read the learning material from the textbook. The difference between taking photos of objects when using a camera and ITR is that the former was used for created learning content, whereas the latter was used for generating an English label for the object of interest.

The Google Images service was employed for the ITR process. This service allows users to search the Web for image content (Yao *et al.*, 2017). The ITR process is based on a Search by Image feature, and it performs reverse image searches. That is, in contrast to traditional image retrieval by typing in keywords, this feature allows users to search by submitting a sample image as their query (Ohtaki, 2018). To use the Google Images service, a user has to click on the ITR icon (item b in Figure 2) in the learning system first. Then, a window opens that asks the user to take a photo of an object or upload a photo of an object from the tablet. After a user takes a photo of an object (e.g. watermelon), the English word and its meaning are shown in a new window.

Like many other recognition technologies (Shadiev, Hwang, Chen & Huang, 2014; Shadiev & Sun, in press), ITR also generates inaccurate output in some cases, especially when a user has limited experience using this technology. Thus, we arranged training on ITR sessions for

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

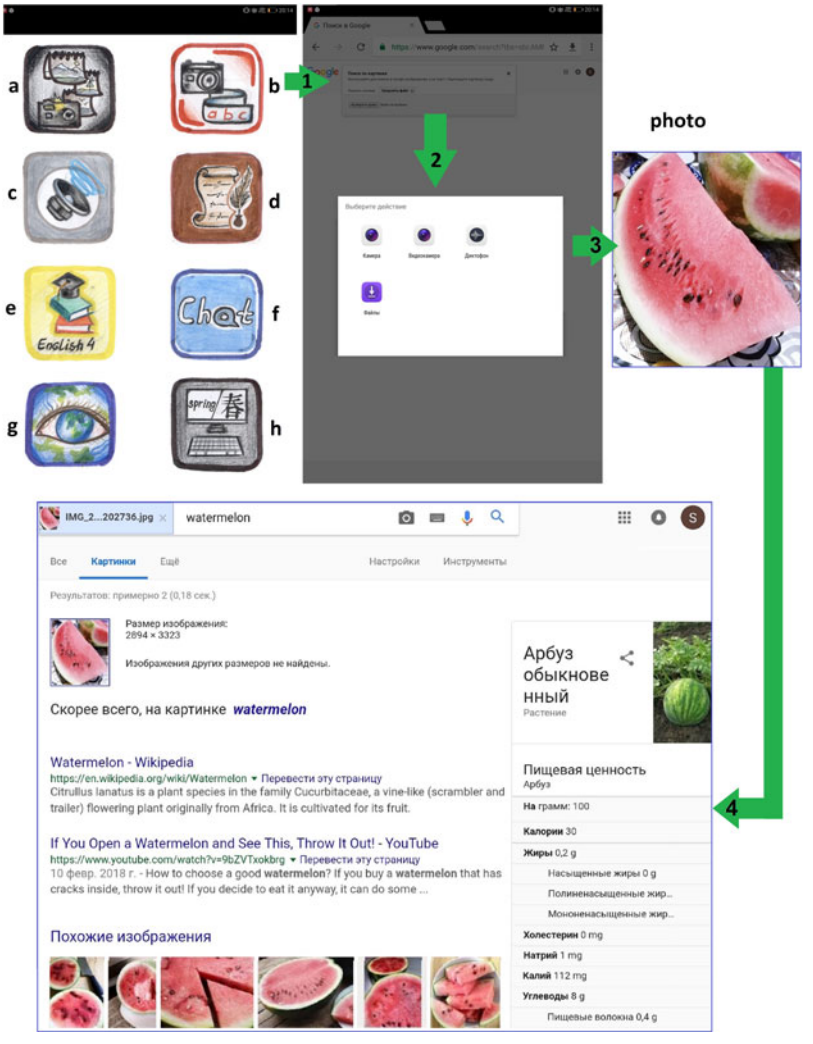


Figure 2. Screenshot of the system and image-to-text recognition process

the learners for a week. During the training, the learners used ITR to discover its strengths and limitations. At the beginning of the training, accuracy rates for the ITR process were very low (i.e. less than 70%); however, they became high (i.e. reached 95%) after one week.

During the training, the learners used several useful strategies to achieve higher ITR process accuracy rates. For example, the images submitted for query had to be clear so that the technology could easily analyze them and find appropriate identifiers such as colors, points, lines, textures, etc. A query image could not be ambiguous; that is, if a user wants to obtain the label for an apple, a query image should contain an apple only and not any other fruit, such as a cherry. Otherwise, the technology will not be able to recognize the apple and generate the correct label. The learners also experienced some issues when uploading pictures. Because cameras are very powerful nowadays, they capture high-resolution images, so it took some time for the learners to upload the photos they took to the ITR when a connection to the Internet was established through mobile communication. One strategy was to set a camera to take pictures at lower resolutions, and another was to trim the photos before uploading them. These two strategies helped upload the photos faster when the tablets were connected to the Internet using mobile communication.

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

To answer the first research question posed in this study related to the effectiveness of the proposed approach to facilitate vocabulary acquisition, a pre-test–post-test/delayed post-test design was used. A pre-test was created to measure the learners' prior vocabulary; a post-test was developed to assess their learning achievement, and a delayed post-test was designed to explore the retention of vocabulary knowledge. Each test contained three questions: (1) Select: Read the vocabulary words and select the correct pictures; (2) Match: Please match each English word with the correct Russian meaning; and (3) Write: Please write down the Russian meaning of each word. Examples of the test items are provided in Appendix A. Question 1 contained 10 sub-items for which the maximum score was 10. That is, each sub-item was scored "1" if it was answered correctly and "0" if it was answered incorrectly. Question 2 and Question 3 also contained 10 sub-items, and they were scored in the same way as Question 1. All test items were related to learning material covered in this study. They were not the same; all test items were similar in structure but different in content. In addition, the items were from different categories (i.e. nouns, verbs, and adjectives).

A questionnaire survey was developed in Russian to answer the second research question related to learner perceptions toward the system. The survey design was based on the general recommendations of earlier related research. Venkatesh and Davis (2000) explored factors affecting acceptance of technology, and Liaw (2008) was concerned about learning satisfaction during an intervention. The survey (see Table 3, section 4, Results) included four dimensions. The first three dimensions measured learner perceptions of the system: (1) perceived ease of system use (six items) – the degree to which a learner believed that using the system would be free of physical and mental effort; (2) perceived usefulness of the system during learning (six items) – the degree to which a learner believed that using the system for learning would enhance his or her learning performance; and (3) behavioral intention to use the system for learning in the future (three items) – a major determinant of whether a learner would feel like using the system again or not. The fourth dimension measured (4) learner perceived satisfaction (six items) – the degree to which a learner was satisfied with the system for learning purposes.

All items were almost the same as in Venkatesh and Davis (2000) and Liaw (2008) except we used simple words to make items understandable, given the age of the participating learners, and replaced some unrelated terms (e.g. "job" with "learning" or "e-learning" with "the system"). We received 20 valid answer sheets from 20 learners in the experimental group. The learners responded to the questionnaire items on a 5-point Likert scale, anchored by the end points "strongly disagree" (1) and "strongly agree" (5).

To answer the third research question, one-on-one semi-structured interviews in Russian were conducted by two researchers with the experimental learners. We aimed to explore learner experiences with using the system during the learning activity. The interview design followed the general recommendations of Creswell (2014). The interview protocol is provided in Appendix B. The interviews lasted for 20 minutes, where interviewees were asked open-ended questions such as "Please describe your learning experience with the system during the learning activity." and "Was the system useful for learning? If yes, please explain why."

An open coding analytic grounded theory process was employed for this study. The interview data analysis was carried out as follows: First, we audio-recorded all interviews with the participants' permission. We then fully transcribed all interviews for our analysis. Next, we highlighted and coded the text segments that met the criteria of providing the best research information – that is, those that showed the learner's learning experience with the system, affordances of the system, and its usefulness for learning. After that, codes with similar meanings were sorted into categories, and categories were established to form a framework by which to report findings. Two experienced researchers were involved in the coding process.

One may argue that the instrument's items (i.e. the questionnaire and interviews) were difficult for elementary school learners. To address this issue, we took the following precautions. First, we

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

Table 2. The results of the assessment and *t*-test

Test	Control group		Experimental group		*t*-test		Cohen's *d* (effect size)
	M	*SD*	*M*	*SD*	*t*	*p*	
Pre-test							
Question 1	5.20	2.31	5.15	2.25	0.069	0.945	
Question 2	4.65	2.56	4.60	2.39	0.064	0.949	
Question 3	3.75	1.68	3.40	1.90	0.616	0.541	
Post-test							
Question 1	7.40	2.14	9.15	0.93	−3.356	0.002	1.06
Question 2	6.65	1.78	8.40	1.14	−3.693	0.001	1.17
Question 3	5.15	1.72	7.50	1.43	−4.686	0.000	1.49
Delayed test							
Question 1	6.05	1.88	8.45	0.94	−5.108	0.000	1.61
Question 2	5.60	1.79	7.65	1.04	−4.431	0.000	1.40
Question 3	4.15	1.50	6.65	1.18	−5.863	0.000	1.85

considered the maturity level of our participants, so the original survey questions were modified to match the age of the participating learners as much as possible (e.g. we used simple words to make the items easily understandable by elementary school learners and replaced unrelated terms). Second, to add more confidence to the appropriateness of the instrument's items, an initial draft of the instrument was subjected to the scrutiny of the two elementary school teachers. Some items were modified based on their advice before the final version of the instrument was produced. Third, all items were in Russian (i.e. the native language of the participants). Fourth, before the learners started answering any item, the instructor ensured that they understood its meaning and rephrased or explained an item when necessary. Finally, we confirmed the validity of the learner responses to the questionnaire during the interviews.

4. Results

4.1 The effectiveness of the intervention on vocabulary acquisition

To answer the first research question, we compared the results of the pre-test, post-test, and delayed post-test given to the learners in the control group with those of the learners in the experimental group. To this end, the independent samples *t*-test and effect size were employed. A *t*-test determines whether there is a statistically significant difference between the means in two unrelated groups, and effect size identifies the practical strength of the conclusions related to group differences (Creswell, 2014). According to the results presented in Table 2, there were no significant between-group differences on the pre-test results: Question 1, $t = 0.069$, $p = 0.945$; Question 2, $t = 0.064$, $p = 0.949$; and Question 3, $t = 0.616$, $p = 0.541$. However, the results showed that the learners in the experimental group performed significantly better compared to those in the control group on the post-test: Question 1, $t = -3.356$, $p = 0.002$, $d = 1.06$; Question 2, $t = -3.693$, $p = 0.001$, $d = 1.17$; and Question 3, $t = -4.686$, $p = 0.000$, $d = 1.49$. In addition, the learners in the experimental group outperformed those in the control group on the delayed post-test: Question 1, $t = -5.108$, $p = 0.000$, $d = 1.61$; Question 2, $t = -4.431$, $p = 0.000$, $d = 1.40$; and Question 3,

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

Table 3. Perceptions and learning satisfaction survey results

#	Items	*M*	*SD*
Ease of use			
1	Learning to operate the system is easy for me	4.35	0.67
2	I find it easy to get the system to do what I want it to do	4.30	0.73
3	Interacting with the system does not require a lot of my mental effort	4.15	0.75
4	My interaction with the system is clear and understandable	4.30	0.86
5	It is easy for me to become skillful at using the system	4.30	0.92
6	Overall, I found the system easy to use	4.15	0.81
Overall		4.26	0.78
Usefulness for learning			
7	Using the system improves the quality of my learning	4.45	0.51
8	Using the system helps me to accomplish learning tasks more quickly	4.50	0.61
9	Using the system increases my productivity	4.60	0.50
10	Using the system enhances my learning effectiveness	4.45	0.60
11	Using the system improves my learning performance	4.60	0.50
12	Overall, I found using the system is useful in my learning	4.60	0.50
Overall		4.53	0.53
Behavioral intention to use the system			
13	I intend to continue using the system in future	4.35	0.49
14	I plan to use the system often	4.40	0.50
15	I will strongly recommend others to use the system	4.35	0.49
Overall		4.37	0.49
Perceived satisfaction			
16	I think using the system for learning is very interesting	4.20	0.41
17	I like to practice my EFL skills with the system in real-life situations	4.20	0.41
18	I feel free and relaxed during my learning with the system	4.20	0.41
19	I think I learned a lot by using the system	4.30	0.47
20	Using the system can increase my learning motivation	4.20	0.41
21	If I had to give a grade to my learning experience with the system, I would give it a high score	4.30	0.47
Overall		4.23	0.42

$t = -5.863$, $p = 0.000$, $d = 1.85$. Thirty-seven words from the textbook were included in the test questions. Twenty-five of them appeared in the experimental group's written homework task, and 16 appeared in the control group's homework. Vocabulary that appeared in either condition were common foods that students consume frequently (e.g. apple, eggs, bread, potatoes, water, etc.).

Our results suggest that using the learning system featuring ITR technology was beneficial to vocabulary acquisition. Thus, we conclude that the answer to the first research question is that

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

Table 4. Affordances of the system from interview data

Category	Code	Definition	Agreement
Technological support	Image-to-text recognition	The system recognizes an object and provides its name	100%
	Image-to-text translation	The system recognizes an object and provides its translation	100%
	Image-to-resources generation	The system recognizes an object and provides resources related to it (e.g. information about an object from Wikipedia online encyclopedia)	94%
Learning	Learn vocabulary	A learner can learn new vocabulary	100%
	Recall vocabulary	A learner can recall vocabulary	94%
	Spelling	A learner can learn spelling of vocabulary	90%
Authenticity	Authentic contexts	Learning takes place in authentic contexts	91%
	Authentic activities	Learning activities are authentic	94%

learners who study using ITR technology learn vocabulary better compared to those who study using traditional methods.

4.2 Perceptions of the learners

We attempted to answer the second research question by using the questionnaire survey. In order to assess the internal consistency of the questionnaire, we employed Cronbach's α. The value was 0.799; this demonstrated the reliability of the items to be acceptable. In addition, the Kaiser–Meyer–Olkin (KMO) index and Bartlett's sphericity test were performed to measure sampling adequacy. According to the results, the KMO value was 0.815, and the significance of Bartlett's sphericity was 0.000, indicating that the samples met the criteria for the analysis.

The results of the questionnaire survey are presented in Table 3. According to the results, most learners in the experimental group expressed high agreement toward ease of system use ($M = 4.26$, $SD = 0.78$). The results also showed that most learners positively perceived the usefulness of the system for learning ($M = 4.53$, $SD = 0.53$). In addition, we found that most learners had high behavioral intention to use the system for learning in the future ($M = 4.37$, $SD = 0.49$). Finally, our results showed that most learners had high levels of learning satisfaction ($M = 4.23$, $SD = 0.42$). These results suggest that most of the learners under consideration had positive perceptions of the learning system. Therefore, it could be concluded that experimental learners positively perceived their learning experiences using ITR technology.

4.3 Affordances of the learning system

To answer the third research question related to the affordances of the system, we carried out interviews with the experimental learners. Table 4 includes the categories, codes, definitions, and inter-rater agreement from the interview data analysis. The inter-rater reliability of the interview data was evaluated using Cohen's kappa; the result ($k = 0.957$) indicated high inter-rater reliability. For example, the highest inter-rater agreement (100%) was achieved for the "image-to-text recognition," "image-to-text translation," and "learn vocabulary" codes. The lowest inter-rater agreement (90%) was achieved for the "spelling" code. The percentages of appearance of the codes identified by the two coders shown in Table 4 allow observation of the weight of each item.

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

The results of the interview data analysis revealed three main categories of affordances: (1) technological support, (2) learning, and (3) authenticity. According to the learners, the system can recognize an object and provide its name (Code: Image-to-text recognition) as well as recognize an object and provide its translation (Code: Image-to-text translation). Furthermore, the learners said that the system can recognize an object and provide them with resources related to this object (e.g. information about the object from the Wikipedia online encyclopedia) (Code: Image-to-resources generation). In terms of the learning affordances of the system, the learners mentioned that they could learn new vocabulary (Code: Learn vocabulary) as well as recall vocabulary (Code: Recall vocabulary).

In addition, the learners said that they could learn the spelling of words (Code: Spelling). Finally, the learners recognized that the system had authenticity affordances. For example, the learners could learn vocabulary in authentic contexts using the system (Code: Authentic contexts), and their learning activities were authentic (Code: Authentic activities) because they were related to real life. Based on our results, we could conclude that the affordances of ITR technology for vocabulary acquisition are as follows: (1) image-to-text recognition, image-to-text translation, and image-to-resources generation in the technological support category; (2) learn vocabulary, recall vocabulary, and learn spelling in the learning category; and (3) experience authentic contexts and engage in authentic activities in the authenticity category.

Below are three extracts from the interviews with the learners that show the system was useful for language learning.

> Student 1
>
> We usually learn from our textbooks in English class. Then, after class we have to complete a textbook assignment. Such classwork and homework very rarely let me use my skills and new knowledge in my daily life . . . Usually my mom is busy with house chores, so she asks me to buy groceries. I feel that this experiment was very useful because I could practice my skills and new knowledge that I learned in English class using new technology when I shopped.
>
> Student 2
>
> Learning English with the system was interesting. I prefer to make up my own content related to what I learned in class when I am in the shop because the textbook vocabulary, pictures, and translations are not real, but they are in the shop.
>
> Student 2
>
> I like to learn with this system because it can show labels of different fruits and vegetables that I take photos of. I can take photos of any objects that I am interested in. Also, I have no limitation on vocabulary to learn.

5. Discussion

Learning activities in contexts with a realistic simulation of the real world were designed in this study to achieve relevant, meaningful, and effective learning (Herrington & Herrington, 2006). We employed ITR technology to assist learner vocabulary acquisition (Ohtaki, 2018; Yao *et al.*, 2017). Our results showed no significant between-group differences in the learners' prior vocabulary. However, the learners in the experimental group (i.e. those who used the proposed learning system during the learning activity) outperformed the learners in the control group (i.e. those who used the traditional aid during the learning activity) on the post-test and delayed test.

These results suggest that EFL vocabulary acquisition was more effective when learners used the system compared to the traditional approach, which is in line with related studies

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

(Kurt & Bensen, 2017; Lin & Yu, 2017; Wu, 2018). One reason for this outcome is that the experimental learners learned vocabulary more actively compared to their counterparts. This type of learning is in contrast to rote learning (Pan, 2017). The advantages of learning a language actively has been emphasized (Blaz, 2018). For example, scholars have claimed that learning and retention takes place through active language input (e.g. acquiring new information) and output (e.g. speaking and writing about related and meaningful topics) (Krashen, 1989; Swain, 1985). In addition, active learning added value to vocabulary acquisition (Tanaka, 2017) because learning with objects of interest was very meaningful and relevant to our learners (Kurt & Bensen, 2017). In this study, the experimental learners located the objects of interest in authentic contexts, used the ITR tool to generate the names, translations, spelling, and related resources from the Internet, and then applied the information to contexts with a realistic simulation of the real world by creating their own learning content (Kukulska-Hulme & Viberg, 2018). This learning activity ensured active learning.

Another reason for the outcome is the fact that the system could recognize any targeted object. For this reason, the experimental learners had no limits related to learning the topic vocabulary, whereas the control learners were limited to vocabulary provided in the textbook. This finding suggests that learners can always use different available resources when they are outside school context, and mobile technology makes it possible to learn languages in all kinds of informal contexts. This notion is in line with seamless learning (Kuh, 1996; Wong, 2013). That is, the in-class and out-of-class learning experiences were linked; the learners learned in class and then applied newly learned knowledge to contexts with a realistic simulation of the real world. Scholars have argued that with more available learning resources, vocabulary can be increased (Peker *et al.*, 2018). It has also been suggested that learning words of interest is useful for learning motivation and satisfaction (Ansari & Sabouri, 2016).

Furthermore, our results suggest that using the learning system was beneficial for the experimental learners to transfer the vocabulary items they learned into their long-term memory. The learners who used the learning system learned actively, focused on vocabulary of interest, and had no limit in terms of selecting vocabulary. Such circumstances encouraged the experimental learners to review and study the vocabulary items on a more regular basis. Similar findings have been reported elsewhere (Alemi *et al.*, 2012; Kim, 2011). For example, Kim (2011) suggested that learners learn more vocabulary and retain it better if they are highly involved in vocabulary acquisition cognitive processes. Thus, the answer to the first research question is that the learners who studied using ITR technology learned vocabulary more easily as compared to their counterparts.

In this study, we developed a mobile learning system. According to Hwang *et al.* (2016), it is essential to examine learner perceptions of the learning system and to objectively understand their degree of acceptance of the system. Venkatesh and Davis (2000) suggested that learner acceptance of a system can be measured through perceived ease of use and perceived usefulness. Venkatesh and Davis (2000) further claimed that if learners feel that the system is easy to use and useful, they will continue using it in the future. The results showed that most learners accepted the learning system and were satisfied with using it, as their perceptions corresponded highly with positive views of the system. Ensuring acceptance of the system is very important, especially when a system is based on signal processing such as ITR technology. The acceptance of such systems is mostly due to its recognition process accuracy rates (Shadiev *et al.*, 2014; Shadiev & Sun, in press). It was very low in the beginning of this study but reached higher rates after one week of training on the learning system. One reason for this achievement was revealed in the interviews with the learners. They mentioned that several strategies applied during the ITR process ensured higher ITR rates of the system. As a result, learners found the system easy to use.

In terms of the usefulness of the learning system for learning, the learners mentioned that learning vocabulary by taking pictures of objects and getting their labels from ITR was useful. Because they could take pictures and learn vocabulary related to objects of their interest, the

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

learning process was very useful, meaningful, and relevant to their learning. This is why the learners had positive perceptions of the learning system in terms of its usefulness for learning. Because the learners perceived the system to be easy to use and useful for learning, their behavioral intentions to use it for learning in the future were also high.

Our findings can be supported by those from earlier related studies. For example, scholars have shown that using pictures with corresponding labels is beneficial for vocabulary learning because learners can more easily learn and remember vocabulary (Lin & Yu, 2017; Peker *et al*., 2018). Scholars have also emphasized the importance of making context and learning inseparable (Herrington & Herrington, 2006; Kiernan & Aizawa, 2004; Kukulska-Hulme & Viberg, 2018). Extending the learning process to contexts with a realistic simulation of the real world made it effective and meaningful. Finally, when technology is being used for vocabulary acquisition, the learning process can be extended from the classroom to contexts with a realistic simulation of the real world and thus facilitated greatly (Comas-Quinn *et al*., 2009; Kurt & Bensen, 2017; Lin & Yu, 2017; Ros i Solé *et al*., 2010). For example, in our study, learners learned both in class and outside of it. They used ITR to take pictures of objects and generated their labels to learn new vocabulary. Similarly, the learners in other related studies also positively perceived their learning with mobile learning systems (Kurt & Bensen, 2017; Lin & Yu, 2017; Wu, 2018). For example, the system was reported to be useful to recall the meanings of new words and improve learning satisfaction (Lin & Yu, 2017). However, our research was different from other studies in that we employed ITR technology for vocabulary learning, whereas other scholars have used quite different tools. The answer to the second research question is that the experimental learners positively perceived their learning experiences using ITR technology.

Several affordances of the learning system to support vocabulary acquisition were found in this study. Our answer to the third research question was that the affordances of ITR technology for vocabulary acquisition include (1) image-to-text recognition, image-to-text translation, and image-to-resources generation in the technological support category; (2) learning vocabulary, recalling vocabulary, and improving spelling in the learning category; and (3) learning in authentic contexts and engaging in authentic activities in the authenticity category. Affordances from the first category are related to the functions of the learning system (e.g. recognition of an object and providing its name); affordances from the second category are related to vocabulary learning processes (e.g. learning or recalling vocabulary); and affordances from the last category are related to learning contexts (e.g. authentic contexts or activities). It is difficult to say now which are the most important of these affordances and why this would be the case because we did not focus on this aspect in this study. Therefore, this may be a promising research direction in the future. For example, learner behavior to use the system functions (e.g. image-to-text translation or image-to-resources generation) or to learn using the system (e.g. learning or recalling vocabulary) can be tracked and recorded, and then correlations between the recorded data and learning achievement can be tested.

In the learning task, the learners had to explain where they shopped, what they bought, and how much they spent for each item using the target language. The experimental learners engaged in image-driven lookup of vocabulary as opposed to the translation/dictionary lookup method used by the control group. The experimental learners had an advantage over their counterparts because of the learning system. For example, if there was an object that the learners had to describe, the experimental learners could easily get its name in English by using the system. Furthermore, the system could also provide a translation of the object as well as additional resources related to it, such as collocations and examples of using it in a sentence. As a result, vocabulary acquisition – that is, learning, recall, and spelling – was facilitated.

The lack of research on usage of ITR in the area of language learning suggests that the possibilities of this technology are not being fully harnessed for education. Scholars have argued that it is possible that educators and researchers simply do not fully understand the potential and

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

affordances of the technology. For this reason, they are not able to make appropriate or innovative use of it (Mishra & Koehler, 2006). Therefore, the findings of this study can be useful to inspire educators and researchers and guide them in designing learning activities supported by ITR technology.

Other promising research directions are as follows. First, an exploration of how ITR supports the learning and remembering of words in a different way on the basis of how words are presented to learners and then learned (see, for example, Lin & Yu, 2017); for example, combinations of words, pictures, and translations as in the different question types in the tests used in the present study: (Question 1) word in English plus picture, (Question 2) word in English matched with word in Russian, and (Question 3) word in English and its translation in Russian. Second, an investigation into how parental involvement (e.g. the amount) or time spent on the task can influence learning outcomes would be warranted.

This novel approach was employed based on a shopping topic in which the learners used images of fruits and vegetables for the ITR process. The approach should be applied to other topics with caution because the ITR accuracy rate may decrease when using images of other objects as queries. In addition, learners have to be trained in using the technology for the ITR process. This will ensure that they understand the strengths and limitations of the technology and then can fully utilize it during vocabulary acquisition. Furthermore, educators and researchers should teach their learners useful strategies for achieving better accuracy rates when using ITR. The following are some strategies used in our study. Strategy 1: Use clear images for queries so that the technology can easily analyze them and find appropriate identifiers such as colors, points, lines, and textures. Strategy 2: Use unambiguous images for queries (i.e. an image with a single object on it) so that the technology will be able to recognize the object and generate the correct label; for example, if a user wants to obtain the label for an apple, a query image should contain an apple only and not any other fruit. Strategy 3: Set a camera to take pictures at lower resolutions when using low Internet bandwidth so that the ITR can process them faster. Strategy 4: Trim the photos before uploading them to make them unambiguous, with lower resolutions; in this case, the ITR can recognize them and generate labels more efficiently and more quickly.

The methodological use of the task was short (i.e. two weeks) because of the realities of classroom-based research. Therefore, longer studies need to be carried out to add more rigor to the validity of our methodological approach. In addition, the study could be replicated with a larger sample. We also suggest that the system can be applied to facilitate vocabulary acquisition related to different language writing systems. Furthermore, this approach may be useful in other domains in which learners have to learn about objects by distinguishing their shape, size, and surfaces, after which they must apply newly learned knowledge to the real world. For example, in a content-based language class, learners learning about plants in French for a biology course may use the system to take photos of plants of interest in authentic contexts to obtain their labels and additional learning content.

Next, we highlight how the present study differs from earlier research. First, learning activities were instructor centered in earlier studies (Lin & Yu, 2017; Wu, 2018) but were learner centered in ours. Second, related studies focused on language input only (Kurt & Bensen, 2017; Lin & Yu, 2017), whereas we considered both language input and language output. Finally, not much attention was paid to the use of ITR technology in earlier studies (Kurt & Bensen, 2017; Lin & Yu, 2017; Wong, 2013; Wu, 2018).

6. Conclusion

The following are our answers to the three research questions. First, the learners who studied using ITR technology learned vocabulary more easily compared to those who studied using the traditional approach. Second, the experimental learners positively perceived their learning experiences using ITR technology. Third, the affordances of ITR technology for vocabulary acquisition include

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

(1) image-to-text recognition, image-to-text translation, and image-to-resources generation in the technological support category; (2) learning vocabulary, recalling vocabulary, and improving spelling in the learning category; and (3) learning in authentic contexts and engaging in authentic activities in the authenticity category.

Based on our results, we suggest designing similar learning activities and employing the learning system featuring ITR technology to facilitate vocabulary acquisition. Such learning activities can help students learn new information and then apply it to the real world. With the support of the learning system, vocabulary acquisition can become more effective, relevant, and meaningful compared to the traditional approach, because learners can learn more actively and use unlimited learning resources of interest from surrounding contexts in the real world.

Ethical statement. (a) The data set will be provided on request after we finish this project; (b) the study was performed following institutional ethical guidelines; and (c) there is no potential conflict of interest in this work.

References

Abdous, M. H., Camarena, M. M. & Facer, B. R. (2009) MALL technology: Use of academic podcasting in the foreign language classroom. *ReCALL*, 21(1): 76–95. https://doi.org/10.1017/S0958344009000020

Alemi, M., Sarab, M. R. A. & Lari, Z. (2012) Successful learning of Academic Word List via MALL: Mobile assisted language learning. *International Education Studies*, 5(6): 99–109. https://doi.org/10.5539/ies.v5n6p99

Ansari, M., & Sabouri, N. B. (2016) The relationship between the use of vocabulary learning strategies and gender of Iranian EFL learners. *Research in English Language Pedagogy*, 4(1): 88–100.

Blaz, D. (2018) *The world language teacher's guide to active learning: Strategies and activities for increasing student engagement* (2nd ed.). New York: Routledge. https://doi.org/10.4324/9781315169514

Collins, A. (1988) *Cognitive apprenticeship and instructional technology* (Technical Report No. 6899). Cambridge, MA: BBN Laboratories Inc. https://doi.org/10.21236/ADA203609

Comas-Quinn, A., Mardomingo, R. & Valentine, C. (2009) Mobile blogs in language learning: Making the most of informal and situated learning opportunities. *ReCALL*, 21(1): 96–112. https://doi.org/10.1017/S0958344009000032

Creswell, J. W. (2014) *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston: Pearson Education.

Duman, G., Orhon, G. & Gedik, N. (2015) Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, 27(2): 197–216. https://doi.org/10.1017/S0958344014000287

Harmer, J. (2007) *The practice of English language teaching* (4th ed.). London: Longman.

Herrington, A. & Herrington, J. (2006) What is an authentic learning environment? In Herrington, A. & Herrington, J. (eds.), *Authentic learning environments in higher education*. Hershey: Information Science Publishing, 1–14. https://doi.org/10.4018/978-1-59140-594-8.ch001

Hwang, W.-Y., Shadiev, R., Hsu, J.-L., Huang, Y.-M., Hsu, G.-L. & Lin, Y.-C. (2016) Effects of storytelling to facilitate EFL speaking using Web-based multimedia system. *Computer Assisted Language Learning*, 29(2): 215–241. https://doi.org/10.1080/09588221.2014.927367

Kiernan, P. J. & Aizawa, K. (2004) Cell phones in task based learning: Are cell phones useful language learning tools? *ReCALL*, 16(1): 71–84. https://doi.org/10.1017/S0958344004000618

Kim, Y. (2011) The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 61(s1): 100–140. https://doi.org/10.1111/j.1467-9922.2011.00644.x

Krashen, S. (1989) We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73(4): 440–464. https://doi.org/10.1111/j.1540-4781.1989.tb05325.x

Kuh, G. D. (1996) Guiding principles for creating seamless learning environments for undergraduates. *Journal of College Student Development*, 37(2): 135–148.

Kukulska-Hulme, A. & Viberg, O. (2018) Mobile collaborative language learning: State of the art. *British Journal of Educational Technology*, 49(2): 207–218. https://doi.org/10.1111/bjet.12580

Kurt, M. & Bensen, H. (2017) Six seconds to visualize the word: Improving EFL learners' vocabulary through VVVs. *Journal of Computer Assisted Learning*, 33(4): 334–346. https://doi.org/10.1111/jcal.12182

Liaw, S.-S. (2008) Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education*, 51(2): 864–873. https://doi.org/10.1016/j.compedu.2007.09.005

Lin, C.-C. & Yu, Y.-C. (2017) Effects of presentation modes on mobile-assisted vocabulary learning and cognitive load. *Interactive Learning Environments*, 25(4): 528–542. https://doi.org/10.1080/10494820.2016.1155160

Mishra, P. & Koehler, M. J. (2006) Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6): 1017–1054.

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

Nah, K. C., White, P. & Sussex, R. (2008) The potential of using a mobile phone to access the Internet for learning EFL listening skills within a Korean context. *ReCALL*, 20(3): 331–347. https://doi.org/10.1017/S0958344008000633
Nation, I. S. P. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524759
Nation, P. (2007) The four strands. *Innovation in Language Learning and Teaching*, 1(1): 2–13. https://doi.org/10.2167/illt039.0
Nation, P. (2013) *What should every EFL teacher know?* Tokyo: Compass Publishing.
Ohtaki, Y. (2018) An image source checker for educational presentation materials. In Barolli, L., Kryvinska, N., Enokido, T. & Takizawa M. (eds.), *Advances in network-based information systems: The 21st International Conference on Network-Based Information Systems (NBiS-2018)*. Springer, Cham, 903–912. https://doi.org/10.1007/978-3-319-98530-5_81
Paivio, A. (1986) *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
Pan, W. F.. (2017) The Effects of using the Kinect motion-sensing interactive system to enhance English learning for elementary students. *Educational Technology & Society*, 20(2): 188–200.
Peker, H., Regalla, M. & Cox, T. D. (2018) Teaching and learning vocabulary in context: Examining engagement in three prekindergarten French classrooms. *Foreign Language Annals*, 51(2): 472–483. https://doi.org/10.1111/flan.12338
Ros i Solé, C., Calic, J. & Neijmann, D. (2010) A social and self-reflective approach to MALL. *ReCALL*, 22(1): 39–52. https://doi.org/10.1017/S0958344009990188
Shadiev, R., Hwang, W.-Y., Chen, N.-S. & Huang, Y.-M. (2014) Review of speech-to-text recognition technology for enhancing learning. *Educational Technology & Society*, 17(4): 65–84.
Shadiev, R., Hwang, W.-Y. & Huang, Y.-M. (2017) Review of research on mobile language learning in authentic environments. *Computer Assisted Language Learning*, 30(3–4): 284–303. https://doi.org/10.1080/09588221.2017.1308383
Shadiev, R., Hwang, W.-Y. & Liu, T.-Y. (2018) Investigating the effectiveness of a learning activity supported by a mobile multimedia learning system to enhance autonomous EFL learning in authentic contexts. *Educational Technology Research and Development*, 66(4): 893–912. https://doi.org/10.1007/s11423-018-9590-1
Shadiev, R., Liu, T.-Y. & Hwang, W.-Y. (in press) Review of research on mobile-assisted language learning in familiar, authentic environments. *British Journal of Educational Technology*. Advance online publication. https://doi.org/10.1111/bjet.12839
Shadiev, R. & Sun, A. (in press) Using texts generated by STR and CAT to facilitate student comprehension of lecture content in a foreign language. *Journal of Computing in Higher Education*. Advance online publication. https://doi.org/10.1007/s12528-019-09246-7
Shadiev, R. & Yang, M. (2020) Review of studies on technology-enhanced language learning and teaching. *Sustainability*, 12(2): 1–22. https://doi.org/10.3390/su12020524
Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass, S. M. & Madden, C. G. (eds.), *Input in second language acquisition*. Cambridge, MA: Newbury House Publishers, 235–256.
Tanaka, M. (2017) Examining EFL vocabulary learning motivation in a demotivating learning environment. *System*, 65: 130–138. https://doi.org/10.1016/j.system.2017.01.010
Venkatesh, V. & Davis, F. D. (2000) A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2): 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926
Wong, L.-H. (2013) Analysis of students' after-school mobile-assisted artifact creation processes in a seamless language learning environment. *Journal of Educational Technology & Society*, 16(2): 198–211.
Wu, T.-T. (2018) Improving the effectiveness of English vocabulary review by integrating ARCS with mobile game-based learning. *Journal of Computer Assisted Learning*, 34(3): 315–323. https://doi.org/10.1111/jcal.12244
Yao, Y., Zhang, J., Shen, F., Hua, X., Xu, J. & Tang, Z. (2017) A new web-supervised method for image dataset constructions. *Neurocomputing*, 236: 23–31. https://doi.org/10.1016/j.neucom.2016.07.066

Appendix A. Examples of the pre-test items

Question 1. Select: Read the vocabulary words and select the correct pictures (10 items/max score 10)

Vocabulary words: (1) Eggs, (2) Peaches, (3) Potatoes



_____

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press



_____



_____

Question 2. Match: Please match each English word with the correct Russian meaning (10 items/max score 10)

1. Cabbage (_____)Баклажан
2. Eggplant (_____)Хлеб
3. Bread (_____)Капуста

Question 3. Write: Please write down the Russian meaning of each word (10 items/max score 10)

1. Delicious _____
2. Strawberry _____
3. Sausage _____

Appendix B. Interview protocol

Time of interview:
Date:
Place:
Interviewer(s):
Interviewee:

Introduction to the interview:

- Mention that the interview is being recorded.
- Mention how long the interview will take.
- Mention that there is no right or wrong answer.
- Mention that the interview is anonymous and that he/she can talk as freely as possible.
- The aim of this research is, as stated, to find out his/her learning experiences with using the system during the activity and affordances of the system for vocabulary learning.

Interview questions:

- Please describe your learning experience with the system during the learning activity.
- Was the system useful for learning? If yes, please explain why.

Additional questions to prompt his/her answer:

- Please tell us whether you used the system during the learning activity or not.
- If you used the system, can you please tell us what functions of the system you used.

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press

– Please tell us for what purpose did you use certain functions of the system.
– Please tell us if you think the system was/wasn't useful for your vocabulary learning.
– If it was/wasn't useful, can you explain why?
– Please tell us where you used this system.

Additional questions to elicit more information:

– Tell us more.
– Can you explain your response more?
– I need more detail.

Closing comment

Thank him/her for co-operation and participation in this interview. Assure him/her of the confidentiality of the responses and the potential for future interviews.

*All correspondence regarding this publication should be addressed to Dr Huang (Email: huang@mail.ncku.edu.tw).

About the authors

Rustam Shadiev is a professor at Nanjing Normal University, China. He is also a distinguished professor of Jiangsu province, China. His research interests include ICT for enhancing language learning and cross-cultural understanding.

Ting-Ting Wu is an associate professor at the National Yunlin University of Science and Technology, Taiwan. Her research focus covers such areas as mobile and ubiquitous learning, information technology-applied instructions, and intelligent learning systems.

Yueh-Min Huang is a chair professor at the National Cheng Kung University, Taiwan. His research interests focus on e-learning, multimedia communications, and artificial intelligence.

Author ORCiD. Rustam Shadiev, https://orcid.org/0000-0001-5571-1158
Author ORCiD. Ting-Ting Wu, https://orcid.org/0000-0003-4970-7042
Author ORCiD. Yueh-Min Huang, https://orcid.org/0000-0001-7052-1272

https://doi.org/10.1017/S0958344020000038 Published online by Cambridge University Press