Liu, P.-L., & Chen, C.-J. (2023). Using an AI-Based Object Detection Translation Application for English Vocabulary Learning. *Educational Technology & Society*, *26*(3), 5-20. https://doi.org/10.30191/ETS.202307_26(3).0002

Using an AI-Based Object Detection Translation Application for English Vocabulary Learning

Pei-Lin Liu[1] and Chiu-Jung Chen[2*]
[1]National Chia-Yi University, Taiwan // [2]National Chia-Yi University, Taiwan // peilin@mail.ncyu.edu.tw // chenc@mail.ncyu.edu.tw
[*]Corresponding author

ABSTRACT: This study aimed to examine the effects of an AI-based object detection translation (AI-based ODT) application (app) on EFL students' vocabulary learning. We developed a system that utilized strategies to facilitate learners' vocabulary learning. The app applied dual code theory to present the objects in picture, word, and pronunciation formats. Seventy-two elementary school students were divided into lower-ability and higher-ability groups according to their English proficiency, and were then randomly assigned to the control and experimental conditions based on their ability. The learners' learning performance in the control and experimental conditions was compared using a pre-test–post-test design. Through two-way ANOVA analysis, we observed that in the experimental group the higher-ability students benefited more from the AI-based ODT app technology than did the lower-ability students. This significant difference could be taken as evidence of the positive effect of the AI-based ODT app technology, particularly for higher-ability students.

Keywords: Artificial intelligence, Dual code theory, English as a foreign language (EFL), Object detection translation, Vocabulary learning

1. Introduction

1.1. Background and problems of the study

Learning foreign languages can be a challenging but rewarding process for young children. Learning a language in countries where that language is not generally spoken can be difficult. Among the language learning skills, the use of a foreign language primarily relies on the use of vocabulary as the building blocks of a language. Without vocabulary, people cannot communicate or continue to develop their language acquisition (Ramos & Dario, 2015).

As a result, improving vocabulary learning is essential and a priority in the language learning process in order to learn a language well (Tanaka, 2017). Creating an impactful educational setting is crucial for increasing young children's motivation for learning (Weiland & Yoshikawa, 2013). Traditionally, English is taught to children using songs, textbook exercises, nursery rhymes, and storybook reading, and a majority of learners apply the rote learning strategy for vocabulary memorization based on repetition (Nation, 2013). However, there are some common difficulties associated with the rote learning of vocabulary, including the learners easily forgetting words and having difficulty recalling them (Chuo & Yen, 2014; Wu & Huang, 2017).

In order to overcome the limitations of rote learning, the provision of effective learning strategies and tools to help learners improve their memorization of vocabulary is important. Today's children are raised in a technologically advanced society, which has influenced how they learn in comparison to previous generations. Lessons which incorporate technology can help to shape the teaching and learning process to be more inventive. It is critical to investigate which technologies can be used in the classroom to engage young non-native-speaking pupils in the learning materials and to encourage them to practice English in order to enhance their foreign language learning skills (Dalim et al., 2020).

1.2. Purpose and scope of this study

This study explored the potential of combining an artificial intelligence (AI) based object detection translation (ODT) application (app) and dual code theory design which presented the objects in picture, text, and pronunciation formats. The AI-based ODT app was adopted as an innovative way to teach basic English vocabulary to young non-native-speaking learners. AI is the general term for the science of artificial intelligence. It uses computers to simulate human intelligent behaviors and it trains computers to learn human behaviors such as learning, judgment, and decision-making (Yigitcanlar et al., 2020; Zhang & Lu, 2021). Object detection is a subset of AI; it uses deep learning to provide a fast and accurate means to predict the location of an object in an image. Deep learning is a powerful machine learning technique that uses multi-layered AI networks in which the object detector automatically learns image features required for detection accuracy in tasks. For example, Google Translate (see Figure 1) allows learners to explore the world in the language that they are familiar with by just pointing the camera lens at the foreign text (Gu, 2019; Liu, 2018). The function of Google Translate is to use AI to train computers through machine learning and deep learning technologies in order to predict the most likely words, thus providing stronger and more accurate translations (Liu, 2018).

Figure 1. The Google Translate app (Fu, 2016)

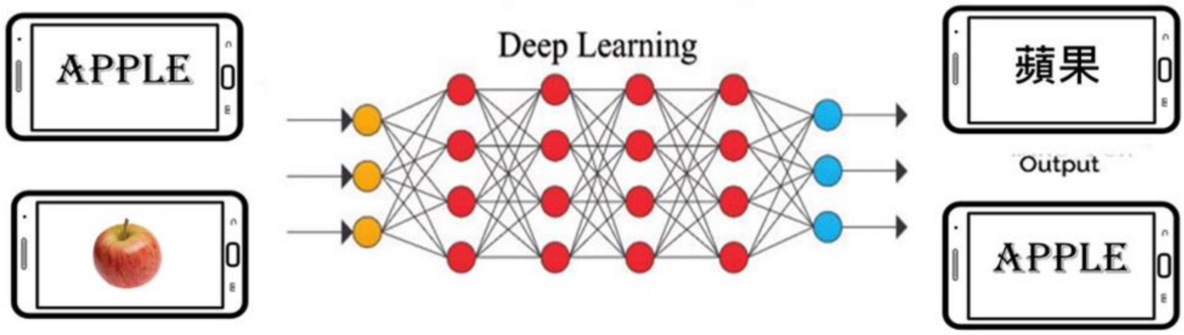
The Google Translate app has been updated to include 60 new languages. When a camera image is sent, the app must first locate the letters in the image. It must filter out the background items and focus on the words that it can identify. Second, the app must be able to identify each letter. Deep learning comes into play here. The third step is to look up the known letters in a dictionary to check what their particular combination means (see Figure 2) (Good, 2015; Vincent, 2019).

Figure 2. Step-by-step process of how the Google Translate app works (Good, 2015)

Although the Google Translate app launched the Instant Camera Translation function for mobile phones in 2014, thus far, there have been few empirical studies on its application in English learning. Most of the related studies have focused on the technical introduction and calculations to improve image recognition (e.g., Chatterjee & Bhattacharjee, 2020; de Carvalho et al., 2018; Giovany et al., 2017; Kong et al., 2015; Mezgec et al., 2019; Xie et al., 2020) and text recognition (Lee et al., 2017; Yousef et al., 2020; Zhong et al., 2018). Both Giovany's et al. (2017) and Kong's et al. (2015) studies concentrated on how to improve the recognition of food images, with a focus on how to remove unnecessary information due to the cluttered background on the menu, and how to improve recognition of the menu text.

In terms of the current computer technology, recognizing "text" is simple for a computer. Because of the fixed writing method of text or numbers, the computer can analyze the characters using a database. However, it is relatively difficult to detect "objects." An object involves different compositions and details, the identification of which requires more complicated calculations by the computer. This is also one of the reasons why Google Translate can only support text recognition and translation functions for objects, but is currently unable to support an object detection translation function (see Figure 3).

Figure 3. Differences between the text-to-text translation app and the object detection translation app



In the present study, we developed a learning system featuring the AI-based ODT app (see Figure 4) based on the dual-coding theory (DCT). The dual-coding theory assumes that information is encoded using both visual and verbal forms. In the human mind, both visual and verbal information are processed separately via different channels. By integrating both visual and verbal channels simultaneously, it will be easier for learners to recall the particular memory in the future (Clark & Paivio, 1991; Kanellopoulou et al., 2019; Kassim, 2018). In the study of English vocabulary, it has been proven that the dual-coding theory can provide learners with the ability to organize and classify words through associations and present the relevance of words with specific pictures, not only allowing learners to memorize these words quickly but also helping them to remember the words for a long time (Liu, 2016; Liu et al., 2020).

Figure 4. Schematic representation of the ODT app developed by the researchers

1.3. Benefits of using the ODT app with AI, and the research question

The combination of the AI-based ODT app and the DCT design in this study enabled the presentation of objects in picture, text, and pronunciation formats. Learners took photographs of objects, and the AI-based ODT app recognized and translated the authentic objects into English/Chinese words with their relevant pronunciation. There are two main benefits of using ODT with the application of AI. First, it is expected that well-validated AI tools can be used to improve the accuracy and reproducibility of vocabulary learning. Moreover, AI systems reproducibly yield the same results when provided with similar inputs, do not suffer from fatigue, and excel at finding patterns in large amounts of data. Second, AI tools can augment the efficiency of object detection for language learning (Steiner et al., 2021).

Driven by the dual needs for improved accuracy and efficiency, as well as by rapid improvements in technology, there are now a growing number of research articles describing promising applications of AI across a wide variety of tasks. However, despite these fundamental advances, there are exceptionally few practical demonstrations of the integration of AI into language learning.

The aim of this study was therefore to explore the potential of an AI-based ODT app for teaching the English names of fruit and vegetables to children who do not speak English as their first language. This study aimed to answer the following research question: Is there a significant difference in the learning gains of children with different abilities due to the teaching platforms in which they experienced second language vocabulary learning?

1.4. Contributions of this study

The main contributions of the study include insights into how object recognition translation affects children's interaction with an AI-based application, and influence their learning experience. The more that is known about the effect of taking pictures and AI on young non-native children's English language learning, the more effective learning approaches that can be developed.

2. Literature review

2.1. Technology-enhanced language learning

Learning and teaching processes have entirely changed as a result of the availability of new interactive computer technologies (Burston, 2014; Eslami & Ahmadi, 2019), leading to the development of several fields of study such as Computer-Assisted Language Learning, Mobile-Assisted Language Learning, and Computer-Mediated Communication. These distinct but overlapping fields of study share a focus on using technology as an assistive tool or mediator to enhance the teaching and learning of a second or foreign language (L2) (Chun, 2016). In most cases, technology has been used to assist students in accomplishing a variety of language learning objectives such as listening comprehension (Ramírez Verdugo & Alonso Belmonte, 2007), reading comprehension (Dreyer & Nel, 2003), and vocabulary acquisition (Oberg, 2011). In terms of terminology, Computer-Assisted Language Learning (CALL) is a well-established field with a long history, and is often used as a catch-all phrase for research on computer use in L2 settings (Bax, 2003; Bax, 2011). In recent years, the term Technology-Enhanced Language Learning (TELL) has often been used interchangeably with or even in preference to CALL (Hubbard, 2013; Walker & White, 2013), as TELL suggests a more inclusive sense of technology interventions enabled by a wide variety of technology devices and software applications. As a result, TELL is used in this article to refer to the employment of technology in various forms in the application domain of L2 education (Chang & Hung, 2019).

Many scholars have discussed the advantages of mobile devices in education, including broadening learning outside the conventional classroom (Wu, 2016), allowing convenient access to learning materials (Kaliisa et al., 2019), flexibility in the time and location of study (Loewen et al., 2019), and creating more interaction and communication between teachers and students (Wu, 2014). The related studies include the use of mobile devices for vocabulary learning (Asmana & Arifani, 2021; Katemba, 2021), reading comprehension (Alnujaidi, 2021; Rahimi & Babaei, 2021), sentence construction (Purgina et al., 2017), listening (Islam & Hasan, 2020; Kamasak et al., 2021), speaking skills (Almadhady et al., 2019; García Botero et al., 2019; Lutfi, 2020; Xu, 2020), writing (Gharehblagh & Nasri, 2020; Krisbiantoro, & Pujiani, 2021), and grammatical development (Ghorbani & Ebadi, 2020). With the development of smartphones, an increasing number of applications have been developed to offer a variety of functions that can assist with English learning. The use of technologies for language learning has become nearly ubiquitous.

New technologies (e.g., AI, virtual reality, augmented reality, or wearable technologies) are increasingly available (Shadiev et al., 2019). Furthermore, emerging technologies are maturing and are very promising for use in language learning and instruction. Studies have shown that technology can promote the learning performance of language learners, increase learning motivation, and provide them with more efficient means of language learning (Jin, 2018; Shadiev & Huang, 2020).

Shadiev and Yang (2020) reviewed 398 research articles on technology-enhanced language learning and teaching from 2014 to 2019. The types of technology used in the articles we reviewed are listed in Table 1 (adopted from Shadiev & Yang, 2020). In these articles, the most used technologies were games (n = 49) and online videos (n = 37). There are, however, few studies that have explored the use of new technologies (e.g., AI n = 0, VR n = 19, AR n = 3, wearable technology n = 2)

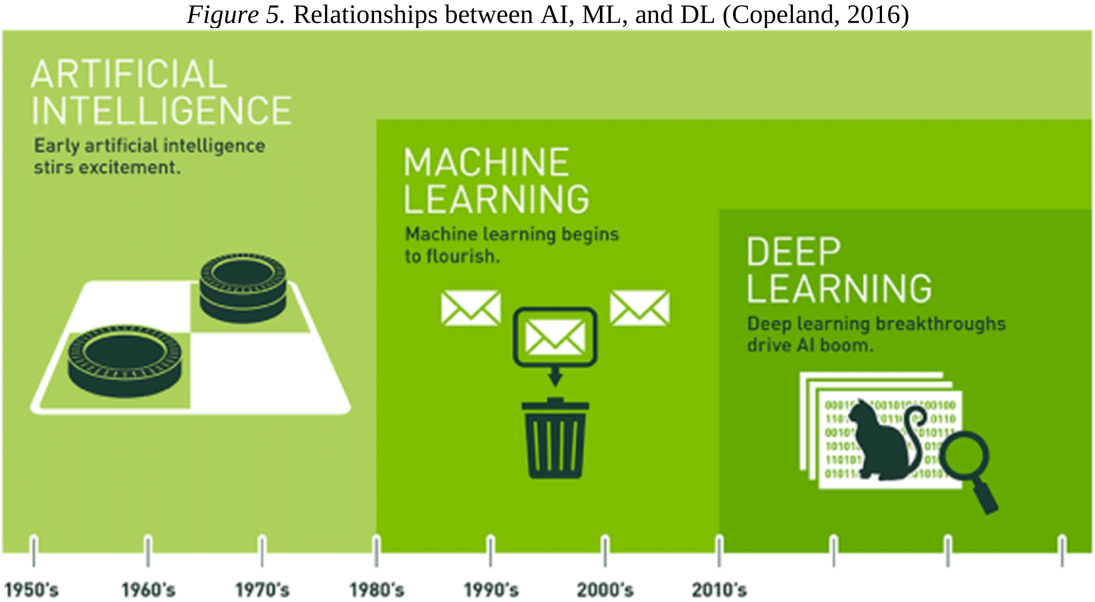
Table 1. Technologies used in language learning and teaching from 2014 to 2019

Technology	Year						Total
	2014	2015	2016	2017	2018	2019	
Game	6	7	6	11	14	5	49
Online video	5	7	4	6	10	5	37
Collaborative writing	3	9	8	5	3	7	35
Corpus	8	5	9	5	2	3	32
Instant messaging	2	4	4	11	5	5	31
Automated feedback	6	4	4	4	5	6	29
Social networking	1	6	4	10	7	1	29
Websites and digital resources	3	2	5	6	3	6	25
Virtual reality	0	5	2	4	6	2	19
Speech recognition	0	3	5	4	3	3	18
Electronic gloss or annotation	3	1	0	3	2	1	10
E-books	0	0	3	3	2	1	9
Electronic dictionary	2	0	3	1	1	1	8
Intelligent tutoring system	1	1	1	1	0	2	6
Voice recording	2	0	0	1	2	1	6
Augmented Reality	0	0	0	2	1	0	3
Robots	0	1	1	0	1	0	3
Clicker	0	0	0	0	1	2	3
Wearable devices	0	0	1	0	1	0	2
Course management system	2	0	0	0	0	0	2
Digital library	1	0	0	0	0	0	1
White board	1	0	0	0	0	0	1
Unidentified technology	7	6	9	5	9	12	48

2.2. AI and language acquisition

Artificial intelligence (AI) is a branch of computer science that focuses on the creation of intelligent machines that think and work like humans (Xie et al., 2021). Its goal is to replicate human intelligence in machines that are programmed to think and act like humans. AI plays an important role in the technology that supports daily social life. Machine learning (ML) is an AI system that can learn on its own by using an algorithm (Rahimy, 2018). Deep learning (DL) is a type of machine learning that is used to analyze large amounts of data by using a cascade of multilayered convolutional neural networks (CNN) (Lee et al., 2021; Litjens et al., 2017; Shen et al., 2017) (see Figure 5).

Figure 5. Relationships between AI, ML, and DL (Copeland, 2016)



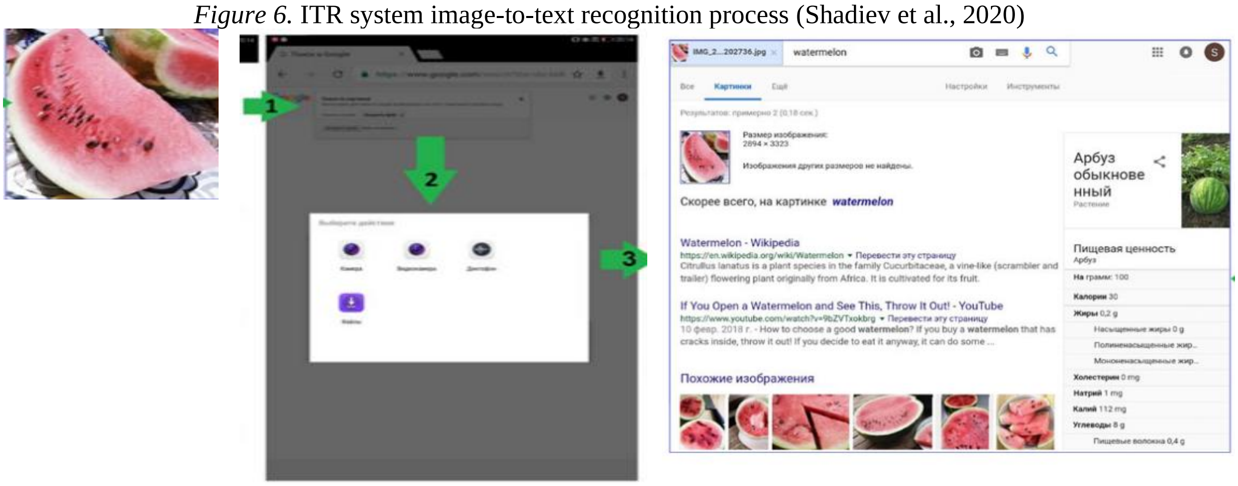
The emerging AI technologies in teaching and studying are used to provide more customized, versatile, inclusive, and engaging learning, as well as to automate everyday learning activities via automated evaluation and feedback (Gulson et al., 2018; Luckin et al., 2016). Theoretically, AI can assist parents in optimizing their children’s early linguistic development, as well as teachers in choosing resources, planning courses, increasing attendance, and delivering customized instruction to their students (Porayska-Pomsta, 2016). In its early stages, AI in education generally referred to intelligent tutoring systems that aimed to solve problems such as automatically improving the operator performance (e.g., Chai et al., 2021; Hwang, 2003; Ross, 1987). Nowadays, AI refers to the use of large amounts of data to complete complex tasks.

DL has sparked a surge of interest in business and science over the last decade, revolutionizing the field of machine learning by achieving state-of-the-art results in perception tasks such as image and speech recognition (Goodfellow et al., 2016). Recently, DL has obtained the top results in a wide range of computer vision issues, and has proved to be particularly effective for image recognition (Mezgec et al., 2019), a visual deep learning model, in which a convolutional neural network (CNN) is applied. A CNN is made up of thousands of individual neurons that can perform complex tasks, such as pixel intensity-based image recognition and classification (Rahimy, 2018). In other words, deep learning techniques allow for this automatic learning by absorbing large quantities of unstructured data such as text, images, or video (Salloum et al., 2020).

To the best of our knowledge, thus far, there have been only a few studies on the application of AI to language learning (i.e., Hasnine et al., 2019; Hasnine et al., 2020; Shadiev et al., 2020). These three studies focused on the development of AI-assisted systems for vocabulary learning. They each developed a system which allowed students to take pictures with smartphones and then upload them to the web, to be translated using Google Translate. Learners could make text or image annotations on the smartphone systems to support their vocabulary learning. However, the systems did not directly provide an object detection function or show how to spell or pronounce the word in addition to the translation function.

Shadiev et al. (2020) developed a learning system featuring image-to-text recognition (ITR) technology to support Russians learning English as a foreign language (see Figure 6). An Android-based mobile learning system was developed in their study and was installed on tablet PCs. In contrast to traditional image retrieval by typing in keywords, this system allowed users to search by submitting a sample image as their query. The Google Images service was employed for the ITR process. This service allows users to search the Web for image content. In comparison, the learners in the control group used the conventional approach, looking at the images that corresponded to the questions in their textbook. The experimental group outperformed the control group on the vocabulary tests.

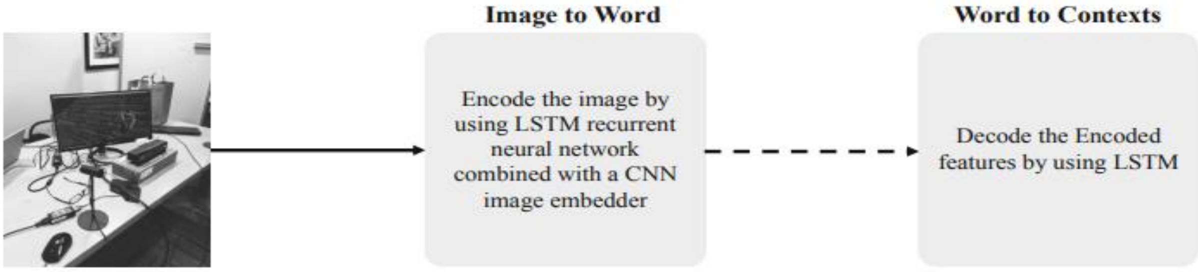
Figure 6. ITR system image-to-text recognition process (Shadiev et al., 2020)



(a) Take photo (b) Upload to the ITR system (c) Use the Google Image service to search the Web for image content

Hasnine et al. (2019) and Hasnine et al. (2020) created a system that allows students to take pictures that they upload as feedback (see Figure 7). The system then examines the visual content of these images and creates unique learning contexts on the basis of the visual content. Language learners could connect their previous knowledge with new knowledge by using this system, allowing them to review and recall the previously acquired vocabulary. Figure 7 illustrates the architecture of the model. In this model, the image encoder is a deep convolutional neural network (CNN), which is widely adopted for object recognition and detection tasks. The architecture of the model uses the Long Short-Term Memory (LSTM) network which is trained as a language model conditioned on image encoding. In the LSTM network, words in the captions are represented with an embedding model where each word is associated with a fixed-length vector representation that is learned during training.

Figure 7. Hasnine's et al. (2019; 2020) system image-to-word recognition process



3. Methodology

This study aimed to examine the effect of the AI-based ODT app on different ability EFL students' vocabulary learning using a quantitative approach. A randomized subjects, pre-test–post-test control group experimental design was adopted. The difference between the experimental group (EG) and the control group (CG) was the aid used to support task completion during the classroom activity sessions. Each of the learners in the EG was given a smartphone with the AI-based ODT app to finish the task guided by the instructor right after the teaching session, whereas the learners in the CG used Google Translate for the task; that is, they were given a vocabulary worksheet related to the teaching content. The dependent variable of this study was the different ability students' vocabulary performance on the post-test, and the independent variable was the different teaching platforms the students experienced (the AI-based ODT app vs. Google Translate).

Each group's average difference between the pre- and post-tests was calculated, and the average difference scores were compared to determine whether the experimental treatment produced a greater change than the control situation.

3.1. Participants

There were 72 participants involved in this research. They were students from a public elementary school located in a rural area in central Taiwan, and their native language was Mandarin Chinese. The average age of the participants was 11 years, with a range of 10 to 12 years. The participants had received formal English education in their elementary school for at least one year.

Each participant took the First Step level examination of Anglia English Speakers for Other Languages International Examinations (First Step Level of Anglia ESOL Exams) (Anglia Examinations England, 2014). The total score for the examination is 100 points. Those who passed the examination (with a passing score of 50) could understand the first useful structures such as classroom commands and how to introduce themselves, as well as approximately 100 words from familiar categories such as numbers, family members, colors, and household items. The mean score for the participants was 76.83.

The participants were divided into higher-level and lower-level groups based on the results of the scores of the First Step Level of Anglia ESOL Exams. Stratified sampling was then used to select samples from different level participants for the two groups—the experimental group and the control group—which ensured that the two groups were equivalent in terms of knowledge and ability.

3.2. Instruments

3.2.1 Vocabulary tests

In this research, a test sheet was designed to test the participants' vocabulary recognition. The 50 target words were selected from a daily market shopping experience, including 29 types of fruit and 21 vegetables. The participants were required to write the Chinese translation of the English words. The total score of the test was

100 points, with every question worth 2 points. The reliability coefficient value of the test was .93, indicating high reliability.

3.2.2. Class instruction and classroom activity

We planned four target vocabulary lessons, which followed the Presentation, Practice, Production model (PPP model) that many foreign language teaching course books are based on (Hutasoit et al., 2020). In the Presentation and Practice stages, both the CG and EG followed the same instruction. However, in the Production stage, the information exchange activity was applied to the CG by using a crossword puzzle, while in the EG, a role play activity was conducted.

Presentation (10 mins): In this stage, the instructor began the class lessons by presenting the target words and model sentences. One of the researchers as the instructor introduced the course first and then started to teach the vocabulary. Each week, the instructor began the presentation with 8-10 target words and the corresponding sentence structures. The instructor used flash cards to lead into the structure to be taught.

Practice and production (80 mins): In the practice stage, the new language was practiced by the students in a controlled manner. Students repeated the new words together and then separately right after the teacher to be able to say them correctly. In the production stage, the students used the language in context. For the EG, each student had a smartphone with the AI-based ODT app, which they could manipulate to scan the fruit or vegetables, listen to the pronunciation, and say the words correctly. The students in the CG were shown the same fruit or vegetables, and used Google Translate by typing the Chinese. They were asked to pronounce the word in English. A snapshot of the exploration time of the EG is shown in Figure 8.

Figure 8. Students used the AI-based ODT app to learn English vocabulary

3.2.3. The AI-based ODT app

The research team utilized TensorFlow to train the object detection model with the daily fruit data, and provided the corresponding English words for the pictures. After training, the model was exported to the mobile app as an English language learning tool. For training the object detection model, the camera was used to take 360° photos of each fruit, and then approximately 500 different angles of the fruit were chosen from the photos, which were arranged in a folder that was labeled with the name of the fruit (Chang et al., 2019; Liu et al., 2021).

Mobile Net was trained using TensorFlow's open source code tensorflow-for-poets-2, and the bottleneck model was used to implement the image recognition model. The training image was 224 pixels wide, and MobileNet's initial relative size was set to 1.0 (see Figure 9). Furthermore, Intel Core i5 2.70 GHz CPU, 4 GB RAM, and a 100-Mbps network environment were used in the research. The trained modules were connected to the mobile phone by using Android Studio 3.0. From the training dataset, two types of fruit, namely oranges and apples, were chosen for testing. The rationale for choosing these two types of fruit was that their appearance is similar (see Figure 9), making it easy to confuse the image recognition model.

We split the dataset into 10 parts and used a 10-fold cross-validation to determine the model's accuracy. Cross validation is a method applied to a model and a data set in an effort to estimate the out of sample errors. The relative sizes and pixel settings of different Mobile Nets were then recorded. Following the training, 16 models were obtained, and the data from TensorFlow were used for further analysis. We used a tensor board, which is a tool for providing the measurements and visualizations needed during the machine learning workflow, to compare the recognition rates of 16 image recognition models trained with different parameters, and then we chose the best image recognition model for use on mobile phones. Figure 10 shows the training results of the app accuracy. The results revealed that this study's image recognition model was very accurate (98%) (Chang et al., 2019; Liu, 2018). Figure 8 shows that the EG students used the app to identify the objects (e.g., cabbage and carrot) and learned English words with English/Chinese word translations and pronunciation.

Figure 9. Image recognition results for orange (93%) and apple (97%)

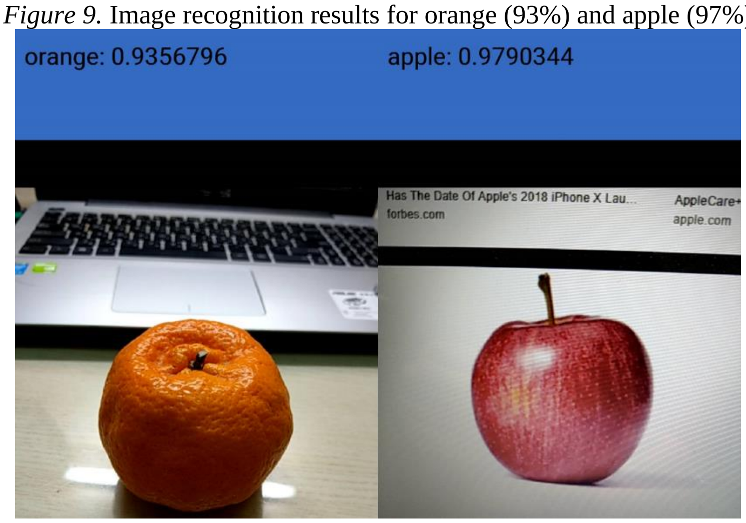
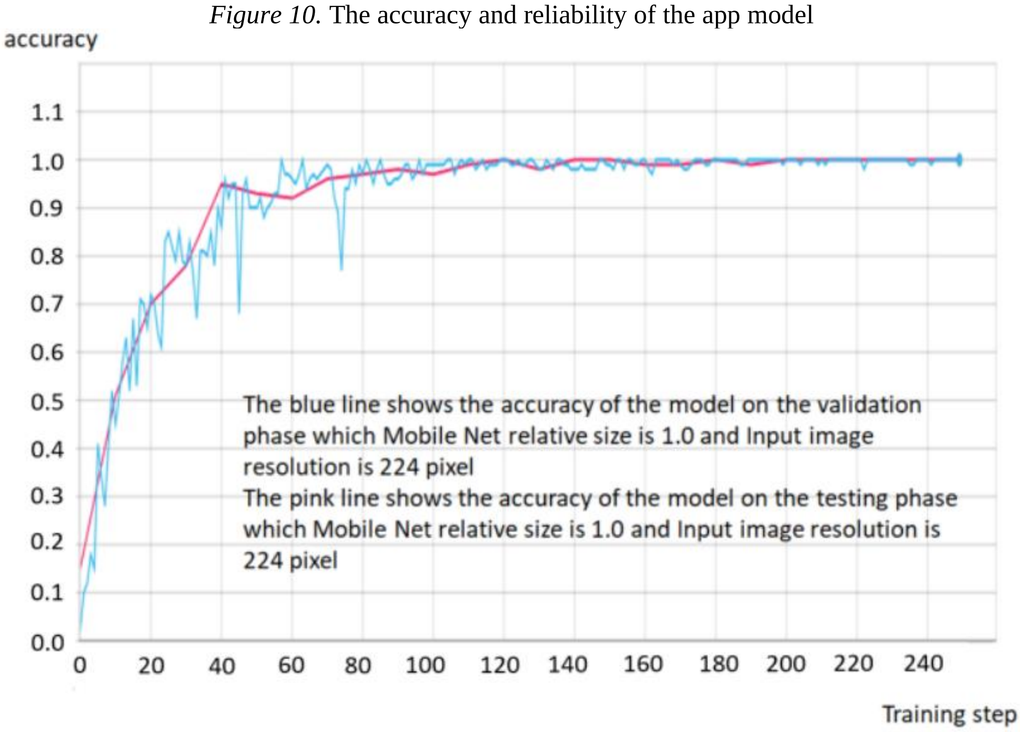


Figure 10. The accuracy and reliability of the app model



3.3. Procedure

The total duration of the experiment was 11 weeks. In the first week, all participants were asked to take the First Step Level of Anglia ESOL Exams to measure their English proficiency at the beginning of the experiment. Stratified sampling was then used to select samples from different level participants for the EG and CG. In the second week, the vocabulary pre-test was administered to represent the learners' prior knowledge of the English learning content. In the following paragraphs, the procedure and the content of six vocabulary lessons will be illustrated.

In order to reduce the practice effects, the researchers provided a two-week warm-up story reading exercise before the experiment began. From weeks 5 to 10, six vocabulary lessons were conducted, with one per week. There were 50 target words in total. The vocabulary related to fruit was taught in lessons 1 to 3, including star fruit, dragon fruit, and grapes. Next, the target vocabulary related to vegetables was taught in lessons 4 to 6, such as cauliflower, cucumber, and garlic. In each 90-min class, we followed the PPP model to teach the target vocabulary and provide opportunities for practice. During the class periods with the PPP model, all the participants learned the target vocabulary and applied it to the model sentences during the presentation stage. In the practice and production stages, the two groups were separated into two different classrooms with different activities for practice during which they applied the vocabulary learnt in class. For the EG, each student had a smartphone with the AI-based ODT app to scan the fruit or vegetables. The students in the CG were shown the same fruit and vegetables and used Google Translate by typing in the Chinese. Both groups were asked to listen to the pronunciation and to try to say the words correctly.

After six vocabulary lessons, in week 11, all participants were asked to take the vocabulary post-test. The post-test had different permutations of the questions from the pre-test. An independent t test was conducted using the vocabulary pre-test and post-test to compare the difference of the EG and CG data as the independent variable.

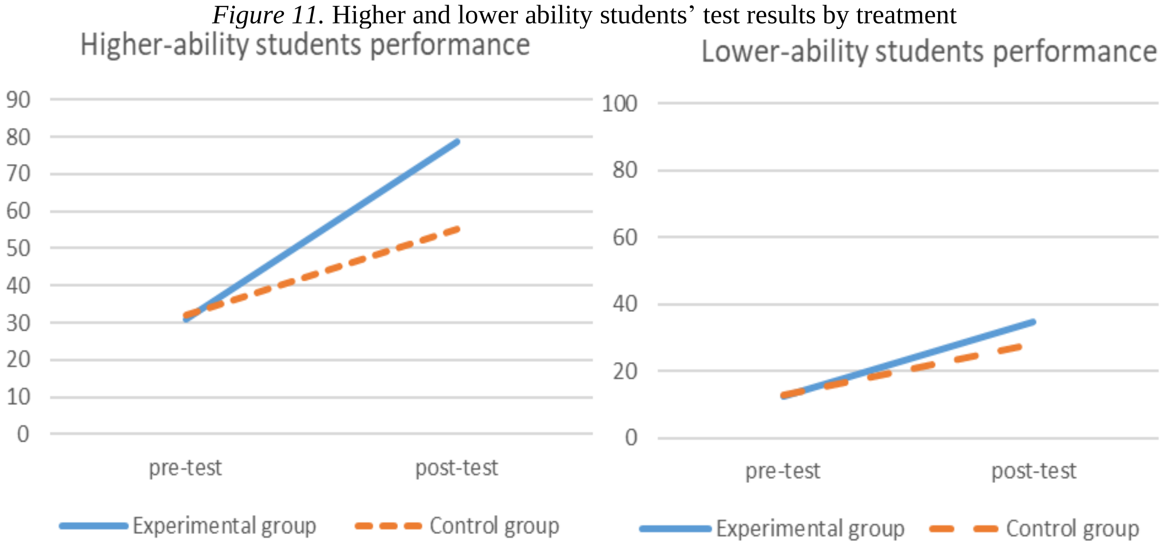
4. Results

The pre-test mean score of the EG group was 18.66 and that of the CG group was 23.66. The mean score of all participants was 21.17. This result indicated that there was no statistically significant difference between the EG and CG groups in the pre-test (t = 1.16, p = .25), indicating that these two groups had similar vocabulary ability prior to the experiment.

Table 2. Descriptive data of the different ability students’ test results by treatment

	N	Pre-test		Post-test		*MD*
		M	*SD*	*M*	*SD*	
Higher-ability						
EG	18	30.89	10.91	78.67	16.92	50.00
CG	18	32.22	24.00	55.33	24.49	22.00
Total	36	31.56	17.96	67.00	23.88	36.00
Lower-ability						
EG	18	12.33	4.92	34.67	16.21	22.34
CG	18	13.00	10.23	28.00	3.50	15.00
Total	36	12.67	7.88	31.33	12.04	19.29

Figure 11. Higher and lower ability students’ test results by treatment



We then employed the pre-test scores for further analysis. The students were divided into different ability groups according to their pre-test scores. Students who scored above the mean score in the pre-test were identified as higher-ability learners and those who scored at or below the mean score were identified as lower-ability learners. This resulted in 36 students in the app treatment (higher-ability M = 28.67, lower-ability M = 12.33) and 36 in the traditional treatment (higher-ability M = 33.33, lower-ability M = 13.00). There was also no significance difference between the higher-ability students (t = .77, p = .44) and lower-ability students (t = .51, p = .62) in the two groups on the pre-test (see Table 2, Figure 11).

The data analysis for student achievement was a 2 (Experimental treatment vs. Traditional treatment) × 2 (Ability Level: Higher-ability vs. Lower-ability) Two-Way Repeated measurement ANOVA on the pre-test and post-test (see Table 3)**.** The analysis included between-subjects variables such as treatment and skill level, as well as within-subjects variables such as test occasion and problem type. Table 2 shows the post-test results for the two types of treatment of the different ability students. The data on each variable's achievement is discussed below. Significant differences in English entering understanding and treatment were found using a 2 x 2 ANOVA. In terms of English entering knowledge level, students with higher-ability English entering knowledge outperformed those with lower-level English entering knowledge by a substantial margin (M = 67.00 and M = 31.33, respectively), $F(1, 72)$ = 78.93, MS = 22898.00, p = .00*. For treatments, the mean correct scores were 41.67 for the OTR translation app subjects and 56.67 for the traditional subjects, $F(1, 72)$ = 13.96, MS = 4050.00, p = .00*. The 2 x 2 ANOVA also yielded a significant two-way interaction for treatment by English ability, $F(1, 72)$ = 4.31, MS = 1250.00, p =.04*. This two-way interaction reflected the fact that the OTR translation app higher-ability students had considerably higher scores than the traditional higher-ability students (M = 78.67 vs. M = 55.33), but the lower-ability students of the two treatments had similar scores (M = 34.67 vs. M = 28.00) on the post-test.

Table 3. Test of between-subject effects

	SS	*df*	*MS*	*F*	*p*
Treatments	4050.00	1	4050.00	13.96	.00*
Ability	22898.00	1	22898.00	78.93	.00*
Treatments x Ability	1250.00	1	1250.00	4.31	.04*
Error	19728.00	68	290.12		
Total	221976.00	72			

Note. *p < .05.

As the follow-up to the two-way interaction, further univariate analysis of variance revealed that the slight difference in post-test scores by treatment for students with lower-ability was not statistically significant (F = 2.91, p = .54), whereas the difference in scores favoring the AI-based ODT app over the traditional treatment for students with higher-ability was significant (F = 11.07, p = .00*).

5. Discussion

In terms of vocabulary attainment, there was a substantial difference between the experimental and control groups; students who learned using the AI-based ODT app technology scored higher on the post-test than those who learned using the non-AI technique. This finding is consistent with the findings of previous research (Asmana & Arifani, 2021; Hasnine et al., 2019; Hasnine et al., 2020; Katemba, 2021; Shadiev et al., 2020). First of all, the experimental learners learned vocabulary more actively than their counterparts. By integrating the word translation and pronunciation, this type of learning based on the Dual-Coding Theory (DCT) differs from rote learning (Chuo & Yen, 2014; Wu & Huang, 2017). It is easier for learners to remember vocabulary if both the visual and verbal channels are integrated simultaneously while learning the input information (Kassim, 2018).

According to DCT (Clark & Paivio, 1991; Kanellopoulou et al., 2019; Kassim, 2018), a cognition theory, a learner's memory consists of two separate but interrelated verbal and visual codes for processing information. Interestingly, there exists an interconnection between the two separate systems which facilitates dual coding of information if not activated independently. Psychologists have demonstrated how our minds respond well to words that form a picture, and some studies have found that individuals who read illustrated text outperformed those who read text alone (Mayer & Sims, 1994).

In this study, we examined the interconnection between the combination of the AI-based ODT app which is responsible for visual capability and speech pronunciation and which provides the verbal capability. The visualized spoken words, based on DCT, may not only improve a learner's recollection of English vocabulary, but also stimulate other information processing such as feature identification and linkage with past knowledge. The detailed examination of the pre-post test results provided an insight into how the combination, actualized through different teaching platforms (AI and Non-AI with object detection-enabled or disabled), scaffolds children's experiences and learning of a second language (Dalim et al., 2020).

In both experiments, the participants who used the AI-based ODT app platform had a significantly higher knowledge gain than those who used the non-AI platform (Google Translate app). Unlike the non-AI teaching platform which provides limited ability to manipulate the visual feedback, participants were able to scan objects with the AI-based ODT app in a more interactive way. The ability for participants to hold the mobile phone to scan and see the input on the screen while remaining in their real environment made the learning more interesting. This AI-based ODT app learning activity ensured active learning. Kim (2011) believed that if students are actively engaged in the cognitive processes of vocabulary acquisition, they will learn more words and retain them for longer. Moreover, because the experimental learning task was considerably more complicated and challenging than the traditional activities, the higher-ability students were more involved in this activity than the lower-ability students.

Another reason for the above-mentioned outcome was that the experimental learners' first encounter with this technology was a novel and exciting experience for them. According to previous research, when new technologies are used in education, they motivate students to participate in the learning process. As a result, students are more engaged in the learning process, making it easier for them to comprehend the material (Sahin & Yilmaz, 2020). Moreover, the AI-based ODT app is interesting for students and attracts their attention (Hasnine et al., 2019; Hasnine et al., 2020; Shadiev et al., 2020). It also helps them achieve their goals. According to the literature, AI may maximize children's early linguistic development, as well as increase attendance and provide students with personalized instruction (Porayska-Pomsta, 2016).

According to Venkatesh and Davis (2000), if learners find a system to be simple to use and useful, they will use it again. The findings revealed that the majority of the experimental group students approved of the learning system and were pleased with its use, as their interaction with the system was simple and clear. The acceptance of such systems is largely attributed to the accuracy of their recognition processes (Shadiev & Sun, 2020). To train the picture recognition model, a camera was used to take 360° photos of each fruit, and then approximately 500 different angles of the fruit were chosen from the photos. The results revealed that the proposed image recognition model was very accurate (98%) (Chang et al., 2019; Liu, 2018). As a result, the learners found the AI-based ODT app to be easy to use. The greatest challenge observed in the use of the Google Translate system was the difficulty the children faced in synchronizing their hand movements when typing the Chinese or English while searching for the translation; this was especially true for the lower-ability students.

Despite the positive results of this study, there are some limitations that should be noted. For example, for training the picture recognition model, we selected only 50 types of fruit and vegetable because of the limitations of availability and time. As a result, larger database studies are needed to add more categories to ensure the validity of the methodological approach. We also found that the AI based-ODT app benefited the higher-ability students more than the lower-ability students. It would be more beneficial if the design of the app could be modified to help the lower-ability students learn vocabulary more easily. This would provide a more significant contribution to the topic of language learning.

6. Conclusion and future work

In this article we have presented an AI based-ODT app for teaching young children who are non-native English speakers the English vocabulary for basic fruit and vegetables. Based on previous studies, our system is the first AI based-ODT app language-learning tool designed to teach young children. The objectives of this study were to ascertain how effective the AI based-ODT app combined with the DCT design was compared to the use of the non-AI Google Translate app, and to explore if the use of object detection translation was able to increase the effectiveness of vocabulary learning. The findings suggest that the AI-based ODT app could be useful as a teaching aid for young children, as it boosts learning engagement and knowledge gain. Real-time interaction encourages children to delve more deeply into the learning materials. The participants had positive experiences of interacting with the AI-based ODT app interface over the non-AR interface. Our findings conclude that the speech-enabled AR interface is usable for younger children even with little or no prior AR experience, and provides a motivating factor for their foreign language learning. The results show significant evidence of knowledge gain and a positive inclination to use the AR interface over the Google Translate interface.

In terms of future work, more research is needed to discover the potential of AI-based ODT for young children's activities. This study could be replicated in other domains where learners could learn about objects by differentiating their knowledge. For example, learners learning sign language for special purposes may use the app to take photos of sign language to learn the corresponding translations. We also suggest that future studies elaborate more on the benefits of ODT for language learning, and determine which of the three functions of the app (picture, text, and pronunciation) is the most beneficial and meaningful for language learning.

References

Almadhady, A. A., Salam, A. R. H., & Baharum, H. I. (2020). The Motivation of Arab EFL university students towards using MALL applications for speaking improvement. *Universal Journal of Educational Research*, *8*(11), 23-36.

Alnujaidi, S. (2021). Adoption of mobile assisted language learning (MALL) in Saudi Arabian EFL Classrooms. *Journal of Language Teaching and Research*, *12*(2), 312-323.

Asmana, C. H., & Arifani, Y. (2021). Mobile-assisted language learning based using scientific approach to improve students' vocabulary. *Journal of English Teaching, Literature, and Applied Linguistics, 4*(1), 43-51.

Anglia Examinations England (2014). *Handbook for teachers: Full examination syllabus and specifications for the revised exams 2014/2015*. Anglia Examinations Syndicate Limited.

Bax, S. (2003). CALL: Past, present and future. *System, 31*(1), 13-28.

Bax, S. (2011). Normalisation revisited: The Effective use of technology in language education. *International Journal of Computer-Assisted Language Learning and Teaching, 1*(2), 1-15.

Burston, J. (2014). MALL: The Pedagogical challenges. *Computer Assisted Language Learning, 27*(4), 344-357.

Chang, H. Y., Liu, P. L., Wang, W. C., & Chen, T. C. (2019). English learning tool for elders-Design of instant fruit identification system with TensorFlow. *Proceedings of* the *2019 IEEE International Conference on Consumer Electronics.* https://doi.org/10.1109/ICCE-TW46550.2019.8991988

Chang, M. M., & Hung, H. T. (2019). Effects of technology-enhanced language learning on second language acquisition: A Meta-analysis. *Journal of Educational Technology & Society*, *22*(4), 1–17.

Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K. F., & Qin, J. (2021). Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students. *Educational Technology & Society, 24*(3), 89-101.

Chatterjee, S., & Bhattacharjee, K. K. (2020). Adoption of artificial intelligence in higher education: A Quantitative analysis using structural equation modeling. *Education and Information Technologies*, *25*(5), 3443-3463.

Chun, D. M. (2016). The Role of technology in SLA research. *Language Learning & Technology, 20*(2), 98-115.

Chuo, T. W. I., & Yen, S. C. H. (2014). The Learning journey of college at-risk EFL students in Taiwan: An exploratory study. *The Asian EFL Journal Quarterly, 16*(2), 44-68.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-210. https://doi.org/10.1007/BF01320076

Copeland, M. (2016). What's the difference between artificial intelligence, machine learning and deep learning? https://blogs.nvidia.com/wp-content/uploads/2016/07/Deep_Learning_Icons_R5_PNG.jpg.png.webp

Dalim, C. S. C, Sunar, M. S., Dey, A., & Billinghurst, M. (2020). Using augmented reality with speech input for non-native children's language learning. *International Journal of Human-Computer Studies, 134*, 44-64. https://doi.org/10.1016/j.ijhcs.2019.10.002

de Carvalho, T. B., Sibaldo, M. A., Tsang, R., Cavalcanti, G. D., Sijbers, J., & Tsang, J. (2018). Intensity patches and pegion patches for image recognition. *Applied Soft Computing*, *62*(2018), 176-186.

Dreyer, C., & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. *System, 31*(3), 349-365.

Eslami, R., & Ahmadi, S. (2019). Investigating the role of educational media on secondary school students' learning process improvement in Jahrom city. *Journal of Humanities Insights*, *3*(1), 13-16.

Fu, C. (2016, May 12). *Google Translate real-time camera translation new function, full support for traditional Chinese recognition*. https://www.cool3c.com/article/106214

García Botero, G., Questier, F., & Zhu, C. (2019). Self-directed language learning in a mobile-assisted, out-of-class context: do students walk the talk? *Computer Assisted Language Learning*, *32*(1-2), 71-97.

Gharehblagh, N. M., & Nasri, N. (2020). Developing EFL elementary learners' writing skills through mobile-assisted language learning (MALL). *Teaching English with Technology*, *20*(1), 104-121.

Ghorbani, N., & Ebadi, S. (2020). Exploring learners' grammatical development in mobile assisted language learning. *Cogent Education*, *7*(1), 1704599. https://doi.org/10.1080/2331186X.2019.1704599

Giovany, S., Putra, A., Hariawan, A. S., & Wulandhari, L. A. (2017). Machine learning and sift approach for Indonesian food image recognition. *Procedia Computer Science*, *116*, 612-620.

Good, O. (2015, July 29). *How Google Translate squeezes deep learning onto a phone.* https://ai.googleblog.com/2015/07/how-google-translate-squeezes-deep.html

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning*. MIT press.

Gu, X. (2019, July 10). *Google Translate's instant Camera Translation gets an upgrade*. https://blog.google/products/translate/google-translates-instant-camera-translation-gets-upgrade/

Gulson, K. N., Murphie, A., Taylor, S., & Sellar, S. (2018). *Education, work and Australian society in an AI world: A Review of research literature and policy recommendations.* Gonski Institute for Education.

Hasnine, M. N., Flanagan, B., Akcapinar, G., Ogata, H., Mouri, K., & Uosaki, N. (2019). Vocabulary learning support system based on automatic image captioning technology. *Lecture Notes in Computer Science, 11587*, 346-358. doi.org/10.1007/978-3-030-21935-2_26

Hasnine, M. N., Mouri, K., Akcapinar, G. Ö. K. H. A. N., Ahmed, M. M. H., & Ueda, H. (2020). A New Technology Design for Personalized Incidental Vocabulary Learning using Lifelog Image Analysis. *In Proceedings of the 28th International Conference on Computers in Education (ICCE2020)* (pp. 516-521).

Hubbard, P. (2013). Making a case for learner training in technology enhanced language learning environments. *CALICO Journal, 30*(2), 163-178.

Hutasoit, E. V., Elfiza, R., & Subroto, G. (2020). The Effect of using PPP method on students' speaking skill. *Student Online Journal, 1*(2), 258-263.

Hwang, G. J. (2003). A Conceptual map model for developing intelligent tutoring systems. *Computers & Education*, *40*(3), 217-235.

Islam, A. B. M. S., & Hasan, M. (2020). The Effectiveness of mobile assisted language learning (MALL) on ESL listening skill. *NOBEL: Journal of Literature and Language Teaching*, *11*(2), 188-202. https://doi.org/10.15642/NOBEL.2020.11.2.188-202

Jin, L. (2018). Digital affordances on WeChat: Learning Chinese as a second language. *Computer Assisted Language Learning, 31*, 27-52.

Kaliisa, R., Palmer, E., & Miller, J. (2019). Mobile learning in higher education: A Comparative analysis of developed and developing country contexts. *British Journal of Educational Technology*, *50*(2), 546-561.

Kamasak, R., Özbilgin, M., Atay, D., & Kar, A. (2021). The Effectiveness of mobile-assisted language learning (MALL): A Review of the extant literature. In A. S. Moura, P. Reis, & M. N. D. S. Cordeiro (Eds.), *Handbook of Research on Determining the Reliability of Online Assessment and Distance Learning* (pp. 194-212). https://doi.org/10.4018/978-1-7998-4769-4

Kanellopoulou, C., Kermanidis, K. L., & Giannakoulopoulos, A. (2019). The Dual-coding and multimedia learning theories: Film subtitles as a vocabulary teaching tool. *Education Sciences, 9*(3), 210. https://doi.org/10.3390/educsci9030210

Kassim, W. (2018). Utilizing dual coding theory and animated images to enhance ESL students'vocabulary learning. *The English Teacher, 47*(3), 81-91.

Katemba, C. V. (2021). Enhancing vocabulary performance through mobile assisted language learning at a rural school in Indonesia. *Acuity: Journal of English Language Pedagogy, Literature and Culture, 6*(1), 1-11.

Kim, Y. (2011) The Role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning, 61*(1), 100–140. https://doi.org/10.1111/j.1467-9922.2011.00644.x

Kong, F., He, H., Raynor, H. A., & Tan, J. (2015). DietCam: Multi-view regular shape food recognition with a camera phone. *Pervasive and Mobile Computing, 19*, 108-121.

Krisbiantoro, B., & Pujiani, T. (2021). The Effectiveness of mobile-assisted language learning to teach writing. In L. Y. Rasnani, & L. Mareza (Eds.), *Empowering Human Development Through Science and Education* (pp. 1-15). https://proceedings.aecon.ump.ac.id/index.php/aecon/article/view/15/2

Lee, C. A., Tzeng, J. W., Huang, N. F., & Su, Y. S. (2021). Prediction of student performance in massive open online courses using deep learning system based on learning behaviors. *Educational Technology & Society, 24*(3), 130-146.

Lee, M. C., Chiu, S. Y., & Chang, J. W. (2017). A Deep convolutional neural network based Chinese menu recognition app. *Information Processing Letters, 128*, 14-20.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., & Sánchez, C. I. (2017). A Survey on deep learning in medical image analysis. *Medical image analysis, 42*, 60-88.

Liu, P. L. (2016). Mobile English vocabulary learning based on concept-mapping strategy. *Language Learning & Technology, 20*(1), 128-140.

Liu, P. L. (2018). *Integrating concept map and AI image recognition translator on elders English vocabulary learning* (Report No. MOST 107-2635-H-415-002-). Ministry of Science and Technology.

Liu, P. L., Chen, C. J., Cheng, Y. H. (2021, June). *The Effect of image recognition mobile APP applied on EFL students' vocabulary learning* [Paper presentation]. 4th Pedagogy and Practice in Technology Enhanced Language Learning (PPTELL 2021), Taipei, Taiwan.

Liu, X., Liu, C. H., & Li, Y. (2020). The Effects of computer-assisted learning based on dual coding theory. *Symmetry, 12*(5), 701-713.

Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL: The Journal of EUROCALL, 31*(3), 293-311.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An Argument for AI in education.* Pearson Education

Lutfi, N. (2020). The Integration of MALL to enhance students speaking skill: An Autonomous learning model. *Journal of Foreign Language Teaching and Learning, 5*(1), 1-19.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Education Psychology, 86*(3), 389-401.

Mezgec, S., Eftimov, T., Bucher, T., & Seljak, B. K. (2019). Mixed deep learning and natural language processing method for fake-food image recognition and standardization to help automated dietary assessment. *Public Health Nutrition, 22*(7), 1193-1202.

Nation, P. (2013). *What should every EFL teacher know?* Compass Publishing.

Oberg, A. (2011). Comparison of the effectiveness of a CALL-based approach and a card-based approach to vocabulary acquisition and retention. *CALICO Journal, 29*(1), 118-144.

Porayska-Pomsta, K. (2016). AI as a methodology for supporting educational praxis and teacher metacognition. *International Journal of Artificial Intelligence in Education, 26*(2), 679-700.

Rahimy, E. (2018). Deep learning applications in ophthalmology. *Current Opinion in Ophthalmology, 29*(3), 254-260.

Ramírez Verdugo, D., & Alonso Belmonte, I. (2007). Using digital stories to improve listening comprehension with Spanish young learners of English. *Language Learning & Technology, 11*(1), 87-101.

Ramos, R., & Dario, F. (2015). Incidental vocabulary learning in second language acquisition: A Literature review. *Profile Issues in Teachers' Professional Development, 17*(1), 157-166.

Ross, P. (1987). Intelligent tutoring systems. *Journal of Computer Assisted Learning, 3*(4), 194-203.

Sahin, D., & Yilmaz, R. M. (2020). The Effect of augmented reality technology on middle school students' achievements and attitudes towards science education. *Computers & Education, 144*(2020), 103710. https://doi.org/10.1016/j.compedu.2019.103710

Salloum, S. A., Alshurideh, M., Elnagar, A., & Shaalan, K. (2020). Machine learning and deep learning techniques for cybersecurity: A Review. In *The International Conference on Artificial Intelligence and Computer Vision* (pp. 50-57). Springer, Cham.

Shadiev, R., & Sun, A. (2020). Using texts generated by STR and CAT to facilitate student comprehension of lecture content in a foreign language. *Journal of Computing in Higher Education, 32*(2020), 561–581. https://doi.org/10.1007/s12528-019-09246-7

Shadiev, R., & Huang, Y. M (2020). Investigating student attention, meditation, cognitive load, and satisfaction during lectures in a foreign language supported by speech-enabled language translation. *Computer Assisted Language Learning, 33*(3), 301-326.

Shadiev, R., Sun, A., & Huang, Y. M. (2019). A Study of the facilitation of cross-cultural understanding and intercultural sensitivity using speech-enabled language translation technology. *British Journal Educational Technology, 50*, 1415-1433.

Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability, 12*, 524. https://doi.org/10.3390/su12020524

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. doi.org/10.1017/S0958344020000038

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, *19*(2017), 221-248.

Steiner, D. F., Chen, P. H. C., & Mermel, C. H. (2021). Closing the translation gap: AI applications in digital pathology. *Biochimica et Biophysica Acta: Reviews on Cancer, 1875*(1), 188452. https://doi.org/10.1016/j.bbcan.2020.188452

Tanaka, M. (2017) Examining EFL vocabulary learning motivation in a demotivating learning environment. *System, 65*(2017), 130–138.

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

Vincent, J. (2019, July 11). *Google's live camera translation is getting better AI and 60 new languages.* https://www.theverge.com/2019/7/10/20688682/google-translate-camera-instant-translation-function-languages-update-ai

Walker, A., & White, G. (2013). *Technology enhanced language learning: Connecting theory and practice.* Oxford University Press.

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, *84*(6), 2112-2130.

Wu, T. T. (2014). The Use of a mobile assistant learning system for health education based on project-based learning. *Computers, Informatics, Nursing*, *32*(10), 497-503.

Wu, T. T. (2016). English reading e-book system integrating grouping and guided reading mechanisms based on the analysis of learning portfolios. *Journal of Internet Technology, 17*(2), 231–241.

Wu, T. T., & Huang, Y. M. (2017). A Mobile game-based English vocabulary practice system based on portfolio analysis. *Journal of Educational Technology & Society*, *20*(2), 265-277.

Xie, C., Tan, M., Gong, B., Wang, J., Yuille, A. L., & Le, Q. V. (2020). Adversarial examples improve image recognition. In L. O'Conner & H. Torres (eds.), *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 819-828). The Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/cvpr42600.2020.00090

Xie, H., Hwang, G. J., & Wong, T. L. (2021). From conventional AI to modern AI in education: Re-examining AI and analytic techniques for teaching and learning. *Educational Technology & Society, 24*(3), 85–88.

Xu, Q. (2020). Applying MALL to an EFL Listening and Speaking Course: An Action Research Approach. *Turkish Online Journal of Educational Technology-TOJET*, *19*(4), 24-34.

Yigitcanlar, T., Desouza, K. C., Butler, L., & Roozkhosh, F. (2020). Contributions and risks of artificial intelligence (AI) in building smarter cities: insights from a systematic review of the literature. *Energies, 13*(6) (2020), 1473-1511.

Yousef, M., Hussain, K. F., & Mohammed, U. S. (2020). Accurate, data-efficient, unconstrained text recognition with convolutional neural networks. *Pattern Recognition, 108*(2020),107482. doi.org/10.1016/j.patcog.2020.107482.

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The State of the art and future prospects. *Journal of Industrial Information Integration, 23,* 100224. https://doi.org/10.1016/j.jii.2021.100224

Zhong, Y., Ma, A., Ong, S. Y., Zhu, Z., & Zhang, L. (2018). Computational intelligence in optical remote sensing image processing. *Applied Soft Computing*, *64*(2018), 75-93.