Chiang, Y. V., Cheng, Y.-W., & Chen, N.-S. (2023). Improving Language Learning Activity Design through Identifying Learning Difficulties in a Platform Using Educational Robots and IoT-based Tangible Objects. *Educational Technology & Society*, *26*(3), 84-100. https://doi.org/10.30191/ETS.202307_26(3).0007

Improving Language Learning Activity Design through Identifying Learning Difficulties in a Platform Using Educational Robots and IoT-based Tangible Objects

Yueh-hui Vanessa Chiang[1], Ya-Wen Cheng[2] and Nian-Shing Chen[3*]
[1]National Chengchi University, Taiwan // [2]Asia University, Taiwan // [3]National Taiwan Normal University, Taiwan // vanessa.chiang@gmail.com // yawencheng@asia.edu.tw // nianshing@gmail.com
[*]Corresponding author

ABSTRACT: Understanding the obstacles and causes students faced while learning with new technologies is the key to inform effective instructional designs. To achieve this aim, this study conducted a qualitative video analysis on language learners' observable behaviors when they took part in learning activities supported by the technology of robots and IoT-based tangible objects. Insightful findings and instructional implications emerge from the attempt to explore learners' learning process in terms of the obstacles learners encountered and the causes of the obstacles. Based on the findings and implications, eight instructional guidelines are proposed for teachers/instructional designers to design effective language learning activities with robots and IoT-based tangible objects. This study contributes to the literature on enhancing learning and teaching by integrating educational robots and IoT-based tangible objects in the field of robot assisted language learning (RALL).

Keywords: Contextualized multimodal language learning, Robot-assisted language learning, IoT-based tangible objects, Qualitative video analysis, Instructional design guidelines

1. Introduction

The rapid innovation and improvement of information technology expands its application to facilitating learning and teaching in various educational domains. One of the major application areas is English as a foreign language (EFL) learning (Cheng et al., 2018). In traditional EFL learning settings, learners usually face two main challenges to master English. The first one is lacking learning contexts that learners can have sufficient exposure to English in daily life, particularly for advancing their listening skills. The second one is lacking opportunities that learners can apply English to real-life situations for enhancing meaning making in English learning process (Chamot, 2004; DeKeyser, 2005; Seedhouse, 2017). As an endeavor to overcome such challenges, this study was grounded in an EFL learning environment integrating humanoid robots and Internet of Things (IoT) based tangible objects as instructional technologies to enhance immersion and engagement for EFL learners at elementary school level. IoT technologies refer to real-world tangible objects, which are embedded with digital sensors for collecting relevant data and connected with other objects via network technologies like intranet or internet for storing, exchanging, and analyzing data to fulfill specific purposes (Madakam et al., 2015).

Humanoid robot can be used to foster active learning (Kim & Baylor, 2006) by playing a role in language learning activities, socially interacting with learners via its utterances, facial expressions and gestures to facilitate learners' learning process. In this study, such social interaction with robots intended to engage young language learners to repetitively practice English listening skills. The inclusion of tangible objects embedded with IoT sensors can serve as embodied learning materials related to learning contexts to increase the degree of modalities that supports situated learning (Hung et al., 2015; Pasfield-Neofitou et al., 2015). In a language learning environment supported by technology without embodiment, learners may just sit on chairs in front of computer monitors to interact with two-dimensional digital learning materials. This study incorporating tangible objects supported by IoT technologies into the learning environment attempted to expand two-dimensional digital learning materials to three-dimensional cyber-physical mixed learning contents. Such learning environment affords language learners to interact with tangible learning materials via various body movements in learning process, which supports embodied cognition that has been regarded as beneficial for helping learning retention and meaning making in learning contexts requiring real-life application scenarios (Shilling, 2017). From the perspective of cognitive science, learning involves the process of cognition. The viewpoint of embodied cognition emphasizes the mutual influences between brain and body in the cognitive process. Learning activities using technologies to elicit learners' body movements can facilitate learning in terms of enriching mental representations and enhancing retention of learned information (Chao et al., 2013). Thus, this study intended to establish a contextualized multimodal language learning environment supporting both social interaction and embodied cognition with a novel integration of robots and IoT-based tangible objects.

Teachers are experts in the subjects they teach. Most teachers are also good at using technology with which they are already familiar to deliver instruction. However, when it comes to integrating new technology into learning and teaching, educators have noticed a challenge that not every teacher is always ready to incorporate state-of-the-art instructional technology into learning effectively and efficiently (Becker, 1994; Cuban, 2001; Hadley & Sheingold, 1993; Tondeur et al., 2013). Ertmer and Ottenbreit-Leftwich (2013) suggested that the integration of educational technology should pay attention to the instructional aspect to facilitate learning and teaching for students and teachers, rather than technological deployment merely. Thus, supports for helping teachers integrate new technology to fulfil their instructional goals are needed (Batane & Ngwako, 2017). A set of instructional guidelines may serve the purpose. Moreover, Wells (2001) suggested that learning process making up the learning experience is likely to provide further insights for enhancing learning. As such, the objective of this study is to provide instructional guidelines based on analyzing students' observable behaviors when they participated in the language learning activities supported by the learning environment with robots and IoT-based tangible objects.

The novel integration of emerging technologies may arise new instructional concerns to be discovered. Video recording students' learning process makes the attempt of gaining insights from the course of learning feasible (Janík et al., 2009). Identifying the obstacles students encountered initiates our video analyses from where students got stuck in their learning process. Moreover, exploring the causes of the obstacles can inform teachers how to support students' language learning in such learning context. Thus, this study aims at addressing the following three research questions: (1) what obstacles did language learners encounter when participating in learning activities with robots and IoT-based tangible objects? (2) what causes led to the obstacles? (3) what can the identified obstacles and causes inform the design of language learning activities with robots and IoT-based tangible objects?

2. Literature review

2.1. Recent relevant studies on educational robots and IoT-based tangible objects

As emerging technologies, robots and IoT have drawn different levels of attention from researchers in education. Prior studies on robots in education are relatively abundant and considered favorable for supporting young learners' language learning in terms of vocabulary learning, language production, listening skills, and oral interaction, as well as reducing learning anxiety (Alemi et al., 2014; Alemi et al., 2015; Cheng et al., 2021; Cheng et al., 2018; Hung et al., 2015; Lin et al., 2022; Pachidis et al., 2019). Prior studies on using tangible objects with IoT technologies in education are relatively few (Domínguez & Ochoa, 2017; Seedhouse, 2017; Skulmowski et al., 2016; Xie et al., 2008). Noteworthy, Seedhouse and his colleagues (2017) conducted a series of studies using IoT-based tangible objects in kitchen with a task-based pedagogical approach to support foreign language learning and reported such environment can be used to facilitate listening skills and vocabulary learning (Pallotti et al., 2017; Roos et al., 2017). Even though the application of IoT technologies is not as popular as educational robots yet, researchers start to recognize its potential for constructing a personalized smart learning environment (Khaddage et al., 2016; Gul et al., 2017; Ramlowat & Pattanayak, 2019). Given that current applications of educational robots and IoT-based tangible objects as learning technologies are on different stages, few studies integrated these two technologies into the same learning contexts (Tanaka et al., 2007). Our study integrating robots and IoT tangible objects served as a pioneering attempt to establish a contextualized multimodal language learning environment leveraging the advantages of robots, IoT, and tangible objects.

Among recent works of using robots and/or IoT objects in educational contexts, many of them supported that using these technologies can attract learners' attention and their presence is acceptable to younger learners (Chen et al., 2011; Hyun et al., 2010; Kanda et al., 2004; Kanda et al., 2007; Kozima et al., 2009; Leite et al., 2009; Leite et al., 2013 Tanaka et al., 2007). However, most of the prior studies tended to conduct experimental inquiries and regarded the use of educational robots and/or IoT objects as treatments in research design to examine the effects of the technologies on learning outcome via summative measurements at a certain time point. Few of them investigated the use of educational robots and/or IoT objects as learning technologies from the formative aspect throughout the process of learning activities (Benitti, 2012; Karim et al., 2015; Li, 2015; Liu et al., 2016; Mubin et al., 2013; Toh et al., 2016; Spolaôr & Benitti, 2017). Therefore, this study attempted to bridge the gap by focusing on analyzing language learners' observable behaviors throughout their learning process.

2.2. The importance of supporting teachers integrating new technology into learning and teaching

Applying emerging technology to learning and teaching may create more opportunities of educational innovation and improvement. However, as Ertmer and Ottenbreit-Leftwich (2013) suggested that the integration of educational technology should pay particular attention to the instructional aspect of how teachers can facilitate learning and teaching with such technology, instead of merely to the development of technology. Simply using new technology in learning activities does not necessarily enhance learning. Jonassen (1996) stated teachers should focus on designing learning activities incorporating technology in facilitating students' learning process. By doing so, the benefits of using technology can be extended to carry out effective and meaningful learning. In order to address this concern, educational researchers proposed that support for teachers to enable their adoption of new technology in learning process is needed especially when more technology is developed and promising for being used in learning contexts (Batane & Ngwako, 2017; Cheng et al., 2021; Ertmer & Ottenbreit-Leftwich, 2013). Additionally, Howland et al. (2012) emphasized helping teachers adopt new teaching practices afforded by new technology should focus on the change of learning process.

This study introducing a novel integration of emerging technology, educational robots and IoT-based tangible objects, may raise new instructional concerns when teachers would like to design effective learning activities with such technology. Responding to suggestions from existing literature, this study attempted to provides guidelines for teachers to design language learning activities with educational robots and IoT-based tangible objects for smoothing the process of adopting new educational technology.

In order to achieve this objective, the guidelines particularly derive from analyzing language learners' observable behaviors when they participated in learning activities with educational robots and IoT-based tangible objects. Analyzing learners' behaviors to identify what may hinder or assist learning process can provide insights for how to engage learners in learning by avoiding hindrance or rendering assistance (Vygotsky & Cole, 1978; Wells, 2001). Accordingly, the analyses started from identifying the obstacles that learners encountered when they participated in the language learning activities involving educational robots and IoT-based tangible objects. Subsequently, identifying the causes of the encountered obstacles may inspire what needs to be taken into consideration when teachers want to redress the obstacles.

3. The design of learning activities

3.1. The robot and IoT-based tangible objects language learning environment

This study was conducted in a technology-enhanced language learning environment that integrates robots, tablets, and tangible objects with IoT sensors and is driven by scripts composed by instructional designers with a visual programming tool. The learning environment was constituted by the frontend and backend system components. The frontend system components refer to the devices, such as robots, tablets, and tangible objects, with which learners actually interact. The backend component called script editor is a visual programming platform on which teachers/instructional designers work for designing learning activity. How the frontend elements interact with learners depends on the scripts composed with the backend script editor.

The script editor is a visual programming website adapted from Google Blocky (see Figure 1). Teachers/instructional designers compose scripts with this visual programming tool to design learning activities involving the frontend system components and to manage the learning materials used in the learning activities. The scripts created by teachers/instructional designers are saved on a cloud server beforehand and deployed to the frontend components to steer what and how they perform when the corresponding learning activities take place.

The frontend system components include a robot, a tablet, and a set of tangible objects embedded with IoT sensors as shown in Figure 2. The role the robot assumes may vary according to teachers'/instructional designers' decisions to carry out their instructional design. Also, an additional tablet, dedicated to displaying multimedia materials is accompanied with the robot to increase the richness of information presented to learners. The IoT-based tangible objects used for this study is a set of supermarket toys, including a supermarket shelf with many tangible food or stationery models (see Figure 2). In addition to the multimedia materials displayed on the tablet, the tangible toys are three-dimensional learning materials that support embodiment for learners in this multimodal learning environment.

Figure 1. The backend script editor

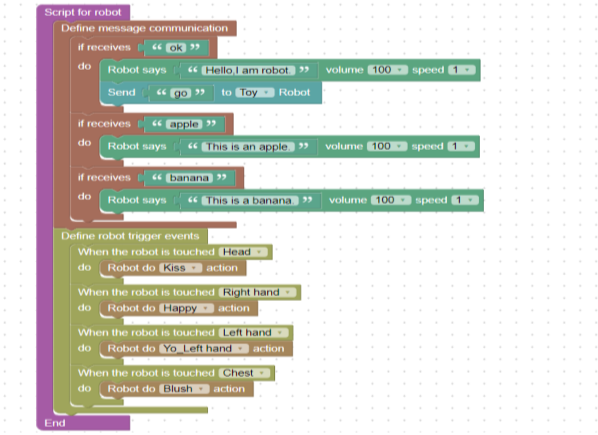


Figure 2. The frontend system components

Figure 3. A toy object with a NFC tag attached

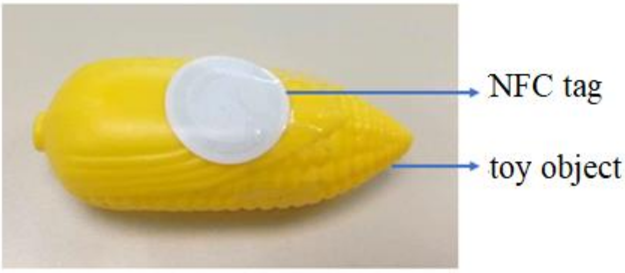
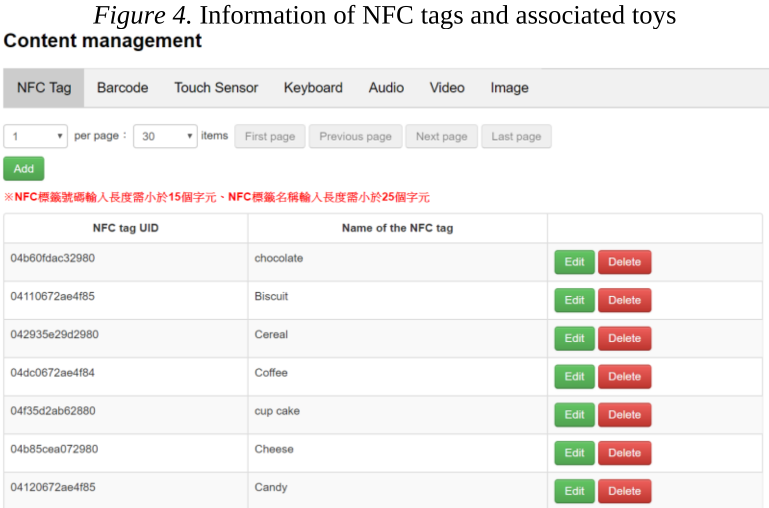


Figure 4. Information of NFC tags and associated toys



Content management

NFC Tag | Barcode | Touch Sensor | Keyboard | Audio | Video | Image

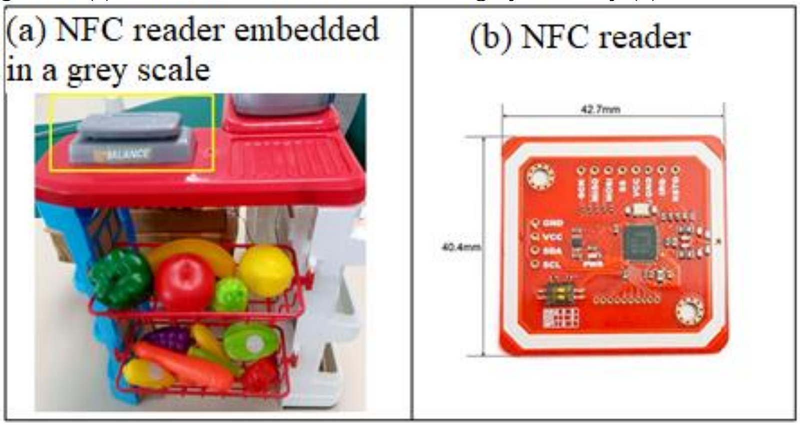
1 per page： 30 items First page Previous page Next page Last page

Add

※NFC標籤號碼輸入長度需小於15個字元、NFC標籤名稱輸入長度需小於25個字元

NFC tag UID	Name of the NFC tag	
04b60fdac32980	chocolate	Edit Delete
04110672ae4f85	Biscuit	Edit Delete
042935e29d2980	Cereal	Edit Delete
04dc0672ae4f84	Coffee	Edit Delete
04f35d2ab62880	cup cake	Edit Delete
04b85cea072980	Cheese	Edit Delete
04120672ae4f85	Candy	Edit Delete

Figure 5. (a) An NFC reader embedded into a grey scale toy (b) An NFC reader

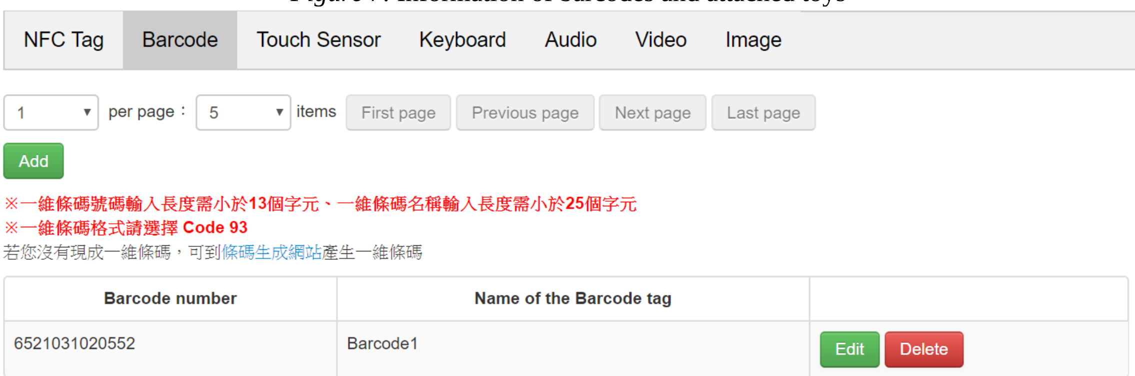


The supermarket toys are adapted to accommodate two types of sensors. First, Near Field Communication (NFC) tags and readers: each toy object is attached with an NFC tag (see Figure 3). The corresponding information of which NFC tag is associated with which tangible toy object is set and kept in the backend script editor (see Figure 4). An NFC reader is embedded into a grey scale toy as shown in Figure 5(a)(b). When learners place a toy object on the grey scale, the robot can tell which object it is and trigger corresponding actions.

Second, one-dimension barcodes and the reader: similar to NFC tags, one-dimension barcodes are stuck on some tangible objects (see Figure 6(a)). Learners can hold a barcode reader to scan the barcodes (see Figure 6(b)). A corresponding table of which barcode represents which toy object is also set in script editor (see Figure 7).

Figure 6. (a) A barcode stuck on a toy object (b) A barcode reader

Figure 7. Information of barcodes and attached toys



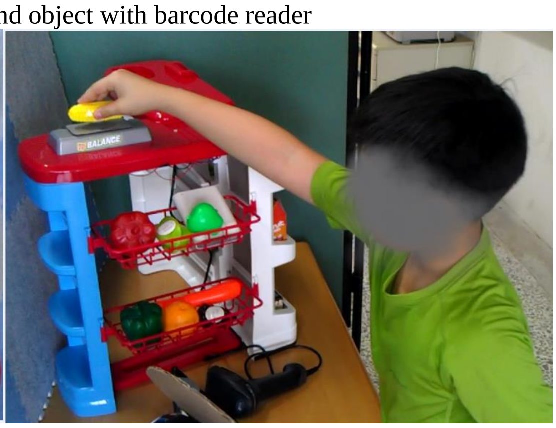
3.2. The learning activities

This study followed the framework suggested by Hubbard (1992) in terms of integrating technology into language learning to design task-based activities, which has been considered suitable for enhance language learning (Pica et al., 2006; Morales, 2017). The language learning activities in this study consisted of four sessions in four weeks. Each session included several tasks for learners to complete one by one. The learning objective of this series of learning activities was to improve learners' English listening skills. Responding to Hubbard's (2017) suggestion of using emerging technologies to establish a three-dimensional immersive environment with digital sensors for learning language listening, this study used robots and IoT technologies for this purpose. With such foundation, most of the instruction and hints of the tasks were intendedly presented by the robot in the audio format, supplemented with some visual materials on the tablet. The vocabularies and conversations used in the learning activities were designed in alignment with the content of the fourth graders' English textbooks in Taiwan and confirmed by the participants' English teachers before the learning activities took place.

The goal for the learning activity in the first week was to prepare learners for the subsequent activities by helping them get familiar with how to use the two types of sensors (i.e., the NFC and the barcode). The goal for the week 2 learning activity was to help learners practice their listening skills of vegetables, fruits, and food. The learning activity of the third week attended to cover the vocabularies of stationery, colors, clothes and drinks in addition to fruits. The goal of week 4 learning activity was to help learners practice their listening skills of colors. Several tangible objects embedded with sensors were used in the learning activities as learning materials, representing the learning content. Such learning activities were not designed to replace the existing English class sessions. Instead, before the learning activities took place, the participants' English teachers previewed the learning activities and acknowledged the potential for supplementing the existing English classes, enhancing students' learning motivation, and further improving students' English listening skills. In addition, the teachers suggested to incorporate more learning supports, such as oral encouragement and visual hints assisting learners in connecting the vocabularies and the corresponding objects, into the learning activities. Teachers' suggestions were all implemented in the learning activities.

Figure 8. (a) The robot provides instruction of a task (b) The learner places the found object on grey scale (c) The learner scans the found object with barcode reader

(a)



(b)

(c)

The general process occurring in the four learning activities was that the robot started to provide instruction describing the target object to complete a task to learners as shown in Figure 8(a). Then, the learners worked on finding the target object based upon their understanding of the instruction. Learners would present the tangible object they found by either placing it on the grey scale or scanning it with the barcode reader as shown in Figure 8(b) and Figure 8(c). The robot would determine if the presented object is the target one and provide corresponding feedback to the learners to indicate their success or failure. In the case of success, the robot would proceed to give instruction of the next task until all tasks were completed. In the case of failure, the robot would encourage the learners to try again. Upon the completion of all tasks, the robot would render a final congratulation to the learners.

4. Method

From the sociocultural perspective (Vygotsky & Cole, 1978), this study attempted to analyze students' observable behaviors during task performance in the real learning situations supported by the learning environment with robots and IoT-based tangible objects to address the three research questions with an inductive qualitative approach (Lincoln & Guba, 1985).

4.1. Participants

This study took place in an elementary school in central Taiwan. Ten fourth graders participated in this study with their parents'/legal representatives' consent. Participants' age ranged between 9 and 10 years old. Five of them were female; five were male. As shown in Figure 9, a classroom was allotted to be the venue of this study. Sound-proofing panels were used to set up 10 cubicles. Participants mainly worked within their assigned cubicles during the learning activities. In each cubicle, a robot with a tablet and a set of supermarket toys embedded with IoT sensors and readers were equipped. Participants attended one learning session, lasting 20-30 minutes, per week for four consecutive weeks in 2019.

Figure 9. Research site

4.2. Data sources and analysis

The primary data source of this study was the video recordings of participants' learning process. To collect the primary data source, each cubicle was equipped with a camera to record what happened in the cubicle during the learning session. The observational field notes served as a secondary data source.

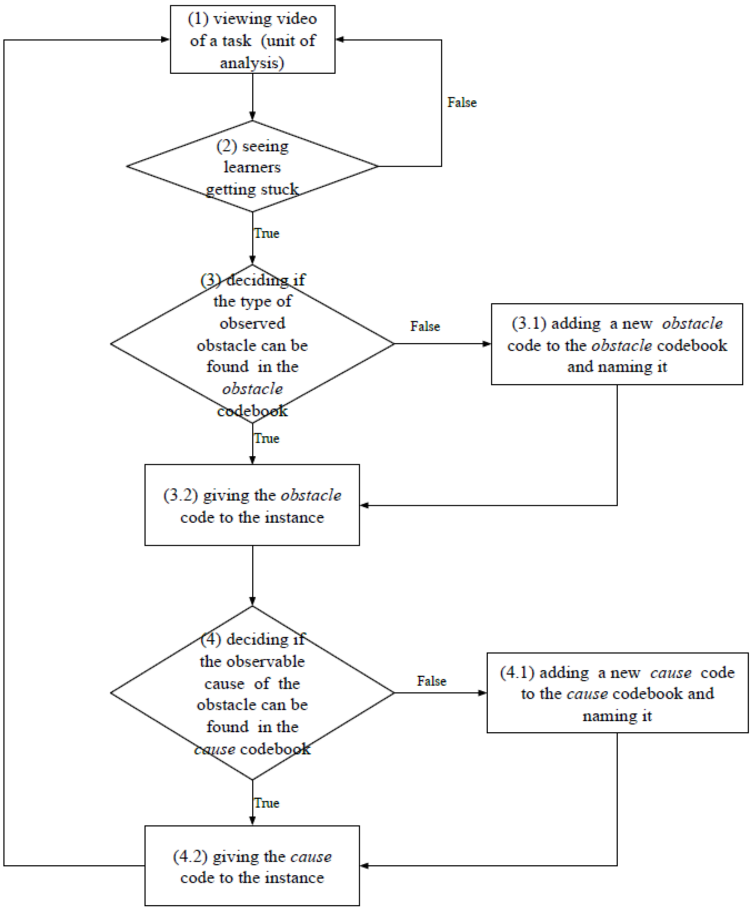
The technique of video analysis was used to analyze the data. All video clips were viewed for the first round to select valid clips to proceed to the next round of analysis. The criteria for selecting valid clips included: (a) the visibility of the participants in the video clips, (b) the audio quality of the video clips, and (c) the visibility of the objects used in the learning activity. The units of analysis for the subsequent analysis were the tasks that the participants were assigned to complete in every learning activity. A total of 27 valid video clips were selected and 135 tasks were analyzed to identify obstacles and causes to address the first two research questions.

The open coding technique for qualitative inquiries (Strauss & Corbin, 1990) was adopted to generate the codebooks in terms of *obstacles* and *causes*. Figure 10 presents the procedure of identifying *obstacles* and *causes* from video recordings in the following steps:

- Step (1): the researchers viewed the video clip of a task.
- Step (2): when seeing the participants getting stuck, the researchers regarded such instance as an obstacle and proceeded to the next step, otherwise, continuing to viewing other video clips.
- Step (3): deciding if the observed obstacle belonged to an existing type of obstacle in the *obstacle* codebook. If not, the researchers proceeded to step (3.1), otherwise, proceeding to step (3.2).
  - Step (3.1): adding a new *obstacle* code which can fit the new instance to the *obstacle* codebook, labelling the new code, and continuing to step (3.2).
  - Step (3.2): giving the *obstacle* code to the instance.
- Step (4): deciding if the observable cause of the obstacle belonged to an existing type of cause in the *cause* codebook. If not, the researchers proceeded to step (4.1), otherwise, proceeding to step (4.2).
  - Step (4.1): adding a new *cause* code which can fit the new instance to the *cause* codebook, labelling the new code, and continuing to step (4.2).
  - Step (4.2): giving the *cause* code to the instance.

Then, the researchers repeated the step (1) to (4) for all selected video clips. As data analysis proceeded, a set of codes of *obstacles* and *causes* emerged and were continuously compared with new instances to see if modifications were needed. Two researchers followed the procedure to code the data separately first and then reach their consensus in terms of their coding results. The Cohen's kappa coefficient of *obstacle* coding was 0.929 and the coefficient of *cause* coding was 0.873.

Figure 10. The procedure of identifying obstacles and causes from video clips



Before answering the research questions, the researchers confirmed the effectiveness of the learning activities with robots and IoT-based tangible objects on learners' learning outcome measured by pre-test and post-test of learners' English listening skills. A Shapiro-Wilk test of normality indicated pre-test (p = .000) and post-test (p = .008) scores were not normally distributed. Consequently, a Wilcoxon Signed-Ranks test indicated the median of post-test scores, 82.5, was significantly higher than the median of pre-test scores, 80 (Z = -2.000, p = .046 < .05). We, therefore, concluded there was a significant positive effect of the learning activities with robots and IoT-tangible objects on participants' English listening test scores. With confirming the effectiveness on learning outcome, the detailed analysis of learners' learning process aimed for improving the instructional design, making learning experience even better.

5. Results

Among the 135 tasks, 84 instances were coded with obstacles during participants' learning process. The results addressing the first two research questions are reported in this section. Five categories of obstacles that prevented learners from completing the tasks and 7 categories of causes leading to the obstacles emerged from data analysis. Table 1 summarizes the identified categories of obstacles and causes along with definitions and examples. The first and second columns show the names of obstacle categories and the features that were observed in the video recordings of the corresponding obstacles. The third and fourth columns present the observable causes of the corresponding obstacles and their features seen in the video clips. The fifth column provides the examples of the obstacles and causes extracted from the data. The naming convention of referring to a specific raw data of video follows the format of [S##-W#-T#]; where S## indicates the participant ID, W# indicates the number of the week, and T# presents the number of the task in each week's learning activity. For example, [S05-W4-T2] indicates the case of participant S05's second task in week 4.

The first four categories were ordered by the sequence of which they could happen in a given task. At the beginning of each task, when the task instruction was provided to learners mainly by the robot in audio mode, the obstacles of failing to understand the assigned task (O1) and missing the instruction of task (O2) and their causes could occur. After the instructions of tasks were given, participants started to find the target objects according to their understanding of the instructions and the obstacle of failing to find the target toy object (O3) could happen. After participants found the objects to complete the tasks, the obstacle of failing to present the found toy objects (O4) could occur. As to the fifth obstacle of being interfered by contextual factors (O5), it could happen throughout the whole process of completing the tasks.

Table 1. The identified obstacles and causes

Obstacles	Definition of obstacles	Causes	Definition of causes	Examples
O1. Fail to understand the assigned task	(a) The participants were present in the cubicle as the robot rendered the instruction of the assigned task, and (b) the participants seemed not completely understand what the assigned task was.	C1. Learners' English ability	Even though the participants seemed to listen to the instruction, their subsequent behaviors observed in the video recordings indicated that they did not exactly comprehend the instruction and did not know for which target object they were looking.	Participant S05 worked on the task of finding and scanning the orange fingerprint with the barcode reader. After the robot provided the instruction of collecting an orange fingerprint, he tried to place the fruit orange on the grey scale twice and failed to complete the task. Then, he tried many different toy objects until another participant came to his cubicle pointing out where the orange fingerprint was [S05-W4-T2].
		C2. Misconceiving the instruction	The participants seemed to misconceive the visual/audio feedback or hints in the learning activities.	After participant S07 completed the task of placing the starfruit on the grey scale, audio and visual feedback both indicated her completion of the task. The visual feedback left a picture of starfruit on the tablet screen. Then the robot started to provide the audio instruction of the next task, placing a pencil on the grey scale. While participant S07

				seemed to listen to the instruction, she looked at the tablet screen as well. Her first attempt was to place the starfruit on the grey scale again and failed to complete the task [S07-W3-T2].
O2. Miss the instruction of task	The participants missed the instruction of task so that they did not know what the assigned task was.	C3. Not attended to the instruction	(a) The participants were not physically present in the cubicle as the robot rendered the instruction of the assigned task. (b) The participants were physically present in their cubicles, but they were seen that their attention was distracted by other things.	(a) Participant S08 was absent from her cubicle when the robot was giving instruction. After she returned to her cubicle, she looked puzzled at what was going on until one researcher came to give her some hints on the target object of starfruit [S08-W3-T1]. (b) Participant S03 looked excited after he completed the first task. As the robot was giving the instruction of the second task, he was still immersed in the excitement of collecting the green fingerprint by holding and looking at the green pepper with the green fingerprint and did not pay attention to the new instruction [S03-W4-T2].
O3. Fail to find the target toy object	When the participants seemed to know what they were looking for but cannot find the target among the toys	C4. Unable to see the target object	The target object was placed in a location which the participants cannot reach.	Participant S10 attempted to find the orange fingerprint to complete a task. However, the orange fingerprint appeared on the highest shelf and was covered by another toy object. Participant S10 was not tall enough to reach the highest shelf. She kept looking for the orange fingerprint until another participant who was taller than participant S10 help remove the object covering the orange fingerprint [S10-W4-T2].
		C5. Unable to recognize the target object	The participants failed to identify the right toy object representing the English vocabulary of target object.	Participant S07 worked on placing the green apple on the grey scale to complete a task. She tried many times with green pepper without recognizing that the toy object she held was not green apple [S07-W2-T3].
O4. Fail to present the found toy object	Participants had found the target object to complete the assigned task but failed to present it via input devices	C6. Flawed operation of input devices	(a) The participant presented the object with wrong input device. (b) Ineffective operation of the	(a) Participant S08 would like to complete the task of scanning the barcode on a jar of milk. However, instead of using the barcode reader to scan the barcode on the milk

			input device	jar, she placed the jar of milk on the grey scale [S08-W3-T5]. (b) The participants used the right input device to present the found object; however, their operations of the input devices failed to take effect. They needed to repeat the operation to complete their tasks.
O5. Interfered by contextual factors	Even though participants stayed with the robot as it was giving instruction of the assigned task, some contextual issues prevented them from proceeding with the task	C7. Surrounding noises	Two sources of noises from the surroundings: (a) the noise from the school-wide environment, and (b) the noise from the research context, i.e., the participants	(a) As participant S10 was listening to the task instruction, a school teacher made an announcement via broadcasting system. The volume of the broadcast was too loud for participant S10 to hear the instruction clearly. In the video recording, she was observed to incline forward to the robot to hear the instruction. After the broadcast, the robot had finished the instruction [S10-W2-T5]. (b) The cubicle in the research site was not a completely closed area. If the participants talked too loudly, their voices can be heard by all participants in the classroom.

Table 2. The counts of identified obstacles and causes

Obstacles	Causes	W1	W2	W3	W4	Sub total	Total
O1. Fail to understand the assigned task	C1. Learners' English ability	1	8	19	6	34	41
	C2. Misconceiving the instruction	0	0	3	4	7	
O2. Miss the instruction of task	C3. Not attended to the instruction	1	2	6	4	13	13
O3. Fail to find the target toy object	C4. Unable to see the target object	1	0	0	18	19	21
	C5. Unable to recognize the target object	1	1	0	0	2	
O4. Fail to present the found toy object	C6. Flawed operation of input devices	2	1	3	0	6	6
O5. Interfered by contextual factors	C7. Surrounding noises	0	2	1	0	3	3
Total		6	14	32	32	-	84

Table 2 reports the numbers of instances per week of each category. Among the instances coded with obstacle of failing to understand the assigned task (O1), they were mainly caused by learners' English ability (C1) (34 out of 41). The highest occurrences of obstacles O1 caused by C1 was in the third week (i.e., 19). This finding suggested that the learning content covered in week 3 posed the most challenges to the participants' existing English ability. The target words to complete the tasks in week 3 were from the most various categories, including fruits, stationery, clothes, drinks and colors. The participants were less likely to rely on the words completing the preceding tasks as contextual clues to infer the target words of the subsequent tasks. According to Vygotsky's (1962) notion of zone of proximal development, we regarded these challenges as learning opportunities for language learners to expand their English ability. Being seen in the data revealing learners' task performance, the participants eventually overcame the challenges with supports provided by the learning environment. The findings provided us with insights in terms of how to assist learners to overcome challenges in their language learning activities so as to expand their English ability, leading to the instructional guidelines proposed in the next section.

The second highest frequency of obstacles was failing to find the target toy object (O3) because learners cannot see the object (C4) in week 4 (i.e., 18). By reviewing these instances, they were related to the multimodal features provided by the learning environment. As the example of this category shown in Table 1, the learning activity in week 4 asked participants to collect five fingerprints. In the research context, the fingerprint stickers were so thin as tiny pieces of paper that they can be hidden in the places where some participants were unable to reach like the orange fingerprint in the case of [S10-W4-T2]. Also, the 7 instances of obstacle of failing to understand the assigned task (O1) due to learners' misconception of the instruction (C2) shed light on the similar issue of multimodality. As the case of [S07-W3-T2] listed in Table 1 exemplifying this category, when providing feedback to indicate learners' success of completing the task, the instructional designer added visual feedback along with the audio one. In this case, the participant S07 seemed to misperceive the visual feedback of the task she just completed as the instruction of the next task. Although listening to the audio instruction of the next task, she misconceived the picture left from the feedback of the previous task as part of the instruction of the new task, which prevented her from fully comprehending the instruction. These instances informed us in designing multimodal English learning activities, and the guidelines will be presented in the next section.

As to the obstacle of missing the instruction of task (O2) because learners did not attend to the instruction (C3), the instances of this category happened in all 4 weeks. It suggested young language learners may be easily distracted when they participated in learning activities. In this study, the robots provided a function of repeating the instruction of tasks to learners. This function was triggered by touching robots' belly and intended to help learners overcome this obstacle whenever needed. Another way to address this obstacle would be allowing learners to control the flow of the activities. When learners were ready to proceed to the next task, they were in charge of letting the robot know that they would like to listen to the instruction of the next task. The instances of the obstacle of failing to find the target toy object (O3) because learners were unable to recognize the object (C5) occurred only in the first two weeks. This suggested that learners may need some time to get familiar with the tangible objects and associate them with the corresponding English vocabularies. The cases coded with the obstacle of failing to present the found toy object (O4) due to flawed operation of input devices happened in the first three weeks and none in the fourth week. The learning activities of the first three weeks required learners to use either the NFC reader or barcode reader as input devices. However, for the learning activities in the fourth week only required learners to use the barcode reader as the input device. This finding suggested when an IoT-based learning environment allowed learners to use more than one input devices for interaction, learners may need assistance in using the right input device effectively. As for the obstacle of being interfered by surrounding noises (C7 of O5), this can interrupt English learners' learning process, especially when the learning objective was to practice listening ability. This issue may be resolved by equipping headset microphones as sound input/output devices to learners. In addition, the aforementioned functionality of allowing learners to ask the robots to repeat the task instruction when needed can serve as another solution.

6. Instructional guidelines for designing learning activities using robot and IoT-based tangible objects

To address the third research question, the findings of the first two research questions were synthesized to derive instructional guidelines for teachers/instructional designers who may consider designing language learning activities involving robots and IoT-based tangible objects.

During learning process, some challenges with appropriate scaffoldings may be beneficial to expand learners' understanding of the subject matters (Vygotsky, 1962; Wood et al.,1976; Nikolaevskaya, 2017). Hence, the obstacles of failing to understand the assigned task caused by learners' English ability (C1 of O1) may particularly indicate the opportunities of advancing learners' learning and inform the instructional designers how to facilitate learners' learning process by providing scaffoldings when applicable. In order to make this occurs, four guidelines were proposed as follows:

Guideline 1: Teachers/instructional designers should use robots to confirm learners' understanding of the assigned task in the early stages of language learning activities and making applicable rectification in time to help learners stay on the right track, avoiding detours to ineffective learning.

Guideline 2: When tangible objects are used to represent the new learning content, teachers/instructional designers should use robots to strengthen learners' comprehension of learning content by providing assistance in associating new content with the corresponding tangible objects in the language learning context.

Guideline 3: Teachers/instructional designers should avoid confounding new learning content with multiple meanings and use lexical items that are suitable for learners' age in terms of word frequency and familiarity.

Guideline 4: Teachers/instructional designers should use robots to provide functionality that allows learners to control the progress of the learning activities, such as learners can ask robots to repeat task instruction whenever they need, and learners can ask robots to proceed to the subsequent tasks whenever they are ready.

Speaking of facilitating learning, designing some sort of functionality enabling the learning environment to keep track of learners' learning process can be beneficial to trigger applicable assistance. For example, with a function to monitor the count of failure to complete the tasks, when it passes a certain amount, robots can automatically provide more guidance to learners. Moreover, with a function to keep track of time span between learners' attempts to complete the tasks, if it is longer than a certain period of time, robots could start to provide assistance as well. These scaffoldings are appealing for facilitating learning process and constructing an adaptive learning environment that can offer multiple assistance in accordance with learners' learning states, leading to the next guideline.

Guideline 5: Teachers/instructional designers should make robots' behaviors responsive to learners' learning states and offer multiple types of assistance available to language learners.

In addition, the obstacle of failing to present the found toy object due to the flawed operation of input devices (C6 of O4) was mainly related to the familiarity of using new technologies. It had a little to do with learners' English ability but can definitely hinder learners' learning process. Whenever introducing new technologies to learning and teaching, teachers/instructional designers should prepare learners for necessary operations of the new technologies in advance; otherwise, it is possible to embrace emerging technology at the cost of learners' learning effectiveness. In this study, the instructional designer had designed the learning activity in week 1 to help learners get familiar with the new input devices. However, some cases of this category of obstacles still occurred. More cases would be anticipated if not implementing this activity. Thus, the sixth guideline was provided:

Guideline 6: Teachers/instructional designers should make sure that the pre-task activities preparing learners' operation of newly adopted technologies and familiarity with related vocabularies are well designed and thoroughly carried out so that learners can attend to the language learning tasks.

Moreover, the obstacle of failing to find the target toy object caused by being unable to see the target object (C4 of O3) occurred the most in week 4 is noteworthy. A multimodal language learning environment affording textual, audio, spatial, and visual modes of interaction is promising for enabling immersive language learning (Moreno & Mayer, 2007; Kirsh, 2013). Meanwhile, when designing learning activities, as suggested by Skulmowski and colleagues (2016), the instructional designers should consider the appropriateness of multimodal interaction between learners and the learning environment, rather than only thinking about the traditional mono-modal communication. The examples of this obstacle highlight the importance of such consideration. Placing the target objects at a physical location where learners cannot reach makes the objects beyond the scope of learners' normal vision and hinders learners from completing the learning activities. In this regard, neglecting the aspect of three-dimensional spatial modality can even create more obstacles for learning, particular for young children. To avoid this pitfall, a comprehensive consideration including learners' profiles and the features of multiple modalities when designing language learning activities is recommended for instructional designers as below:

Guideline 7: Teachers/instructional designers should take learners' physical attributes into consideration when arranging tangible objects physically.

Furthermore, the obstacle of failing to understand the assigned task caused by misconceiving the instruction (C2 of O1) also relates to multimodality. It sheds light on the concerns about proper sequence of the multimodal materials used in language learning activities and the consistency of messages delivered to learners via multiple modalities. In a language learning environment supporting multimodal materials, confusing arrangement of the materials and inconsistent messages delivered by them at the same time can hold back the advantages of multimodality and even deteriorate the learning process, which results in the eighth guideline:

Guideline 8: When using robots to deliver messages serving the same purpose via different modalities, such as robots' utterances, facial expressions, and gestures, teachers/instructional designers should make sure that these messages are delivered consistently. Likewise, robots should deliver the messages serving different purposes in a way distinct enough for learners to differentiate their references correctly.

7. Conclusion

This study proposed instructional guidelines for teachers/instructional designers to design learning activities involving robots and IoT-based tangible objects for language learning. The guidelines emerged from analyzing learners' learning process to understand what obstacles learners encountered and the causes of the obstacles as learners participated in language learning activities. These attempts may contribute to supporting teachers/instructional designers who are interested in integrating robots and IoT-based tangible objects to design effective language learning activities for learners. By adopting educational robots and IoT-based objects together, this study introduced an innovative English learning environment that can accommodate multimodal pedagogical practices. Such endeavors may open up new opportunities to enhance language learners' exposure to English and to connect what they learned to real-life scenarios.

Even though the qualitative video analysis granted us rich details of learners' learning process, this study has limitation due to the data collection of video recordings. The researchers set up the recording equipment at a location that can cover students' cubicles mostly. However, in the real learning activities, students might walk outside the camera scene. This methodological limitation has been commonly seen in the studies using video recordings as data sources (Janík et al., 2009). To recognize this limitation, the researchers made the findings and suggestions based upon the observable evidences in the video recordings. Another limitation is that this study only included the language learning activities with single learning strategy, task-based learning, in the learning environment with robots and IoT-based toys. The future direction will be conducting more studies with various learning strategies supported by this learning environment to develop more guidelines for teachers/instructional designers. Making the language learning environment using robots and IoT-based tangible objects more applicable and effective to support learners' learning process is the ultimate goal. To make this happen, the researchers recognize the importance of teachers' perspective on using robots and IoT-based tangible objects to design and carry out learning activities. Thus, in addition to proposing a set of instructional guidelines for teachers, the researchers will conduct future studies to understand teachers' perceptions and collect their feedback in terms of using robots and IoT-based tangible objects to enhance students' learning.

Acknowledgement

This research is supported in part by the Ministry of Science and Technology, Taiwan under project numbers MOST 111-2410-H-003 -028 -MY3, MOST-109-2511-H-003-053-MY3, MOST 109-2511-H-004-004, and MOST-108-2511-H-003-061-MY3. This work was financially supported by the "Institute for Research Excellence in Learning Sciences" of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan

References

Alemi, M., Meghdari, A., & Ghazisaedy, M. (2014). Employing humanoid robots for teaching English language in Iranian junior high-schools. *International Journal of Humanoid Robotics*, *11*(03), 1450022. https://doi.org/10.1142/S0219843614500224

Alemi, M., Meghdari, A., & Ghazisaedy, M. (2015). The Impact of social robotics on L2 learners' anxiety and attitude in English vocabulary acquisition. *International Journal of Social Robotics*, *7*(4), 523-535.

Batane, T., & Ngwako, A. (2017). Technology use by pre-service teachers during teaching practice: Are new teachers embracing technology right away in their first teaching experience? *Australasian Journal of Educational Technology*, *33*(1). https://doi.org/10.14742/ajet.2299

Becker, H. J. (1994). How exemplary computer-using teachers differ from other teachers: Implications for realizing the potential of computers in schools. *Journal of Research on Computing in Education*, *26*, 291–321.

Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A Systematic review. *Computers & Education*, *58*(3), 978–988.

Chamot, A. U. (2004). Issues in language learning strategy research and teaching. *Electronic Journal of Foreign Language Teaching*, *1*(1), 14-26.

Chao, K. J., Huang, H. W., Fang, W. C., & Chen, N. S. (2013). Embodied play to learn: Exploring K inect-facilitated memory performance. *British Journal of Educational Technology*, *5*(44), E151-E155. https://doi.org/10.1111/bjet.12018

Chen, N.-S., Quadir, B., & Teng, D. C. (2011). Integrating book, digital content and robot for enhancing elementary school students' learning of English. *Australasian Journal of Educational Technology*, *27*(3). https://doi.org/10.14742/ajet.960

Cheng, Y. W., Sun, P. C., & Chen, N. S. (2018). The Essential applications of educational robot: Requirement analysis from the perspectives of experts, researchers and instructors. *Computers & Education*, *126*, 399-416.

Cheng, Y. W., Wang, Y., Yang, Y. F., Yang, Z. K., & Chen, N. S. (2021). Designing an authoring system of robots and IoT-based toys for EFL teaching and learning. *Computer Assisted Language Learning*, *34*(1-2), 6-34. https://doi.org/10.1080/09588221.2020.1799823

Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Harvard University Press.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A Review of issues. *Language learning*, *55*(S1), 1-25.

Domínguez, F., & Ochoa, X. (2017). Smart objects in education: An Early survey to assess opportunities and challenges. In *4th International Conference on eDemocracy & eGovernment-ICEDEG 2017* (pp. 216-220). IEEE. https://doi.org/10.1109/ICEDEG.2017.7962537

Ertmer, P. A., & Ottenbreit-Leftwich, A. (2013). Removing obstacles to the pedagogical changes required by Jonassen's vision of authentic technology-enabled learning. *Computers & Education*, *64*, 175-182.

Gul, S., Asif, M., Ahmad, S., Yasir, M., Majid, M., & Malik, M. (2017). A Survey on role of internet of things in education. *International Journal of Computer Science and Network Security*, *17*(5), 159-165.

Hadley, M., & Sheingold, K. (1993). Commonalties and distinctive patterns in teachers' integration of computers. *American Journal of Education*, *101*, 261-315.

Howland, J. L., Jonassen, D., & Marra, R. M. (2012). *Meaningful learning with technology* (4th ed.). Pearson.

Hubbard, P. (1992). A Methodological framework for CALL courseware development. In *Computers in Applied Linguistics: An International Perspective* (pp. 39-65). Multilingual Matters.

Hubbard, P. (2017). Technologies for teaching and learning L2 listening. In *The Handbook of Technology and Second Language Teaching and Learning* (pp. 93-106). https://doi.org/10.1002/9781118914069.ch7

Hung, I., Hsu, H., Chen, N., & Kinshuk. (2015). Communicating through body: A Situated embodiment-based strategy with flag semaphore for procedural knowledge construction. *Educational Technology Research and Development*, *63*, 749–769.

Hyun, E., Yoon, H., & Son, S. (2010). Relationships between user experiences and children's perceptions of the education robot. In *Proceeding of the 5th ACM/IEEE international conference on human-robot interaction* (pp. 199–200). ACM.

Janík, T., Seidel, T., & Najvar, P. (2009). Introduction: On the power of video studies in investigating teaching and learning. In T. Janik, & T. Seidel (Eds.), *The Power of Video Studies in Investigating Teaching and Learning in the Classroom* (pp. 7-22). Waxmann.

Jonassen, D. H. (1996). *Computers in the classroom: Mindtools for critical thinking*. Prentice-Hall, Inc.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A Field trial. *Human–Computer Interaction*, *19*(1), 61-84.

Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A Two-month field trial in an elementary school for long-term human-robot interaction. *IEEE Transactions on Robotics*, *23*(5), 962-971.

Karim, M. E., Lemaignan, S., & Mondada, F. (2015). A Review: Can robots reshape K-12 STEM education? In *2015 IEEE international workshop on advanced robotics and its social impacts (ARSO)*. IEEE. https://doi.org/10.1109/ARSO.2015.7428217

Khaddage, F., Müller, W., & Flintoff, K. (2016). Advancing mobile learning in formal and informal settings via mobile app technology: Where to from here, and how? *Educational Technology & Society*, *19*(3), 16-26.

Kim, Y., & Baylor, A. L. (2006). A Social-cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development*, *54*, 569–596.

Kirsh, D. (2013). Embodied cognition and the magical future of interaction design. *ACM Transactions on Computer-Human Interaction*, *20*(1),1-30. https://doi.org/10.1145/2442106.2442109

Kozima, H., Michalowski, M., & Nakagawa, C. (2009). A Playful robot for research, therapy, and entertainment. *International Journal of Social Robotics*, *1*, 3-18

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: A Survey. *International Journal of Social Robotics*, *5*(2), 291-308.

Leite, I., Martinho, C., Pereira, A., & Paiva, A. (2009). As time goes by: Long-term evaluation of social presence in robotic companions. In *RO-MAN 2009-the 18th IEEE international symposium on robot and human interactive communication* (pp. 669-674). IEEE. https://doi.org/10.1109/ROMAN.2009.5326256

Li, J. (2015). The Benefit of being physically present: A Survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies*, *77*, 23–37.

Lin, V., Yeh, H. C., & Chen, N. S. (2022). A Systematic review on oral interactions in robot-assisted language learning. *Electronics*, *11*(2), 290. https://doi.org/10.3390/electronics11020290

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Sage.

Liu, D.-J., Huang, R.-H., Chen, N.-S., & Fan, L. (2016). *The Current situation and trend of global development of educational robots*. Posts & Telecommunications Press.

Madakam, S., Ramaswamy, R., & Tripathi, S. (2015). Internet of Things (IoT): A Literature review. *Journal of Computer and Communications*, 3, 164-173. http://doi.org/10.4236/jcc.2015.35021

Morales, S. (2017). Locating the European digital kitchen in its research context. In *Task-based Language Learning in a Real-World Digital Environment: The European Digital Kitchen* (pp. 19-42). Bloomsbury Publishing

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*, 309-326.

Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J.-J. (2013). A Review of the applicability of robots in education. *Journal of Technology in Education and Learning*, *1*(209–0015), 13. http://doi.org/10.2316/Journal.209.2013.1.209-0015

Nikolaevskaya, I. (2017). Multidimensional model of the ZPD as tool of analysis of the child's cognitive-personal dynamics of development while overcoming learning difficulties. *CRIRES International Review: Innovating in the Vygotsky Tradition*, *4*(1), 154-160.

Pachidis, T., Vrochidou, E., Kaburlasos, V., Kostova, S., Bonkovic, M., & Papić, V. (2019). Social robotics in education: State-of-the-art and directions. In *International Conference on Robotics in Alpe-Adria Danube Region* (pp. 689-700). Springer.

Pallotti, G., Niemants, N., & Seedhouse, P. (2017). Vocabulary learning in a real-world digital environment. In *Task-based Language Learning in a Real-World Digital Environment: The European Digital Kitchen* (pp. 207-229). Bloomsbury Publishing

Pasfield-Neofitou, S., Huang, H., & Grant, S. (2015). Lost in second life: Virtual embodiment and language learning via multimodal communication. *Educational Technology Research and Development*, *63*, 709–726. https://doi.org/10.1007/s11423-015-9384-7

Pica, T., Kang, H. S., & Sauro, S. (2006). Information gap tasks: Their multiple roles and contributions to interaction research methodology. *Studies in Second Language Acquisition*, *28*(2), 301-338.

Ramlowat, D. D., & Pattanayak, B. K. (2019). Exploring the Internet of Things (IoT) in education: A Review. In S. Satapathy, V. Bhateja, R. Somanah, X. S. Yang, & R. Senkerik (Eds.), *Information Systems Design and Intelligent Applications*. *Advances in Intelligent Systems and Computing* (Vol. 863). Springer. https://doi.org/10.1007/978-981-13-3338-5_23

Roos, J., Reshöft, N., Hartung, L., & Bußwinkel, J. (2017). Assessing and promoting language development in an interactive learning environment. *Task-based Language Learning in a Real-World Digital Environment: The European Digital Kitchen*, 137-156.

Seedhouse, P. (Ed.). (2017). *Task-based language learning in a real-world digital environment: The European digital kitchen* (Vol. 4). Bloomsbury Publishing.

Shilling, C. (2017). Body pedagogics: Embodiment, cognition and cultural transmission. *Sociology*, *51*(6), 1205-1221.

Skulmowski, A., Pradel, S., Kühnert, T., Brunnett, G., & Rey, G. D. (2016). Embodied learning using a tangible user interface: The Effects of haptic perception and selective pointing on a spatial learning task. *Computers & Education*, *92-93*, 64-75.

Spolaôr, N., & Benitti, F. B. V. (2017). Robotics applications grounded in learning theories on tertiary education: A Systematic review. *Computers & Education*, *112*, 97–107.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory and procedures and techniques*. Sage.

Tanaka, F., Cicourel, A., & Movellan, J. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(46),17954-17958.

Toh, L. P. E., Causo, A., Tzuo, P.-W., Chen, I., & Yeo, S. H. (2016). A Review on the use of robots in education and young children. *Educational Technology & Society*, *19*(2), 148-163.

Tondeur, J., Kershaw, L. H., R. Vanderlinde, R., & van Braak, J. (2013). Getting inside the black box of technology integration in education: Teachers' stimulated recall of classroom observations. *Australasian Journal of Educational Technology*, *29*(3). https://doi.org/10.14742/ajet.16

Vygotsky, L. S. (1962). *Thought and language* (E. Hanf-mann & G. Vakar, Trans.) MIT Press.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

Wells, G. (2001). The Case for dialogic inquiry. In G. Wells (Ed.), *Action, talk, and text: Learning and teaching through inquiry* (pp. 171-194). Teachers College Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17*, 89-100.

Xie, L., Antle, A. N., & Motamedi, N. (2008). Are tangible more fun? Comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. *Proceedings of the Second International Conference on Tangible and Embedded Interaction* (pp. 191-198).