Collaborative kinesthetic EFL learning with collaborative total physical response

Fan Wang, Wu-Yuin Hwang, Yung-Hui Li, Pin-Tse Chen & Kinnosuke Manabe

To cite this article: Fan Wang, Wu-Yuin Hwang, Yung-Hui Li, Pin-Tse Chen & Kinnosuke Manabe (2019): Collaborative kinesthetic EFL learning with collaborative total physical response, Computer Assisted Language Learning, DOI: 10.1080/09588221.2018.1540432

To link to this article: https://doi.org/10.1080/09588221.2018.1540432

Published online: 07 Feb 2019.

Submit your article to this journal

Article views: 44

View Crossmark data
Collaborative kinesthetic EFL learning with collaborative total physical response

Fan Wang\textsuperscript{a}, Wu-Yuin Hwang\textsuperscript{b}, Yung-Hui Li\textsuperscript{c}, Pin-Tse Chen\textsuperscript{b} and Kinnosuke Manabe\textsuperscript{b}

\textsuperscript{a}Department of Educational Technology, Jiang Su Normal University, Xu Zhou, Jiang Su, China; \textsuperscript{b}Graduate school of Network Learning Technology, National Central University, Taiwan; \textsuperscript{c}Department of Computer Science and Information Engineering, National Central University, Taiwan

ABSTRACT
This study aims to integrate kinesthetic technology and collaborative learning into total physical response (TPR) and to investigate the effect of collaborative total physical response (CTPR) on English language acquisition by high school students. To accomplish this, we utilized Microsoft Kinects to develop Collaborative Kinesthetic English Learning (CKEL) system designed to accommodate students in pairs during collaborative English learning activities. Thus, we observed significantly improved learning outcomes in the experimental group relative to the control groups in both total score and vocabulary part of the post-test. However, tests carried out for learning different kinds of vocabulary based on CTPR were also analyzed, and the results showed no significant difference between verbs, nouns or adjectives. This experiment revealed that this system would possibly be helpful for English learning regardless of the kind of vocabulary targeted. Therefore, we concluded that CTPR is not only more helpful for English learning than the traditional individual TPR but also worth being employed and requires further exploration in facilitating English learning.

KEYWORDS
Total physical response; English learning; collaborative learning; human-computer interaction

1. Introduction
Language and hand gesture are two types of systems that are highly dependent and mutually influential. In the course of language evolution, our ancestors developed hand gestures to facilitate their vocal communication. The knowledge of native language and linguistic competence was acquired by kinesthetic experience nurtured by observing interactions between caregivers or the environment (Hwang, Chen, Shadiev, Huang, & Chen, 2014; Kuhl, 2010; Lee, Yeung, & Ip, 2016; Tomasello &
Therefore, hand gestures can be viewed as a scaffolding before and during the process of learning a native language. Many related works have shown that body movement helps facilitate one’s learning process. A particularly notable mechanism of interest is the total physical response (TPR). TPR is a language teaching method proposed by Asher, an American psychologist, who emphasized that the essence of TPR lies in having learners respond to verbal instructions using body movements (Asher, 1968). Following the proposal of TPR, Gardner and Hatch proposed the theory of multiple intelligences. They believed that learning and solving problems using the concept of bodily-kinesthetic intelligence, integrating said concept into course material, and using body movement to achieve the goal designed for interaction would be more beneficial to learners (Gardner and Hatch, 1989).

Owing to the development of advanced technologies, there are currently several devices that can detect human body movement without requiring users to wear or hold any remote controller. Typical examples are the Kinect sensors from Microsoft, and Leap Motion Controllers from Leap Motion Corp. Although research for Kinect application mainly focuses on medical care and entertainment (Alexiadis et al., 2011; Chang, Chen, and Huang, 2011; Lange et al. 2012), some can also be applied in education and learning, for example, Kinect for pre-employment training proposed by Chang et al. for intellectual disabilities (Chang, Chou, Wang, and Chen, 2013). Among all the researches done using the TPR method, action or directive phrases using single verbs or noun-verb compounds happened to be the most popular method used in performing specific actions (Asher, 1969; Wolfe and Jones, 1982; Lindstromberg & Boers, 2005; Morton & Mervyn, 2005; Su & Lin, 2013; Chao et al., 2013; Reinders, 2014). So far, it is still rare to see a TPR application that requires learners to use body movement or hand gestures to learn other types of vocabularies and phrases like adjectives, prepositions or adverbs. Hence, this study is designed to influence the way we explore vocabularies and phrases using technologies like Microsoft Kinect and Leap Motion.

Since technologies have potentials to help students learn languages more effectively, it is necessary and important to examine the capabilities and limitations for further collaborative activities. Using the right combinations, technologies like Microsoft Kinect and Leap Motion can possibly be utilized to improve or facilitate learning in details, rather than just its ‘entertainment only’ – perceived function. In the field of language education, collaborative learning has become one of the most popular teaching strategies. However, much of the past researches that applied TPR and sensing technology methods to language learning mainly focused on
single learners and neglected the impact of collaborative interactions. Therefore, this study majorly focuses on using advanced technologies in language learning with the collaborative mechanism. We proposed the application of human motion technology in collaborative learning mode (two persons per group) to verify if this approach could facilitate language learning more effectively. Furthermore, we also explored the efficiency of TPR in both kinesthetic- and non-kinesthetic-related vocabularies and know whether there were differences between them.

1.1. Cognitive theory of multimedia learning with kinesthetic channel

The Cognitive Theory of Multimedia learning was first proposed by Mayer, and it incorporates several concepts including how people learn and how to design instructions (Mayer, 2005). There are three main assumptions in this theory (Figure 1A shows the original diagram of the theory). First, when we process information, there are two separate channels involved; auditory and visual channels. The auditory narration goes into the verbal system while animation goes into the visual system.

![A. Diagram of cognitive theory of multimedia learning](image)

![B. Extended diagram of cognitive theory of multimedia learning](image)

**Figure 1.** Diagram of cognitive theory of multimedia learning and extended diagram of cognitive theory of multimedia learning.
Second, both of these two channels have limited capacity. Third, learning is an active process of selecting, organizing and integrating information based on our prior knowledge. Students would filter the incoming verbal/visual information to yield a text-based/image-based production. Then, the production would be organized into the verbal based/visual based model to make the information more coherent. Finally, the integration would occur when the students connected both of the two separate models with the corresponding events. In Figure 1B, another channel, Kinesthetic, was added to enhance the two basic channels (auditory and visual) in learning, particularly when learning in self or collaborative way. In neuroscience, it is generally known that the motor area of the brain is activated by actions through the kinesthetic channel, therefore, static learning through auditory and visual ways seem insufficient. Actions like motor movement or mouth speaking have to be highlighted, thereby making learning more complete.

2. Literature review

2.1. Language learning with action

For language learners who have already gone through important language-sensitive period like puberty, one of the major factors that facilitate their language learning is social interaction with others (Blakemore & Frith, 2005). According to Blakemore and Frith (2005), even if the learner only observes the motion of other people without doing it himself, the motor area in his brain would still get activated.

Figure 2. The learning activity module and the Kinesthetic interactivity mechanism of the proposed CKEL system.
Learning from observation is easier than learning from oral communication, and the reason might be because the learner’s brain is ready to imitate the gestures when he is observing. Another study by Macedonia and von Kriegstein (2012) reported that making a gesture and imagining a gesture will activate the same cortical areas in the brain. Some research works in neuroscience already affirmed that there are images related to semantic symbols of every vocabulary. If gestures and action can be integrated into foreign language lessons, it may be effective in memorizing new words (Zimmer, 2001), and can also enhance long-term memory compared to using verbal explanation alone (Tellier, 2008).

The TPR proposed by Asher (1966) for teaching a second language (i.e. L2) gave students a context, an environment and fun for learning languages through physical interactions, and it was proven to be more effective in engaging students in learning vocabularies (Er, 2013). After studying the implementation of TPR principle in teaching vocabulary with game designs, Safitri, Setiyadi, and Huzairin (2017) concluded that it is more effective in the improvement of students’ vocabulary. Asher (1966) had however previously suggested that students respond to the L2 command given by the teacher with body movements, as TPR emphasizes more on listening ability and outward behavior. Hence, students are able to understand the given English commands more when TPR is used in teaching vocabularies (Krashen, 1987). In addition, comparing TPR with the traditional method without TPR, Naeini and Mohsen (2016) found out that teaching English vocabularies with TPR approach could lead to better L2 vocabulary learning ability. The theory of TPR does not only help students to understand the L2 command but also enhances the long-term memory of the L2 command. Further proofs from Bahtiar (2017) confirmed that TPR can improve students’ achievement in memorizing vocabulary.

In 1981, Carels proposed the use of hand gestures to systematically to facilitate L2 learning. He suggested that both teachers and students should learn using hand gestures as a good strategy for memory enhancement. Later on, an experiment by Tellier in 2008 further validated the use of hand gestures to be more effective in memorizing new vocabularies than using only visual presentations. In another experiment conducted by Kelly, McDevitt, and Esch (2009), they classified the vocabulary learning process into four modes. Their experiment showed that the most effective of the modes was combining oral pronunciation with meaningful hand gestures. On the other hand, Vygotsky claimed that all knowledge begins with a visible form of social interaction, which is then gradually internalized, forming the thoughts of the learners (Sawyer, 2006). Therefore, it is important to note that there are big
potentials in the exploration of body actions for the purpose of facilitating language learning.

2.2. Language learning with TPR using kinesthetic technology

Recently, due to the progress made in kinesthetic technology, there is an increasing number of researchers applying Microsoft’s Kinect device to learning activities. There are also many applications of language learning that are being developed based on TPR. We give some examples in the following paragraph.

In a recent report by Chao et al. (2013), Kinect was utilized to facilitate English memorization of words containing 13 verb phrases (combinations of verbs and nouns). They showed that for the immediate cued recall test, the performance of the experimental group was significantly better than that of the control group, and similar results were recorded for the delayed free recall test. In another study carried out by Kuo, Hsu, Fang, and Chen (2014), their learning scheme contained 15 vocabulary words for fifth-grade participants from an elementary school. They utilized Kinect for TPR learning activities for an experimental group, while conventional TPR learning strategies without Kinect was employed in teaching another group (control). After conducting the learning activities for 30 minutes, they discovered that there was no significant difference in the results of the post-test and delayed test between the two groups. However, the experimental group that used Kinect and ITC devices had a longer retention period than the control group that only used conventional TPR learning activities.

One particular feature of TPR is allowing students to execute actions or speak after they had heard and understood the meanings of words. A common way of doing this is to use verbs or a phrase consisting of verbs and nouns to form a behavioral command, then students attempt to respond to the commands with voice and action concurrently (Asher, 1969; Wolfe & Jones, 1982; Lindstromberg & Boers, 2005; Su & Lin, 2013; Chao et al., 2013; Reinders, 2014). A particular study by Kashiwagi, Xue, Sun, Kang, and Ohtsuki (2009) claimed to have applied TPR to language learning, where they used a computer system to give English commands by asking a learner to grab a target object and scan its RFID tag into the system. These command sentences consisted of adjectives and nouns. However, they did not integrate hand gestures into the learning process and the commands were all about grabbing something. Therefore, the empirical study of applying body behaviors, actions or hand gestures to learn other kinds of words remains exclusive (for
example, adjectives or adverbs) in the research of computer-assisted learning.

To sum up all results from the above literature, it will be promising to combine vocabularies with customized hand gestures or body motions (which correlates with words to be learned) to help students learn or memorize other words besides verbs. In this study, we will conduct the empirical study with a Kinesthetic device (Microsoft Kinect) to investigate the issue. This is a novel research direction that is worthy of being explored.

2.3. Collaborative learning with TPR

After repeated social interactions, an individual’s expressions and concepts will be internalized in another individual (Jorczak, 2011; Vygotsky, 1986), and the learners who are lacking self-confidence may tend to collaborate with peers who have deeper knowledge and skills; so that they can acquire the ability to achieve the same goal afterward (Nikolhelperou, 2012). Several studies including Vygotsky (1986), Donato (1994, 2000) and Swain (2000) have emphasized the importance of peer collaboration and scaffolding in the process of collaborative dialog (Liang, 2010). In fact, peer effect appears to be very strong in the process of language learning. Children tend to learn accent from their peers instead of their parents (Blakemore & Frith, 2005). Collaborative learning has been applied extensively in the learning process that covers a variety of different domains. Examples include music learning assignment (Nikolhelperou, 2012), English reading and writing for EFL participants (Liang, 2010), and English learning by PDA phone game in the way of context awareness and collaborative learning (Liu & Chu, 2010).

Although TPR emphasizes personal movement, it is also possible to apply it to collaborative learning. Garcia, Asher, and Asher (1988) suggested the encouragement of group participation at the commencement of TPR. Based on this TPR theory, new applications can be developed, for example: acting or TPR-Storytelling (TPR-S). Acting is a teaching method that gives students opportunities and atmosphere to learn with TPR. The strategy of TPR-S integrates the vocabularies learned in TPR into a story. This facilitates students’ ability to deeply understand new vocabularies when they use hand gestures, pictures or other stage materials to act according to the new plots. After a deep understanding of the story, students can perform the story collaboratively, or they can imagine themselves as other characters when watching the performance. Both TPR and TPR-S use the scaffolding teaching strategy of Vygotsky. In the task execution process which desperately requires assistance, teachers or peers can offer help to the learners (Cantoni, 1999). Therefore, incorporating
TPR into storing telling, acting or games can facilitate effective learning among peers through reciprocal teaching and assisting. Hence, it is indeed promising to deeply investigate this method with collaboration.

In another study, it was reported that positive results were achieved because of the embodiment designed into classroom learning and the high degree of collaboration which helps in science classes (Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014). They further explained that learning with advanced technologies for a single-user model can, in some cases, lead to an isolating experience that runs counter to students’ highly socio-collaborative experiences. Therefore, immersive learning environments can be ‘technological’ and ‘collaborative’ and can foster as well as support quality face-to-face, co-present collaboration. Indeed, this study was decisive in our consideration of how to use embodiment design and collaboration to improve sensorimotor activation, gestural congruency with content, and perception of immersion.

From the above literature review, it would be a promising exploit to combine all the views and opinions into a research direction for the purpose of achieving a more effective way of learning by designing collaborative learning with TPR in computer-assisted learning i.e. using Kinect’s recognition techniques based on TPR to implement collaborative English learning. From this, it is expected that learners could acquire and memorize more vocabularies through physical interactions. So far, there are very few past empirical studies where researchers applied body behaviors and actions or gestures to learn words, particularly with collaborative activities. In most studies, researches did not present a broad statistical analysis of such kind of physical behaviors and their effects on English learning, particularly for different kinds of English vocabularies. Here, we deeply investigate the physical behaviors of collaborative activities and their influences on English learning.

3. System design

3.1. Overview

In this research, there are two main modules in the proposed Collaborative Kinesthetic English Learning (CKEL) system: the learning activity module and the kinesthetic interactivity mechanism as shown in the following Figure 2.

In this design, members in each group need to cooperate to compete against other groups. To encourage the cooperative activities between members of a group, members are allowed to communicate face to face, and the result of the competition between each group is shown immediately after. We constructed the hardware structure as follows: a group
has two members and each group has one PC, two screens and two Kinects. Each person uses one screen and one Kinect but cannot watch the screen of his/her teammate.

### 3.2. Body interactive mechanism

There are three modes for learning English language in interactive mechanism: (1) motion mimic, (2) direction recognition and (3) object recognition. All of these three modes exist in all learning modules. Each module uses all of the three modes to facilitate learning. Because these three modes are connected to learners’ physical interaction, they may possibly facilitate learning different types of vocabularies like verbs, nouns, preposition, etc. Meanwhile, we also engaged in an interactive survey and collaborated with experienced English teachers to design learning activities using these three modes. The goals were to help students familiarize with our Kinect system design and help their English vocabulary/sentence learning gradually. That is, the combination of ‘motion mimic’, ‘direction recognition’ and ‘object acquisition’ will be
utilized to connect their physical activities with different types of vocabularies, verb, direction preposition, and noun respectively. Further considerations are being explored to include more vocabularies like conjunctions in future studies.

3.2.1. Motion mimic
This mechanism helps the students to experience the meaning of the words through motion or posture. Motion mimic requires learners to practice some special postures which will include movements of the body in a stepwise manner, a typical example is locking the limbs into a special pose. The system displays an example video that illustrates the meaning of the vocabulary, and the learner is asked to imitate the motion or posture. If the system detects the learners’ motion or posture which fits in with the requirement, the system gives the learner a positive response. We expect that learners would better understand and internalize the vocabulary by interpreting the meaning of the vocabulary through the execution of motions and poses.

3.2.2. Object recognition
This system setup the pair between the target words and the corresponding objects. During this process, the learner is asked to pick up the object that fit the meaning of the words shown on the screen, then the object is placed in front of the Kinect, and the system performs automatic object recognition. If the object picked by the learner is correct, positive responses will be given to the learner. We believe that the act of picking up the corresponding object can enhance the learners’ impression and perception of the vocabulary.

3.2.3. Gesture-based relative positioning
In this module, the system teaches the learner vocabularies based on the learner’s relative position. For example, the phrase ‘in front of’ asks the learner to put his right arm in front of his body. Through the repositioning of the arms relative to the body, the learner learns the reposition vocabularies and it is helpful to memorize them.

3.3. Learning activity module
For this research activity, we designed four major modules for learning. The learner enters the desired module by pushing the corresponding air button when standing in front of Kinect device and the learning activity begins. Based on the previously explained basic mechanisms (motion mimic, direction recognition, and object acquisition), we designed 4 curricula: (1) Vocabulary Learning, (2) Peer Vocabulary Practice, (3) Peer
Vocabulary game and (4) Peer Sentence Game. On the one hand, the first two (Vocabulary Learning and Peers Vocabulary Practice) were designed for participants to learn/practice vocabulary individually and to familiarize with the system. On the other hand, the last two activities (Peer Vocabulary game and Peer Sentence Game) were designed to stimulate their motivation towards learning because the game (competition) involves helping their partners to learn more vocabulary and applying the vocabularies into sentence formation. Hence, the higher their performance, the higher the scores recorded for them in the game.

3.3.1. Vocabulary learning

This module in learning activity was designed for individuals. Words were divided into two parts: ‘Learning by Actions’ and ‘Learning by Objects’. In ‘Learning by Actions’, when learners select a new word, the screen shows a sample video of the selected word on the right side, and the learner’s real-time scene will be visible on the left side. Concurrently, the system broadcasts the pronunciation of the word twice. Learners have to imitate the action shown in the sample video. Furthermore, a correct reaction from the learner prompts the display and pronunciation of the word in both English and Chinese and the system records a positive mark for the learner. The use of this method will encourage learners to practice more by imitating and speaking, and the relation between the meaning, pronunciations, and action of the new word is easily established in learners’ memory.

In the ‘Learning by Objects’ method, we prepared all the objects needed by the learner and placed them on top of the learner’s desk. Here, new words are displayed with a message asking the learner to answer with the object closely related to this word. The prompt of the Chinese translation of the target word is also displayed on the screen. When the learner takes an object to the assigned area in front of the Kinect camera, it triggers the system to perform object recognition. If the object selected by the learner does not correspond to the target vocabulary, the system will show the message ‘try again’. If the object selected by the learner is the correct one, the system will show the meaning of the target vocabulary in Chinese, a sentence example in both English and Chinese, and the pronunciation of the sentence.

Every time the system detects the correct answers from the learner, the exact word is displayed or pronounced and a positive response is given to the learner. This method, if adopted, will encourage the learner to practice more by imitation and speaking, and the relationship between the meaning, pronunciations, and body movement of the new word is established in the learners’ memory.
3.3.2. Peers vocabulary practice
This model is for teams comprising a pair of students. The system randomly selects a word and pronounces it. As team members listen to the pronunciations, a member from each team will be selected as either a respondent or a helper. The respondent is then asked to do the action corresponding to the word. If the respondent does not know what to do, he calls for help, then the helper can help the respondent by demonstrating the proper action, telling tips, or triggering the tips button on the screen. This system gives three kinds of hints sequentially. These hints include broadcasting the pronunciation twice, showing the word and displaying the example video or tips. The ‘tips’ button has a three-second countdown. If the helper triggers the button again in three seconds, the system shows the message ‘Try to think again’. Every hint given by the system is shown to both the respondent and the helper.

Once the respondent completes the assignment, the system shows the word in Chinese, the example sentence in both English and Chinese, and the pronunciation of the example sentence. Subsequently, the system automatically picks another word, then the respondent and helper will exchange in turns.

3.3.3. Peers vocabulary game
This is a competitive gaming activity that requires cooperation from team members. Scores are accumulated based on the team’s performance. A randomly selected word by the system is displayed on the screen of a member of each team. The member with the selected screen has to pronounce it to the partner. When the partner understands what the new word means, he/she should perform the action corresponding with the vocabulary. However, there is a time limit for each question as well as for the whole game. When the team correctly responds to the system or when the time for the question is up, the two members switches roles in turns. When the game is over, the system lists vocabularies which learners did not answer correctly. Team members can help or remind each other by selecting the word in the list as they listen to the pronunciation of the word.

3.3.4. Peers sentence game
This is another competitive gaming activity which needs cooperation in a team. Scores are accumulated based on team performance. The system selects words randomly and speaks them to both team members. One of the team members will be designated by the system as the respondent to the randomly selected words. The screen shows the total number of assignment, then the team member who is not selected as the respondent
will be responsible for reading out new words to his partner. Two seconds after the system has pronounced the question, it displays the example sentences on the screen of the helper, who should find out the new word in the sentence and read it out to his partner in order to help him answer the question correctly. However, there is a time limit for the whole game as well as the allowable time for answering each question. During the period at which the system speaks the questions, it will not detect the users’ behavior, and the timer will remain fixed. After the question is completely read out, the timer will restart and the system will initiate the detection of users’ behavior. At the end of the game, the system shows the list of the sentences that can be generated from the words that the team did not answer correctly. The team members can click on them and the system will display the complete English and Chinese explanation of the example sentences. The team members can then review and discuss their performance to help each other.

4. Method

4.1. Research participants and research questions

Seventy-nine participants from the eleventh-grade of a vocational high school participated in this research experiment. The participants were mainly female students between the ages of 16–17 and were taught by the same English teacher. Applied language major students were selected for this research due to their interest in learning English language. They were divided into three groups: (1) an experimental group comprised of 30 participants, (2) a control group I with 30 participants and (3) a control group II which consisted 19 participants. The content of learning was inspired by an English magazine which was assigned as an additional teaching material in the school. After several discussions with the English teacher of the participants, words were selected from the magazine. Hence, the syllabus was declared suitable for students.

The objectives of this study were to facilitate EFL learning with our proposed CKEL and to study students’ learning behaviors, as well as the influence of CKEL on learning achievement. Meanwhile, with respects to different kinds of vocabularies, we also examined the possible significant difference in learning improvement across the groups.

Our research aim included:

1. Exploring potential significant differences or improvements in English learning achievements with the use of post-test and delay test, peer-group collaborative learning with TPR and Kinect support,
individual learning with TPR and Kinect support, and traditional learning without Kinect support.

2. Probing the correlations in learning behaviors using Collaborative TPR and learning achievements; as well as studying the effect of these behaviors on learning achievement using multiple regression analysis.

3. Investigating the differences between the usage of collaborative learning and individual learning with TPR and Kinect support to improve learning of different types of vocabularies.

4.2. Activities design

Figure 3 shows experimental procedure and activity design. Forty-eight vocabularies and the corresponding sentence examples were the contents of this design and were chosen by the participants’ English teacher. This teacher was familiar with participants’ English levels and possesses 10 years of teaching experience. Therefore, based on their levels, the teacher confirmed these vocabularies to be suitable for their lessons. The confirmed vocabularies were classified as either ‘motion mimic’ and ‘direction recognition’ \( (n = 37) \) or as ‘object acquisition’ \( (n = 11) \). Hence, these vocabularies were included for each category (please see Appendix I). The experimental group and control group I used CKEL, while Control group II used the original learning mode which featured watching videos and learning the same vocabularies as the other two groups. In the experimental group, the participants assigned to the same team were designated to sit at the same table.

4.2.1. Pre-test

The exam had four parts with a total of one-hundred marks (Appendix II). In part A, there were ten multiple-choice questions designed to test students’ listening ability while hearing the questions with reading the figures. Part B included a total of 10 questions designed to test students’ listening skills as each question is being read out. Part C also contained ten questions designed to test the listening ability during conversations. While part D contained five questions to test listening understanding of a short speech. We emphasized on the listening because, in the proposed collaborative activities, the meanings of vocabulary/sentence by action were demonstrated with voice to their partner (e.g. Peer Vocabulary Game/Peer Sentence Game). Thus, students need to have a good English listening ability to be able to give a correct answer. Therefore, we evaluated their listening ability with these tests. After extensive discussions, these test contents (including pre-test and post-test) were
collaboratively designed with their teacher. Both pre-test and post-tests have 35 questions each and were graded by the English teacher.

4.2.2. Activities for week 1
At the beginning of the experiment, we introduced the CKEL device and explained the functions of Vocabulary Learning and Vocabulary Practice to members of the experimental group and the control group I. This was the first time the participants would be experiencing this learning system, therefore, the participants were allowed to explore and familiarize themselves with how to construct the corresponding motion of a relative vocabulary, posture or object in our system. Each attendant was equipped with a CKEL device and a headphone for learning.

4.2.3. Activities for week 2
Two new activities were conducted during the second week, namely; vocabulary Game and Peer Vocabulary Game. In each experimental subgroup, participants shared the same PC but separate screens designed to interact with the Kinect device individually. The experimental group used the Peer Vocabulary Practice module for collaborative learning while control group I used the Vocabulary Practice module to study individually. In each experimental sub-group, one participant was designated as the ‘respondent’, and the other participant as the ‘helper’. The respondent performed the action corresponding to the word selected for him/her. In any case, where the respondent could not perform the correct action, the helper assisted the respondent by demonstrating the proper action, telling tips, or triggering tips button on the screen for the system to give some helpful hints to the respondent.

At the end of the game, the participants were shown their score list and the list of incorrect answers. Meanwhile, ranking was used to motivate the participants to be proactive.

4.2.4. Activities for week 3
Sentence Game activities and Peer Sentence Game activities were introduced into the experiments in the third week. The experimental group used the Peer Sentence Game while control group I used the Sentence Game. We changed the angle of the screens used by the experimental group to ensure that members of the same team were unable to see their partner’s screen while the control group I had to use headphones during the Sentence Game. For the Peer Vocabulary Game, any time the respondent failed to perform the action corresponding to the sentence, she beckoned on the helper. Subsequently, the helper responded by
reading out new words to make the sentence more understandable or demonstrated the proper action corresponding to the sentence, and in other cases, she showed the telling tips or triggered tips button on the screen.

At the end of the game, scores and incorrect answers were displayed to participants of the experimental group and control group I. Ranking was also used to motivate students.

### 4.2.5. Post-test and interview

In the fourth week, all of the participants were given an hour-long post-test. The post-test was divided into two parts, namely: the vocabulary test and the sentence test. In the vocabulary test, participants were asked to write down the spelling of the English words they learned and their translations in Chinese by free recall. In another section of this test, the participants listened to a broadcast and were asked to write down the spelling of the English words and their translations in Chinese. This was done to influence their learning speed and word retention. Meanwhile, because the participants were vocational high school students whose English proficiency was quite poor, we also tried to enhance and evaluate their basic knowledge of English language by spellings.

The sentence test included sentence listening and reading (Appendix III). We focused on participants’ understanding of the vocabularies rather than the accuracy of their grammar. There were 35 multiple-choice questions divided into three parts in the sentence test, these included pure listening test, listening test with figure support, as well as conversation or a short speech. The sentence test was designed to evaluate whether students’ vocabulary ability (memorizing and listening) had improved after using our proposed system with kinesthetic support and peer help.

After analyzing the efficiency of learning, we used a purposive sampling method to interview the participants face-to-face. We selected students for interview based on their performance and engagement in the proposed system. The best trio, the bottom 5 percent of the post-test score, and the best students in the activity engagement were selected for an interview. The reason for the interview was to further verify the validity of our statistical results and findings. For example, how peer activities can benefit students’ English learning through interaction. In the interview, they were asked questions like; ‘What part of the proposed system is preferable for collaboration with your partner?’, ‘What type of vocabularies do you think is most effective when using kinesthetic action in the proposed system?’, ‘would you prefer this system to the traditional way of learning?’, ‘what differences did you notice in your learning abilities?'
how is kinesthetic learning different from other systems of English learning?, ‘How do you feel about learning via this system?’ and ‘did you have fun or not?’

5. Results

5.1. Analysis of learning effects

ANOVA variance analysis was used to analyze the differences in the English proficiency between the experimental group, control group I and control group II. Prior to the introduction of the new learning tool to interfere with their English performance, we expected students to speak and use actions in their collaborative activities. Students needed to deliver the meanings by actions and speaking in peer activity (e.g. Peer Vocabulary Game/Peer Sentence Game). Thus, we used listening comprehension as a part of our pre-test. We also utilized their grade in the pre-test to represent the prior knowledge of English, and the result is shown in Table 1.

Although there was no statistically significant difference in the pre-test, we still used ANCOVA analysis on the post-test data. Therefore, the pre-test scores of prior knowledge were used as an interfering variable to adjust the post-test scores. Because there were two scores for vocabulary test and sentence test in the post-test, we normalized each of them to a maximal value of 100, which brings the maximum summed score to 200. The result is shown in Table 2.

Regarding the analysis of the result in post-test, the learning efficiency of the experimental group was significantly better than that of the control group I and II. However, the learning efficiency of control group I was shown to be significantly better than control group II. Furthermore, we specifically analyzed the vocabulary test in the post-test and we confirmed the performance ranking to be the same. There was a significant gap in the learning performance of control group II compared with the others. This may be due to the flawed traditional learning method used by control group II. Although we also provided the list of vocabularies and sentences used in CKEL to control group II, control group II cannot really practice them because CKEL system was not provided to control

| Table 1. The pre-test Analysis by ANOVA. |
| --- | --- | --- | --- | --- |
| N | Mean | SD | F |
| Experimental | 30 | 49.90 | 16.919 | .168 |
| Control (I) | 30 | 44.90 | 16.845 |  |
| Control (II) | 19 | 54.11 | 15.165 |  |

*p < .05.
group II. Thus, this may be a reasonable lacuna responsible for the poorer learning performance of the control group II.

The activities performed by the experimental group and control group I were also different. Although both groups used CKEL system for assistive learning, control group I performed their activities in single-player mode while the experimental group not only performed activities in pairs but also established good cooperation with themselves. From the observation and interview, it was found that the experimental group helped each other more and also acquired more knowledge by helping their partners when there are difficulties in answering questions in the peer sentence and vocabulary games. Taken together, the result of both total scores and scores in vocabulary test in post-test showed that learning efficiency of the experimental group is obviously better than control group I. Such result suggests that learning, when combined with cooperative activity can improve the learning efficiency.

In the analysis of results in sentence test of post-test, the learning efficiency of the experimental group and control group I was significantly better than control group II. Although the learning efficiency of the experimental group was slightly higher than that of the control group I, there was no statistically significant difference recorded. This might be due to the type of questions in the sentence test, including the vocabularies which might have been excluded in the given target list. If the students do not understand the words or grammar, it affects their ability to understand those questions. In addition, the students spent more time on learning vocabularies while ignoring the grammatical rules related to the words when they are used in a sentence. All of these factors are culpable in the analysis showing there was no significant difference in post-test score between experimental group and control group I.

On the other hand, for the control group II, their poor scores signified that they did not learn the target vocabularies well. Considering the reason above as stated in the previous paragraph, it is pretty reasonable that their average scores in post-test are much worse than the other two groups.

### Table 2. The post-test analysis by ANCOVA.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted Mean</th>
<th>SE</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Post-test</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>30</td>
<td>127.35</td>
<td>20.396</td>
<td>126.801</td>
<td>3.415</td>
<td>93.80***</td>
</tr>
<tr>
<td>Control(I)</td>
<td>30</td>
<td>114.36</td>
<td>29.211</td>
<td>115.650</td>
<td>3.418</td>
<td></td>
</tr>
<tr>
<td>Control(II)</td>
<td>19</td>
<td>55.91</td>
<td>22.052</td>
<td>54.733</td>
<td>4.293</td>
<td></td>
</tr>
<tr>
<td><strong>Vocabulary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>30</td>
<td>62.01</td>
<td>13.050</td>
<td>61.737</td>
<td>2.198</td>
<td>118.57***</td>
</tr>
<tr>
<td>Control(I)</td>
<td>30</td>
<td>53.98</td>
<td>17.553</td>
<td>54.632</td>
<td>2.201</td>
<td></td>
</tr>
<tr>
<td>Control(II)</td>
<td>19</td>
<td>10.50</td>
<td>10.470</td>
<td>9.900</td>
<td>2.763</td>
<td></td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>30</td>
<td>65.33</td>
<td>11.298</td>
<td>65.064</td>
<td>1.901</td>
<td>23.18***</td>
</tr>
<tr>
<td>Control(I)</td>
<td>30</td>
<td>60.38</td>
<td>14.330</td>
<td>61.018</td>
<td>1.903</td>
<td></td>
</tr>
<tr>
<td>Control(II)</td>
<td>19</td>
<td>45.41</td>
<td>13.261</td>
<td>44.833</td>
<td>2.390</td>
<td></td>
</tr>
</tbody>
</table>

***p < .001.
5.2. *Analysis of the different vocabulary learning effects by actions*

In this experiment, there were 48 vocabularies provided for learning. Among them, 37 were related to using motion mimic or direction recognition to help learners create memory association with vocabularies. We partitioned these 37 vocabularies into four categories: 15 verbs (including five verbal phrases), seven nouns, nine adjectives, and six direction prepositions. The scores were given based on the accuracy of two parts separately: spelling and Chinese translation. For each part, learners can get one point if their answer was correct. In the end, we summed up the points and calculated the accuracy in four categories. Next, we used ANOVA to analyze the data. From Table 3, we found that the result of direction preposition was better than other categories. This may be due to the lesser number of direction preposition or the direct meaning of those words. Furthermore, some learners said they had already learned most of the vocabularies in the direction preposition before the experiments. Hence, we concluded that TPR can enhance learning direction preposition vocabularies when meaningful gestures were used and associated with the target vocabularies.

5.3. *Analysis of the relationship between system usage and learning effects*

5.3.1. *Pearson’s relation analysis between post-test effect and the system usage*

In the first week, both control group I and the experimental group were taught in similar conditions, however, after a week, only the experimental group used computers in pairs. It is difficult to track the time spent on learning for each separately. So, the analysis of learning time is only for the first week. We performed Pearson’s Relation Analysis between the pre-test scores and the number of actions performed during vocabulary learning process in the first week. The results are shown in Table 4. We found that there was no significant correlation. Besides, we also performed Pearson’s Relation Analysis between the pre-test scores and the

Table 3. The analysis of the different vocabulary learning scores by actions.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD.</th>
<th>F</th>
<th>LSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>30</td>
<td>68.889</td>
<td>15.094</td>
<td>12.200***</td>
<td>D.P. &gt; Verb</td>
</tr>
<tr>
<td>Noun</td>
<td>30</td>
<td>66.429</td>
<td>14.548</td>
<td></td>
<td>D.P. &gt; Noun</td>
</tr>
<tr>
<td>Adjective</td>
<td>30</td>
<td>62.593</td>
<td>18.277</td>
<td></td>
<td>D.P. &gt; Adjective</td>
</tr>
<tr>
<td>D.P.</td>
<td>30</td>
<td>85.278</td>
<td>14.464</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control(I)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>30</td>
<td>61.000</td>
<td>19.021</td>
<td>11.645***</td>
<td>D.P. &gt; Verb</td>
</tr>
<tr>
<td>Noun</td>
<td>30</td>
<td>62.857</td>
<td>23.459</td>
<td></td>
<td>D.P. &gt; Noun</td>
</tr>
<tr>
<td>Adjective</td>
<td>30</td>
<td>54.630</td>
<td>19.524</td>
<td></td>
<td>D.P. &gt; Adjective</td>
</tr>
<tr>
<td>D.P.</td>
<td>30</td>
<td>83.611</td>
<td>18.242</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***p < .001.
time spent on learning in the first week. The steps in measuring the learning time are as follows: the timer starts when a learner makes some gestures but pauses if the learner did not make any motion gesture or did not operate the system in two minutes. Then the timer restarts when the learner resume motion gestures or operate the system. This result showed that there was no significant correlation as well in Table 4. The two results above show that the prior knowledge had no correlation with how much time the learners used with the system.

Next, we performed Pearson’s Relation Analysis between the learners’ post-test scores and the number of gestures they used during vocabulary learning in the first week. We noticed that there was still no significant correlation in Table 4. However, after performing Pearson’s Relation Analysis between the learners’ post-test scores and the time spent during the learning process, we found that there exists a significant correlation between them in Table 4. Furthermore, we observed that there are some learners who did not concentrate during the learning process. Although they did perform the corresponding gestures required by the system, in the learning session they kept talking with other learners without paying attention to the information shown on the screen. The gestures and motions they performed became a routine instead of a meaningful learning activity.

We interviewed the learners who made more gestures but got lower scores in the post-test. Their reply was that they only read through roughly and did not read the example sentence, and also did not try harder to remember the spelling. Then we interviewed the learners who made fewer gestures but got higher scores and observed that they concentrated on remembering the meaning and spelling when they made gestures. Some said they already knew those vocabularies before, and they also paid more attention to reading on the screen and memorizing the words. Besides making gestures, during the learning process, one should also spend time listening and reading the information in the learning system. The information includes the vocabularies, their corresponding gestures, Chinese translation and example sentence. Blakemore and Frith (2005) pointed out that even if the learner only observes the motion of other people without doing it himself, the motor area in his brain would also get activated. Learning from observation is easier than

<table>
<thead>
<tr>
<th></th>
<th>Learning spent time in first week</th>
<th>#Action for vocabulary learning in first week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test Pearson</td>
<td>.186</td>
<td>−.003</td>
</tr>
<tr>
<td>Post-test Pearson</td>
<td>.315*</td>
<td>.162</td>
</tr>
</tbody>
</table>

*p < .05.

Table 4. Pearson’s relation analysis between post-test effect and the system usage (N = 60).
from oral communication, and the reason might be because the learner’s brain is ready to imitate the gestures when he is observing. Macedonia and von Kriegstein (2012) also noted that some research works reported that making a gesture and imagining a gesture will activate the same cortical areas in the brain. Interestingly, neuroscience research also confirmed that there are images related to semantic symbols of every vocabulary. When the vocabulary is recalled, the corresponding images will also be recalled. Thus, when learners make an association between gestures, vocabulary and its translation in their first language, the memory effects last longer. Furthermore, learners who spend more time tend to get higher post-test scores. Hence, we concluded that using this system does help learners to acquire English knowledge better.

In the second week, the major activities were games. Since the conduct of activities for the control group and the experimental group was different, we focused on the 30 students in experimental groups when analyzing the correlation between the gaming activities and learning effectiveness. The more frequent the learner used the system, the more familiar they got with the process, the vocabularies, and the gestures provided by this system. Choosing the round where the learners have the highest total score, we performed Pearson’s Relation Analysis on the number of correct answer with total scores, vocabulary scores and sentence scores in the post-test, the result is shown in Table 5. In Peers Vocabulary Game, the number of correct answers significantly correlated with total scores, vocabulary scores and sentence scores in the post-test. In this section, learners have to remember the pronunciation of the vocabulary, Chinese translation and its corresponding gestures to get high scores in the game. The students who got higher scores in Peers Vocabulary Game also got higher scores in the post-test. This may be because Peer Vocabulary Game uses only one word for each question, and it is easier for the learners to focus on one word at a time. Moreover, teammates also relied on each other. This can also contribute to learners’ ability to get higher scores in the post-test. The other reason is that Peer Vocabulary Game had only 48 vocabularies and they are the same vocabulary used in the post-test. Although there are new

<table>
<thead>
<tr>
<th>Table 5. Pearson’s relation analysis between the number of correct answers (both in total and before showing text) and post-test results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Peers Vocabulary Game correct answer in highest Post-test</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>#Peers Sentence Game correct answer in highest Post-test</td>
</tr>
<tr>
<td>before showing text</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01.
vocabularies in the question sentences in the post-test, there exists a higher correlation between the number of correct answers in Peer Vocabulary Game and the vocabulary scores in the post-test.

Regarding the correlation of the highest number of correct answers in Peers Sentence Game, together with total scores, vocabulary scores, and example sentence score in the post-test in Table 5, there was statistical significance only in the sentence score in the post-test. To get higher scores, a learner needs to understand the vocabulary in the questions appearing in the form of an example sentence in the game and link it to its corresponding gestures. The learners who had a higher number of correct answers in Peer Sentence Game also appeared to get higher scores in the sentence questions in the post-test. We believe that if a learner can know more vocabularies outside the target set (the given 48 words) and familiarize with the example sentence, it will be easier to get higher scores in the game. However, if a learner only knows the target vocabularies, though she can get higher scores in the vocabulary test in the post-test, it is difficult to predict that she will also score high points in the Peer Sentence Game. Therefore, there is no significant correlation between the highest number of the correct answers and the vocabulary test in the post-test.

We also applied Pearson’s Relation Analysis to the number of correct answers before showing the text, the results are also shown in Table 5. There exists a significant correlation between the sentence scores of the post-test. This is reasonable because if a learner can answer the sentence questions correctly even before the target vocabulary appears, it means the learner is extraordinarily familiar with both the target vocabulary and its example sentence. On the other hand, to score higher marks in the vocabulary test in the post-test, one only needs to be familiar with the 48 target vocabularies. However, for one to score higher marks in the Peer Sentence Game before the target text shows up, it takes far more than being familiar with the 48 target vocabularies.

5.3.2. The forecasting ability of the system usage for post-test effects

Using stepwise regression method, we analyzed the relationship between (a) learning time in the first week; (b) the number of gestures performed during vocabulary learning in the first week; (c) the number of correct answers in the highest scores in Peer Vocabulary Game during the second and the third week and (d) the number of correct answer before showing text in Peer Sentence Game during the third week. Then we evaluated how much all of those four factors affect post-test scores. From the results in Table 6, we observed that only factor C could effectively forecast the total scores, vocabulary scores and sentence scores in
the post-test. This means that in each team, if one member tries hard to read clearly for her partner and the other tries as hard to focus and understand what is being read, it enhances their ability to jointly work together and increases the efficiency of learning. The finding is consistent with Cantoni (1999) where it was reported that peers can offer help to the learners in the task execution process which desperately requires assistance. On the other hand, if the learners can score high points in Peer Vocabulary Game, it means they are familiar with the vocabularies, which is the foundation of getting high scores in the post-test. In addition, we believe that one needs to know more out-of-scope vocabularies to score higher marks in the sentence test. Such factors would impact the number of correct answers from the learners. This explains why factor C had less impact on the scores of sentences test than that of vocabulary test in the post-test.

### 5.3.3. Performance analysis of different types of vocabulary

To further determine whether there were significant differences in learners’ performance in different types of vocabularies between experimental group and control group I, a t-test analysis was conducted to compare four types of vocabularies between the two groups; verb (V), noun (N), adjective (Adj) and directional preposition (D.P.). Our analysis results in Table 7 indicated that there was a significant difference in V performance between the two groups. On the other hand, the remaining three did not show any statistical differences. We also observed that the reason why statistical differences were only found in V was that the experimental group used a lot of action in their learning activities, particularly in peer activities. Hence, they could get more improvement via these actions, compared to the other three kinds of vocabularies.

### 5.4. The discussion and educational implication

Based on our experimental results and the result of the interview, we advise the educators who would like to use this medium as a study method that; applying such learning method will make learners feel

| Table 6. Stepwise regression analysis to predict post-test results. |
|------------------------|----------------|----------------|-------------|----------|----------------|
| Forecast variable       | B             | SE            | Beta        | T        | Adj. R²        |
| Post-test (F = 15.391)  | #Peers vocabulary game correct answer in highest | 3.971 | 1.012 | .596 | 3.923 | .332** |
| Vocabulary of post-test (F = 12.076) | #Peers vocabulary game correct answer in highest | 2.342 | .674 | .549 | 3.475 | .276** |
| Sentence of post-test (F = 6.764) | #Peers vocabulary game correct answer in highest | 1.629 | .626 | .441 | 2.601 | .166* |

*p < .05, **p < .01.
Table 7. Independent samples test (Verb, Noun, Adjective and D.P.) for experimental group and control group I.

<table>
<thead>
<tr>
<th></th>
<th>Levene’s test for equality of variances</th>
<th>t-test for equality of means</th>
<th>95% Confidence interval of the difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>Verb</td>
<td>2.170</td>
<td>.149</td>
<td>-2.057</td>
</tr>
<tr>
<td></td>
<td>Equal variances assumed</td>
<td></td>
<td>Equal variances not assumed</td>
</tr>
<tr>
<td>Adj</td>
<td>.076</td>
<td>.784</td>
<td>-1.537</td>
</tr>
<tr>
<td></td>
<td>Equal variances assumed</td>
<td></td>
<td>Equal variances not assumed</td>
</tr>
<tr>
<td>D.P.</td>
<td>.273</td>
<td>.604</td>
<td>.079</td>
</tr>
<tr>
<td></td>
<td>Equal variances assumed</td>
<td></td>
<td>Equal variances not assumed</td>
</tr>
</tbody>
</table>
satisfied. It also enhances the motivation of the learners towards learning English if the TPR theory is combined with collaborative learning. Using this method is also helpful in enhancing the vocabulary learning effect if the learning activities are designed based on the proposed methodology. However, for the sentence learning, although the experimental group scored higher than the control group I, the two groups have no statistically significant difference. This may be because the grammar understanding or vocabulary quantity affects the testing result. In subsequent studies, we will focus on how to design more effective experimental activities in order to enhance sentence learning.

The use of TPR combined with motions or gestures to learn different types of vocabularies made us realize that it is significantly more helpful in learning directional preposition than other methods. The reason might be that most learners can master the vocabularies of directional preposition easily. This implied that the application of TPR to learning is not simply by asking learners to perform more motions. Learning is an activity which requires not only performing gestures or motions but also the complete attention of the learners on the connectivity to the target content provided by the system. Lack of the connectivity to the target content will render the method a routine and hence, it becomes boring. This may not achieve the aim of building links between the gestures and the target language as well as improving the learning effects. This finding is consistent with other studies (Carels, 1981; Tellier, 2008, Johnson-Glenberg et al., 2014) which reported that using gestures is a good strategy for memory and learning enhancement as it gives further explanations on the reasons behind each action.

The content used to link vocabularies and example sentences to the gestures need to be designed by experts. In the proposed CKEL system, we used the motion recognition and object-oriented development protocol to design the new patterns of behavior. The designers can easily design new gesture patterns by simply detecting several basic movements from users and combining them together. Besides, there are open source codes available for recording the information of motions using Kinect. These can be utilized in designing the template for behavioral recognition. Then, machines can later be used iteratively in learning to enhance the accuracy of the template. Based on these two methods of behavioral recognition, in the future, we would make the system more convenient and user-friendly for the developer to customize the gestures corresponding to the target vocabularies.

From the results collected during the interview, helper and respondent were peers and they like to help one another to solve a problem. From our observation and the interview conducted, helpers usually give
suitable hints or help to respondents who cannot answer the question immediately, like speaking loudly with related gestures. Therefore, they felt very interested and enjoyed helping each other in peer sentence and vocabulary games. This is one of the advantages of collaborative learning in this kind of learning environment with kinesthetic support. Learning activities that combine TPR theory with kinesthetic recognition and collaborative activities make learners feel interested and entertained during the learning process. It also motivates the learners to get involved more actively in teamwork and joint learning in a team. In summary, the collaborative TPR learning is more suitable for learning a language than the traditional TPR method, because, in such environment, the dialogues and interaction among the team members successfully motivate the learning and enhances their confidence so that they become more aggressive. It is the aggressiveness that helps the learners to achieve higher goals in learning. The findings indeed compared individual TPR to collaborative TPR and showed that collaborative TPR is much more useful (Garcia, Asher, & Asher, 1988).

6. Conclusion

When learning the first language (L1), we have to understand the meaning behind the words then perform the appropriate action. According to this viewpoint, Asher proposed the use of TPR theory to learn the second language. In neuroscience, Macedonia and von Kriegstein (2012) claimed that when learning new vocabularies, the brain is activated on the same region that is responsible for the words with similar meanings in L1. On the one hand, if we use the same sensory input and behavioral information during learning, it enhances the association between new vocabularies and L1. Therefore, the acquired memory lasts longer. On the other hand, Vygotskiǐ (1978) believed that learners should develop social interaction with the teachers during the learning process; that this will connect the acquired language with daily life and becomes more relevant. Through the assistance from the teachers and the learners’ effort, it can help to promote the Zone of Proximal Development.

Therefore, in this research work, we presented a collaborative total physical response which combines TPR theory with collaborative learning with the use of Kinect to develop CKFL system in helping the students to learn in pairs. After the students had used this system for three weeks, they were given post-tests and interviews. The results were collected and analyzed statistically.

The total score and the vocabulary score of the experimental group in post-test were significantly higher than that of the control group I.
However, there was no significant difference in sentence score in post-test (meanwhile, the average scores of experimental groups were still higher than that of the control group I). This implies that the proposed system helps learners in pure vocabulary learning, but was not as effective in sentence learning. This is so because, to score high marks in sentence learning, one has to master a larger amount of vocabularies and grammatical rules. Hence, more attention should be directed towards the improvement of this system design so as to help learners to study sentences better.

Based on the opinions of Macedonia and von Kriegstein (2012), we might get similar results when the learning target is abstract vocabularies. For example, if a meaningful facial expression can solidify the meaning of an abstract vocabulary, it helps retain the memory too. In this research, we analyzed the learning effect of combining TPR with motion and gesture in learning a different kind of vocabulary. The result showed that the learning effect of directional prepositions is better than others because the task is easier and most learners are familiar with those vocabularies. Meanwhile, other results showed no significant difference in learning verbs, nouns and adjectives. Therefore, the result of our work is in concert with reports from Macedonia and von Kriegstein (2012). We can apply TPR in learning as long as we can create motions associated with the target vocabularies appropriately.

Based on the experimental result, applying TPR to learning is by subjecting learners to repetitive gestures and motions. For effective learning, a learner also needs to spend time and pay attention to the learning content provided by the system. Performing gestures or motion continuously without concentration on the material is just a meaningless body movement activity which cannot help to intensify the memory through enhancing the connection between gestures and L1, as suggested by the cognitive psychologists.

In this work, there is still room for improvement. For example, it would be better if we can (1) increase the number of subjects and try to balance the number of males with the number of female students; (2) upgrade the object recognition accuracy; (3) improve the system pronunciation and (4) automatically detect students’ pronunciation. All these would be addressed in future works.

**Disclosure statement**

No potential conflict of interest was reported by the authors.
Notes on contributors

Fan Wang is a professor of School of Education Intelligent Technology Jiang Su Normal University, Xu Zhou, Jiangsu, China. The major research interests of Dr. Wang include new social learning, Computer Supported Collaborative Learning.

Wu-Yuin Hwang is currently a distinguished professor of the Graduate Institute of Network Learning Technology, National Central University, Taiwan. The major research interests of Dr. Hwang include contextual learning, human-computer interaction for collaboration, multimedia tools/systems design and high level of cognition development and construction.

Yung-Hui Li is a professor of Graduate school of Network Learning Technology, National Central University, Taiwan, China.

Pin-Tse Chen is a Master Degree Candidate of Department of Computer Science and Information Engineering, National Central University, Taiwan, China.

ORCID

Fan Wang [http://orcid.org/0000-0003-0658-0563]
Wu-Yuin Hwang [http://orcid.org/0000-0001-5684-3590]
Yung-Hui Li [http://orcid.org/0000-0002-0475-3689]

References


## Appendix I. Vocabularies for kinesthetic learning

Motion mimic, directional preposition and adjective

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Category</th>
<th>Vocabulary</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encounter</td>
<td>Verb</td>
<td>Disappointed</td>
<td>Adjective</td>
</tr>
<tr>
<td>Run into</td>
<td></td>
<td>Massive</td>
<td></td>
</tr>
<tr>
<td>Explore</td>
<td></td>
<td>Insane</td>
<td></td>
</tr>
<tr>
<td>Agree on</td>
<td></td>
<td>Mouthwatering</td>
<td></td>
</tr>
<tr>
<td>Swoop</td>
<td></td>
<td>Agile</td>
<td></td>
</tr>
<tr>
<td>Banish</td>
<td></td>
<td>Impossible</td>
<td></td>
</tr>
<tr>
<td>In tune with</td>
<td></td>
<td>off-the-chart</td>
<td></td>
</tr>
<tr>
<td>Leap</td>
<td></td>
<td>Extended</td>
<td></td>
</tr>
<tr>
<td>Ditch</td>
<td></td>
<td>Asleep</td>
<td></td>
</tr>
<tr>
<td>Pick up</td>
<td></td>
<td>Strength</td>
<td>Noun</td>
</tr>
<tr>
<td>Jump</td>
<td></td>
<td>Exhaustion</td>
<td></td>
</tr>
<tr>
<td>Swim</td>
<td></td>
<td>Ankle</td>
<td></td>
</tr>
<tr>
<td>On foot</td>
<td></td>
<td>Infant</td>
<td></td>
</tr>
<tr>
<td>Twist</td>
<td></td>
<td>Appeal</td>
<td></td>
</tr>
<tr>
<td>Approach</td>
<td></td>
<td>Makeup</td>
<td></td>
</tr>
<tr>
<td>In back of</td>
<td>Directional preposition</td>
<td>Rib</td>
<td></td>
</tr>
<tr>
<td>In front of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On left side</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On right side</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Object recognition

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voyage</td>
<td>Grilling</td>
</tr>
<tr>
<td>Emperor</td>
<td>Commentator</td>
</tr>
<tr>
<td>Fortune</td>
<td>Paw</td>
</tr>
<tr>
<td>Entertainment</td>
<td>Feline</td>
</tr>
<tr>
<td>Flame</td>
<td>Hairspray</td>
</tr>
<tr>
<td>Hickory</td>
<td></td>
</tr>
</tbody>
</table>
Appendix II. Pre-test content

Part A 30%

Look at the pictures below. For each, you will hear one or two questions and four possible answers, A to D. Write down the letter of the best answer (A, B, C, or D).

Questions 1 and 2

1. ________ 2. ________

Question 3

3. ________

Question 4

4. ________
Questions 8 and 9

8. _______ 9. _______

Question 10

10. _______

Part B 30%
In part B, you will hear several questions or statements. Choose the best response or reply.

11. (A) She was very happy.
    (B) She saw it last night.
    (C) We both visited her yesterday.
    (D) Yeah, you can come over tonight.

12. (A) How long has she been sick?
    (B) I’m glad she finally built it.
    (C) It’s nice of her to help out.
    (D) Does she get paid much money?

13. (A) I just got a new one last week.
    (B) Give me one of those tissues.
    (C) That’s not very nice to say.
    (D) I don’t smell anything strange.

14. (A) It’s my favorite sport.
(B) Our seats are over there.
(C) The team will travel by bus.
(D) It’s the middle of the game.

15. (A) Yes, and he caught it, too.
(B) Yes, and he won the race, too.
(C) He is a great swimmer.
(D) It cost a lot of money.

16. (A) In that case, it’s an easy thing to solve.
(B) Yes, many people go hungry these days.
(C) True, it’s only seen in a few places now.
(D) You’re right; it’s no longer a big problem.

17. (A) Which forms do we need to fill in?
(B) Great, I enjoy watching those on TV.
(C) I’m glad they’re keeping things the same.
(D) Good, we need some big changes.

18. (A) Because it’s so interesting.
(B) Because it’s about science fiction.
(C) Because it uses lots of facts.
(D) Because it is about technology.

19. (A) The guides will give us all the maps.
(B) The plane ticket costs NT$10,000.
(C) We are going to Thailand.
(D) We will buy our own food.

20. (A) Why did you only get one?
(B) I can’t believe you forgot them.
(C) Great, share some with me.
(D) Two isn’t enough. Please get more.

**Part C 30%**

In part C, you will hear one or more conversations between a man and a woman, and then a question following each conversation. Choose the best answer to that question.

21. (A) The cloth used to make the couch.
(B) The price of the comfortable couch.
(C) The place where she bought it.
(D) The weight of the couch.

22. (A) Steve got a good grade on his test.
(B) Steve will go to a good college.
(C) Steve is acting differently.
(D) Steve is the same as always.

23. (A) Yes, he goes almost every day.
(B) Yes, he likes to run in races.
(C) Yes, he is one of the fastest.
(D) Yes, he hasn’t gone jogging recently.

24. (A) She will take a short nap.
(B) She will buy new shoes.
(C) She will stretch her legs.
(D) She will get her feet wet.

25. (A) Leave the country.
(B) Lead his country.
(C) Get married.
(D) Fight a battle.

26. (A) Her interests are the same.
(B) She is going to leave Taichung.
(C) Her interests have changed.
(D) She recently started painting.

27. (A) The play the school is performing.
(B) Others saying what they think of him.
(C) Having to perform in front of others.
(D) People that copy his own ideas.

28. (A) It was very interesting and useful.
(B) It was about the government.
(C) He said the speech made him think.
(D) He said the speech was very stupid.

29. (A) He believes they will get married.
(B) He thinks they will break up soon.
(C) He doesn’t think things will improve.
(D) He doesn’t want to give an opinion.

30. (A) He wants to go to the hospital.
(B) He wants to forget it happened.
(C) He wants the criminal to go to jail.
(D) He wants to forgive the person who did it.

Part D 10%
In part D, you will hear one or more short talks and then some questions about them. Choose the best answers to those questions.

Short Talk 1

31. (A) Floating in the water.
(B) In a popular museum.
(C) On a school playground.
(D) Behind a 7-Eleven.

32. (A) He wants people to laugh out loud.
(B) He wants people to rethink things.
(C) He wants people to pay him money.
(D) He wants people to become artists.

33. (A) Hofman shows the duck in public.
(B) Hofman created a fun art piece.
(C) The duck visits many countries.
(D) The duck won’t be displayed in Taiwan.

Short Talk 2

34. (A) They want their team to play better.
(B) They want Dilma Rousseff to stay.
(C) They weren’t happy with bus fares.
(D) They want to see the 2014 World Cup.

35. (A) The 2013 FIFA World Cup.
(B) The 2013 Confederations Cup.
(C) A reform of the Brazilian government.
(D) Changes to the Brazilian infrastructure.
Appendix III. Sentence test in the post-test

Given one picture, listen to each question and choose one correct answer

1. What is the boy doing?
   (A) He's running into the ditch.
   (B) He's leaping over the ditch.
   (C) He's taking a walk.
   (D) He's going for a ride.

2. What happened?
   (A) The ship is floating on the sea.
   (B) The men are swimming in the sea.
   (C) The lifeboat picked up all the survivors.
   (D) They are on the FEMA voyage.

3. What's wrong with the man?
   (A) He's exhausted after a long day.
   (B) He dropped asleep in the armchair.
   (C) He's leaping over the chair.
   (D) He is murmuring to himself.

4. How did the boy get to school today?
   (A) He went to school on foot.
   (B) He ran into the school.
   (C) He took a bus on his own.
   (D) He is carrying a backpack.

5. Where is the letter C?
   (A) It's above the letter A.
   (B) It's in front of the letter B.
   (C) It's under the letter A.
   (D) It's on your left side.

6. What did the boy get hurt?
   (A) He plays the ball on foot.
   (B) He heard someone knocking.
   (C) He gets something to eat.
   (D) His ankle got hurt.
7 Where is the tree?
(A) It’s on the house right side.
(B) It’s in front of the house.
(C) It’s at the back of the house.
(D) It’s above the house.

8 What’s the fish doing?
(A) It’s on the hook.
(B) It’s swimming out of the sea.
(C) It’s jumping out of the water.
(D) It’s moving in the water.

9 Who are these things for?
(A) They are for the insane.
(B) They are for a woman.
(C) They are for a family.
(D) They are for an infant.

10 Where is the airplane?
(A) It’s below the cloud.
(B) It’s in front of the cloud.
(C) It’s above the cloud.
(D) It’s in the cloud.

Listen to the question, and choose a right response.

1 A: Hey, you look beautiful today. B:_______
(A) Yeah, I put on a bit of makeup.
(B) Well, I don’t like it.
(C) Yes, I feel like to do it.
(D) I look forward to the gift, and hope it will be a hairspray.

2 A: I really want to go on a sailing voyage from North America to Europe. B:_______
(A) Going sailing is a good hobby for me.
(B) That is dangerous you might encounter big storms!
(C) It is twenty dollars. You won’t be disappointed!
(D) The band’s popularity is off-the-chart.

3 A: This band’s popularity is off-the-chart. It’s hard to get tickets to one of their shows. B:_______
(A) The show will take 3 hours at least.
(B) Yes, it’s hard to get there.
(C) The tickets cost me NT$500.
(D) To be honest, I got the ticket from my friend last week.

4 A: What happened to your ankle? B:_______
(A) I saw a massive hole in the ground yesterday.
(B) I want something to do first.
(C) I hit it on a table.
(D) It is easy to be happy.

(continued)
A: Hey, Peter. Your idea is not in tune with the rest of the group.
B: (A) Well, I need to make up my mind.
   (B) Yeah, I’m afraid we need more discussions.
   (C) It doesn’t matter.
   (D) Sorry for interrupting you.

6 A: You should explore the whole city, especially the White House.
B: (A) Good idea! I can’t wait to do it.
    (B) He broke the window of the house.
    (C) I found the people there are hospitable.
    (D) I ran into an old friend there.

7 A: Did you see the feline in the kitchen yesterday? B: (A) Yes, the flame is pretty good.
    (B) Yes, the home was in flames after a fire started in the kitchen.
    (C) Yes, I saw the cat leap.
    (D) Yes, well done.

8 A: Is there a bus stop near here? B: (A) Go straight for 1 block. It’s on your right side.
    (B) Don’t stop at the corner. It’s dangerous.
    (C) The bus stops suddenly.
    (D) It’s below the bus, not above it.

9 A: Amy, you always make mouthwatering meals for every guest.
B: (A) It’s a piece of cake because I am good at cooking.
    (B) I know I won’t be disappointed!
    (C) I don’t think you like the meals.
    (D) Maybe I need to water the plants.

10 A: What did you think of the commentator of last night’s baseball game?
B: (A) Well done. He liked looking at their big paw prints.
    (B) I don’t think he was in tune with his viewers.
    (C) I was rushing to ditch the plan.
    (D) I think he was too busy looking at the infant pandas.

Part E Blank filling
1 During the times of kings and ________, explorers had to sail to see the world.
   (A) noble (B) emperors (C) commentators (D) appeals

2 Gulliver feels ________, so he runs away, finds a boat and heads out to sea.
   (A) exhausted (B) extended (C) ________, explore (D) disappointed

3 The sailors went on an exciting ________ to an unknown island.
   (A) voyage (B) encounter (C) banish (D) leap

4 Hikers sometimes ________ bears while walking in the forest.
   (A) entertainment (B) infant (C) encounter (D) hairspray

5 My mom was grilling her homemade, ________, smoked pork sandwiches.
   (A) feline (B) mouthwatering (C) extended (D) ditch

6 A baseball player must have great ________ to hit a home run.
   (A) appeal (B) encounter (C) agile (D) strength

7 I went to bed early from ________ after a long day.
   (A) emperor (B) insane (C) exhaustion (D) extended

8 I missed the infant pandas because I could not get by my group to ________? what to see next.
   (A) in tune with (B) agree on (C) fill in (D) off-the-chart

9 The man made a ________ by starting his own successful company.
   (A) massive (B) ditch (C) flame (D) fortune

10 It’s ________ to drink and talk at the same time.
    (A) impossibly (B) possibly (C) impossible (D) possible

11 It is amazing how ________ tigers are even though they are so big!
    (A) agile (B) infant (C) twist (D) banish

12 Did you see all the birds ________ around their cages at the bird show?
    (A) twist (B) ditch (C) run into (D) swoop

13 I only use a little makeup and a little bit of ________ everyday.
    (A) appeal (B) hairspray (C) hickory (D) fortune

14 He spends the rest of his days speaking to the horses in his stables
    while everyone else considers him ________.
    (A) infant (B) disappointed (C) insane (D) entertainment

15 A: Have you ever been to visit Washington D.C on foot?
    B: No, I have only driven past it quickly by car, but my vacation was ________? so
    I am thinking of going again soon.
    (A) extended (B) entertainment (C) disappointed (D) banished