

Effects of Language Learning Game on Korean Elementary School Students

Kyusong Lee¹, Soo-Ok Kweon², Sungjin Lee¹, Hyungjong Noh¹, Jinsik Lee¹, Jonghoon Lee¹,
Hae-Ri Kim³, Gary Geunbae Lee¹

¹Department of Computer Science and Engineering,

²Division of Humanities and Social Sciences,

Pohang University of Science and Technology, Korea

³Department of English of Education,

Seoul National University of Education, Korea

{¹Kyusonglee, ²soook, ¹junion, ¹nohhj, ¹palcery, ¹jh21983, ¹gblee}@postech.ac.kr,
³hrkim@snue.ac.kr

Abstract

This paper introduces a preliminary study on affective effects of Spoken Dialog-Based Language Learning Game (DB-LLG) currently under the development in our research group. DB-LLG is an educational game designed for language learners to convey interactive conversations with in-game characters in interactive immersive environments, such as post office, library, shops, on the street, etc. Students have meaningful interactions with Non Player Characters to complete tasks in each game mission. We used the Wizard-of-Oz method as a preliminary technique to explore various affective effects resulting from students' dialog behaviors during the game. A total of 10 elementary school students, ranging in age from 12 to 13, participated in the experiment. The pre/post survey questionnaires were conducted to investigate the effect of DB-LLG approach on the affective factors of learning English. The results showed that DB-LLG significantly encourages students' satisfaction, interest, and learnability, while slightly decreasing students' pre-existing confidence.

Index Terms: language learning game, computer-assisted language learning, CALL, Wizard-of-Oz, WOZ

1. Introduction

It has been a widely accepted idea that one-on-one tutoring has a much better educational impact than group instruction [1]. Conceivably, one-on-one English conversation would be the best way to improve speaking ability. However, due to the expensive cost and lack of native speaker English teachers, among other things, many EFL (English as a Foreign Language) learners, like students in Korea, usually have limited opportunities to speak English in a natural language learning setting. Therefore, despite spending enormous time and energy in learning English, many Korean students are faced with difficulties in communication in English.

Game-based learning is currently getting attention as an upcoming educational method because of its benefits of high user motivation, attention, and interest [2]. One advantage of the Spoken Dialog-Based Language Learning Game (DB-LLG) approach is that the learner becomes more actively engaged with tutor in the game than in class. Our ultimate goal is to give students the opportunity to speak English in an interactive environment with high interest, motivation, and attention as in popular on-line games. Our group is currently involved in a project called POSTECH Immersive English Study (Pomy). The program allows users to exercise their visual, aural and tactual senses to receive a full immersion experience to

develop into independent EFL learners and increase their memory and concentration abilities to a greatest extent [3]. It employs spoken-dialog technology which enables computers to speak with humans. Moreover, we created an important in-game character called the Ghost Tutor which makes the game more interesting and educational.

The Wizard of Oz (WOZ) method is used for various purposes in spoken dialog systems, especially to build Language Model (LM) and Acoustic Model (AM). However, our research group developed the WOZ system not only to build LM and AM, but also to explore the educational effect of DB-LLG. Moreover, we believe that the study of users' behaviors would provide a special insight into designing our system. In this paper, we will introduce some design principles based on users' reaction, behavior, and affective effects.

2. Related Work

Tactical Language and Culture Training System (TLCTS) is one of the most successful systems. This targets the members of the U.S. military who need to acquire the basic communicative skills in Arabic and the knowledge of cultural differences in the given zone of operations like Iraq. TLCTS is also believed to be improved based on the user feedback [4]. With this respect, as a preliminary study, we assume that a user study is essential for initial prototyping of our system, which leads us to explore the positive effects of affective factors on language learning, such as user satisfaction and motivation.

Another system is the Spoken Electronic Language Learning (SPELL) system [5]. SPELL provides the opportunities for learning languages in functional situations and its key element has been developed to recognize grammatical errors, especially made by non-native speakers.

The DEAL, a spoken dialog system developed at KTH, using



Figure 1: Game Screen in the Post Office

language learning in virtual environment as TLCTS and SPELL do, focuses more on creating entertaining gameplay [6].

Since the ultimate objective of our system is to assist language learning processes in a virtual game environment, user feedback and data collection are extremely important for the development of the system. Therefore, we analyze users' behaviors and affective effects they demonstrate based on their behaviors. Additionally, we explore the effect of using a virtual English tutor, the Ghost Tutor, as we name it, which guides learners during the game.

3. POSTECH Immersive English Study

3.1. Speech and Language Technology

Our research group is currently developing spoken dialog systems for language learning. There are many unique difficulties associated with the language learning dialog system compared with the information seeking dialog system.

First of all, speech recognition is one of the biggest challenges for Korean EFL learners. Their perception and production of the sound is different from those of native speakers, resulting in numerous pronunciation errors. Therefore, we not only use the speech of Korean students to adapt the AM model, but also study an ASR combination technique to derive better performance.

Second, since language learners commit numerous and various errors, we needed to develop an intention recognition system that should be able to understand language learners' utterances in spite of these obstacles. To accomplish this purpose, we statistically inferred the actual learners' intention by considering not only the utterance itself but also the dialog contexts just as human tutors do [7].

The role of the dialog manager (DM) is to generate system responses according to the learner's intention and generate corrective feedback accordingly, if needed. We believe that the educational DM should be able to generate diverse responses to teach various useful English expressions. Unlike information seeking dialog management, we generate the system response by choosing one of n-best system responses

3.2. Scenario

For the domains that students were exposed to, we selected such domains as path-finding, market, post office, library, and movie theater to ensure having students practice conversations in everyday life setting. So far, five missions have been developed as a form of game and each mission consists of three main tasks, which again is composed of pre- and post-courses. Between missions, a path-finding pre-task is implemented. Before they started with the game, students were supposed to be familiar with the mission objective and the particular tasks, then, they were introduced with important vocabulary and useful expressions during the pre-course. For example, students should understand the meaning of some key words, including 'zip-code', 'insure', and 'over-night letter' and utter some key sentential expressions like "How much does it cost to send a package?" to successfully accomplish missions in the post office (Figure 1). These lessons were conducted during the pre-course before the main mission was introduced. The first mission that occurred in the post office was to send a camera to one's uncle in England. The package must be insured and delivered by the next week. In order to send the package, a student must fill in the zip-code properly. After completing each task, students were asked to review what they had learned in the previous task. Then, some comprehension questions were given to check whether or not

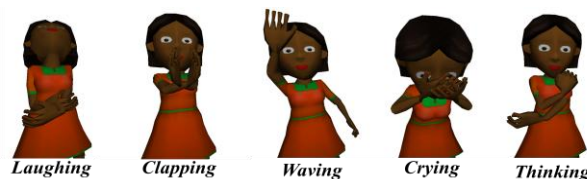


Figure 2: Character Animation.

they fully understood the content (e.g., "If the insured package is lost, what will happen?")

For essential items, students learned the same words and expressions three times: First, during the pre-course (preview), second, during the mission (main task in real situation), lastly, during the post-course (review). Through the repetition of the important expressions in different steps of learning, the effectiveness of second language acquisition could be maximized [8].

3.3. Teaching Strategies

3.3.1. Role of Ghost Tutor

The Ghost Tutor was a key character in the game who helped the students in various situations when they encountered difficulties and who guided them to move toward the next step throughout the whole procedures of the game so that the students could complete the mission successfully. The Ghost Tutor played several special roles, some of which will be introduced below (Figure 1).

First, the Ghost Tutor cast himself in the role of English tutor as a player and helped the students to use more appropriate words and expressions during the game. When a student produced ungrammatical utterances or made pronunciation errors, the Ghost Tutor provided both implicit and explicit negative and positive feedback in a form of recast, which was manifested as effective ways in the second language acquisition processes [9]. He provided a full sentence for students' fragmental responses. Although a student may speak without errors, the Ghost Tutor sometimes gave alternative examples of expressions to help students learn various new forms.

Second, the Ghost Tutor was the guide to the game. Even though the explanation of the mission goals was given clearly in the pre-course section, some students did not remember the mission objective during the game. In this case, the Ghost Tutor reminded them of the previous objective by going back to the particular objective.

Third, the Ghost Tutor was an intimate partner of the students. Whenever a student spoke to the Ghost Tutor, he provided with felicitous responses. Additionally, during a long interval without students' speaking, the Ghost Tutor would strike up a casual talk unrelated to the mission (e.g., "How is the weather today?"), which could give students a chance for bonus credits during the game, and therefore made them practice small talk style conversational skills.

3.3.2. Character Animation

All Non-Player Characters (NPC's) were provided with a number of communicative animations such as talking, laughing, waving, crying, thinking, getting angry, etc (Figure 2). The total number of animations was over thirty from which the operator could select one based on the response of a student. When the NPC spoke with animation, the student could recognize which character was currently talking. Communicative animations were expressions that made the

game more natural. The operator generated positive expressions such as clapping and laughing based on the user's correct answers, negative expressions such as crying and getting angry for incorrect answers.

4. Oracle Dialog System using WOZ

The affective effects can be importantly used to design the system parameters. We also wanted to explore the effect of the Ghost Tutor character we created on providing feedback for educational purpose. The Ghost Tutor unobtrusively helps the students correctly answer the system prompts in DB-LLG. Therefore, we used the WOZ technology to investigate various human-computer interactions in developing an oracle dialog system which will further guide our future development.

We implemented an operator's control panel as a tool to support the experiment and collect dialog corpus on-line. The operator saw exactly the same screen as the student's. The operator could generate speech and control the animation of the NPC and the Ghost Tutor. The background color of the dialog list turned from blue to red as the utterances were presented by the operator. In addition, the operator could select a task that fits a student's English level before the mission got started. If there was no proper answer to the student's utterance, the operator had to type the response directly in text field. The typed text was transformed into speech, and was added to the dialog list.

5. Experimental Design

We conducted an experiment at several elementary schools in Korea to investigate the affective effects of DB-LLG. A total of 10 elementary students were recruited by an English teacher of each school for the experiment and each of the participants was met in a quiet classroom with one hour interval. The proficiency level of the participants was equivalent to advanced EFL learners. All the participants were in the 6th grades, ranging from 12 to 13 years old.

Participants were instructed to speak with the computer characters who responded based on the input students provided to them. For data collection, the participant (user) and the operator (English teacher) were put in a separate room and played the game via network communication.

6. Results and Discussion

6.1. Experimental Results in Affective Effects

To investigate the affective factors before/after learning English through playing computer games, we asked students to complete a questionnaire in which they assessed their likelihood on a 5-point Likert scale (1 is least likely and 5 is most likely). Students were highly satisfied in using DB-LLG for language learning in general (Table 1). Students reported that using the Ghost Tutor was a successful teaching strategy for them to learn in a game-environment (e.g., "Computer's directions were clear for me to do the missions." and "The Ghost Tutor was helpful to guide me when I did not know what to say.") In comparison with the other questions, the questions about the NPC's outer appearance (e.g., "Avatar and the buildings on computer screen looked good.") and voice quality (e.g., "I like the computer's voice.") showed the lowest level of satisfaction. This suggests the need to develop a better computer graphics and a more natural text-to-speech system.

Table 2 indicates student's responses to the questions about the learnability in language learning game on pre- and post-test. One of the problems in game-based learning has been focused on the learnability issue whether or not actual learning

Table 1. *Satisfaction* in language learning after game (SD denotes standard deviation).

Question	N	Mean	SD ¹
Computer's directions were clear for me to do the missions.	10	4.9	0.31
It was clear to know what the mission was.	10	4.2	1.31
The Ghost Tutor was helpful to guide me when I did not know what to say.	10	4.8	0.63
I liked the computer's voice.	10	3.4	1.07
I could easily understand the computer's pronunciation	10	4.3	0.94
I think the avatar could understand me very well.	10	3.9	0.73
Avatar and the buildings in computer screen looked good.	10	3.9	0.87
I want to learn English though computer-game-based activities at school	10	4.8	0.42
It was fun as a whole.	10	4.7	0.48
Generally I enjoyed it.	10	4.1	0.5
Total		4.35	0.7

Table 2. *Learnability* in language learning before and after game. (* p < .05, ** p < .01)

Question	N	Mean		p
		Pre	Post	
I like to study English through computer at school.	10	4	4.2	0.44
I like to play computer games.	10	3.3	4	0.04*
I want to successfully finish a mission in computer games.	10	3.9	4.5	0.14
I can do better when I study with computer	10	3	3.9	0.001**
I like to listen in English through computer	10	3.6	3.8	0.50
I feel more comfortable when I study through Computer System rather than through books	10	3.5	3.9	0.37
I can more easily remember the words and expressions I learned with computer	10	2.7	3.2	0.09
I can concentrate more easily when I study with computer	10	3.1	3.9	0.003**
I can understand better when I study through computer rather than through books	10	3.2	3.5	0.43
Total		3.36	3.88	0.017*

Table 3. Students' *confidence* with English before and after game.

Question	N	Mean		p
		Pre	Post	
I am not afraid of speaking English	10	4	3.9	0.59
I am not afraid of making mistakes in English	10	3.1	3.3	0.16
I am not worried about doing English homework.	10	4.1	4.3	0.55
I think I can speak English better if I study harder	10	4.8	4.7	0.59
I can actively participate in English learning activities (role-play, presentation, game) in school.	10	4.8	4.7	0.59
I can speak in English with my friends at school.	10	4.3	4.1	0.34
I can buy things in English at the market.	10	4.1	3.7	0.10
I will help my friends if she/he has a problem in studying English	10	4.3	3.9	0.16
Total		4.2	4.07	0.34

does occur during the language learning game. Previous studies have shown that students are engaged just for enjoyment and pay attention only to the mission completion and game play. We also found similar behaviors in which some students answered "yes" in order to just complete a mission, even though they couldn't understand what the NPC said. However, during the interview most students

remembered well what they had learned during the game. The positive effect of computer assisted language learning was observed when students showed increase in the effects of learnability after playing the game.

Interestingly, however, students' confidence level in English did not show significant increase after the language learning game (Table 3). In fact, the score of all the items were lowered except for one item ("not afraid of doing English homework"). Students' prior confidence score was quite high considering they were all selected top level learners in their classes. However, in the interview conducted after the experiment, these learners mentioned that they had never used English for shopping at a market (e.g., "I can buy things in English at the market"), but that they had first thought it would be easy to buy things in English. However, they said it was more difficult than they had expected because speaking in a native-like environment made them nervous. Ironically, the learners all had had abundant exposure to English through classes at private institutions as well as their schools at the time of experiment, but the input they received from these sources was not meaningful enough to facilitate their speaking ability in an immersion environment like the one used in our experiment. In conclusion, abundant amount of quality (meaningful) input is more important to learn English than physical time and energy simply spent learning English. Given that engaging students only once in learning English using our DB-LLG system may not be sufficient to get significant increase in student's confidence in English, we aim to develop our system to an extent that various aspects of affective factors of language learning, including confidence, of course, will show significant increase after learners participate in the game-based language learning.

6.2. Design Guidelines Based on Students' Reaction

In this section, the design principles of DB-LLG based on student's oral production, behavior and interview after the game-like experiment will be described.

Principle 1: Familiarization. Make students familiar with interactive games. Some students failed in the interaction with the computer in speech modality. Many Elementary school students already play lots of computer games at home using mainly the keyboard and the mouse, hence during the experiment, students were focusing only on the keyboard and the mouse to control their game characters without listening and speaking. Therefore, we developed voice commands to control user's avatar (e.g., "Go straight", "Turn left", and "Turn right") to make learners familiar with speech modality.

Principle 2: Proper response time. It was also very important for the operator to give response in the right time. Some advanced level students complained that the operator's response time was too long and were not patient enough to wait until the operator provided responses. On the other hand, low level students complained about the short response time. If the Ghost Tutor gave an incomplete answer while a student was still thinking, it made the student frustrated at his inability to answer in time. Thus, the operator controlled response time according to learners' English level.

Principle 3: Educational Purpose. Do not make a game that students play only for enjoyment, neglecting the educational purpose. Although they strongly wanted to accomplish the missions with a high success rate, students sometimes simply looked for shortcuts to finish the missions without having a conversation with the computer. For example, they just said "Yes" even though they did not understand exactly. However, we will be able to design the game to prevent this problem by checking their understanding strictly.

Principle 4: Transcription of Speech. Vary the usage of the on-screen speech bubbles according to students' proficiency level. For low level students, we showed the speech bubbles more frequently and for a longer duration to assist their understanding of the spoken dialog. However, we found them to be distracting for high level students who wanted to concentrate on listening. Hence, the speech bubbles were provided only when students had problems with understanding the utterances provided by the operator.

7. Conclusions

In this paper, we have investigated the affective effects of DB-LLG using WOZ technology. This study lays the foundation for future work on DB-LLG using spoken dialog system. Our results proved DB-LLG would help students improve their satisfaction and learnability in English. Moreover a number of design issues were found based on students' interviews and reactions. For further study, we plan to investigate cognitive effects of DB-LLG and collect more dialog corpus with plentiful contents for a longer period of time. Moreover, we will explore users' reactions using spoken dialog system instead of WOZ.

8. Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0019523), and the Industrial Strategic technology development program, 10035252, development of dialog-based spontaneous speech interface technology on mobile platform, funded by the Ministry of Knowledge Economy (MKE, Korea), and was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Knowledge Economy (MKE).

9. References

- [1] Bloom, B. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring, *Educational Researcher*, 13(6), 4-16.
- [2] Neville, D., Shelton, B., & McInnis, B. (2009). Cybertext redux: using digital game-based learning to teach L2 vocabulary, reading and culture. *Computer Assisted Language Learning*, 22, 409-424
- [3] Lee, S., Noh, H., Lee, J., Lee, K., & Lee, G. G. (2010). POSTECH Approaches for Dialog-based English Conversation Tutoring. *Proceedings of the 2010 APSIPA annual summit and conference*, Singapore, December 2010.
- [4] Johnson, W.L.; Wang, N.; & Wu, S. (2007). Experience with serious games for learning foreign languages and cultures. *Proceedings of SimTecT*, June, 2007.
- [5] Morton, H., & Jack, M.A. (2005). Scenario-based spoken interaction with virtual agents. *Computer Assisted Language Learning*, 18, 2005, 171-191.
- [6] Brusk, J. Wik, P., & Hjalmarsson, A. (2007). DEAL: A Serious Game for CALL Practicing Conversational Skills in the Trade Domain. *The Proceedings of SLATE-Workshop on Speech and Language Technology in Education*, Pennsylvania, USA.
- [7] Lee, S., Noh, H., Lee, J., Lee, K., Lee, G. G., Sagong, S., & Kim, M. (2011). On the Effectiveness of Robot-Assisted Language Learning, *ReCALL Journal*, 23(1), 25-58
- [8] Gass S. M., & Varonis, E. M. (1994). Input, Interaction, and Second Language Production. *Studies in Second Language Acquisition*, 16, 283-302.
- [9] Long, M., Inagaki, S. & Ortega, L. (1998). The Role of Input and Interaction in Second Language Acquisition. *The Modern Language Journal*, 82, 357-371.