2005 IEEE/RSJ International Conference on Intelligent Robots and Systems

Learning by Doing: Space-Associate Language Learning Using a Sensorized Environment

Yoshifumi Nishida
Digital Human Research Center, AIST
2-41-6 Aomi, Koto-ku, Tokyo 135-0064, Japan
y.nishida@aist.go.jp

Makoto Hiramoto
Tokyo University of Science
2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
m.hiramoto@aist.go.jp

Fusako Kusunoki
Tama Art University
2-1723 Yarimizu, Hachioji, Tokyo 192-0394, Japan
kusunoki@tamabi.ac.jp

Hiroshi Mizoguchi
Tokyo University of Science
2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
hm@rs.noda.tus.ac.jp

***Abstract*— This paper proposes a Learning-by-Doing (LBD) system for assisting learners of English as a foreign language, by providing educational content generated based on real-world behaviors in a sensorized environment. The LBD system consists of an ultrasonic location sensing system developed by the authors and an educational content server. The location of the student is monitored by ultrasonic sensors embedded around the learning room. The system voices English verbs related to objects around the student so that the student can obtain the language skills for expressing various activities of everyday life in English by performing real-world behaviors. The system was implemented experimentally in an English language school. Six subjects participated in an experiment evaluating the educational efficacy of the proposed system. Recall rates of learned verbs were compared among three methodologies, "Learning by text," "Learning by pictures and text," and "Learning by doing". Recall rates were measured at the following intervals: soon after the learning task, 2 days after, 4 days after, 7 days after, and 14 days after the learning task. The results indicate that the proposed LBD system increased the recall rate of learned words by 113% 14 days after the learning task. Overall, the system significantly increased the average recall rate by 102% in comparison to the other learning methodologies.**

I. Introduction

Learning by Doing (LBD), or learning through experience, is not a new method of learning. Rather, it is a natural method of learning that has been practiced throughout human history. The results of several psychological experiments support the theory that the human brain memorizes skills most effectively during LBD tasks. This concept is known, for example, as episodic memory [1]. A learning technique based on a similar concept was proposed 40 years ago by Asher [2]. According to Asher's methodology, called the "Total Physical Response (TPR)", a student learns a new language by responding physically to a teacher's commands. A TPR based virtual reality (VR) technology was proposed by Rose [3]. In addition, Glove-Talk II, which can translate hand motion into text information using a data glove was proposed by Fels [4], although its primary purpose is not to assist foreign language learners.

If the environment around a student could detect the location of the student and somehow recognize his or her behavior, as is possible using ubiquitous computing technologies [5] such as location sensors [6], the environment could be used to assist in the learning of a foreign language by associating language concepts with location-specific behaviors or information and by supporting episodic memorization. The authors have developed an ultrasonic location sensing system capable of detecting an individual's behavior [7] as a basic component of an intelligent environment. This paper proposes a Learning-by-Doing (LBD) system that assists a learner of a foreign language by locating the student and recognizing his or her real-world behavior.

II. Learning by Doing

Memory is classified into semantic, procedural, and episodic memory. Semantic memory contains the meaning of language and concepts. Procedural memory contains memories learned through the actions of the body, but not language elements. Episodic memory is memory associated with a performed action.

For example, if we are asked "What did you have for breakfast?", we will remember the meal eaten this morning and will answer, "I had boiled rice, soup, and sliced salmon". At this instant we simultaneously remember this image, this location, and associated behavior rather than remember only the image itself. Thus, episodic memory is egocentric and has time, spatial, and context restrictions. General research on memory has come to be investigated in many clinical experiments. Furthermore, the place neuron [8] is specifically involved with the mechanism of episodic memory. The importance of place cognition is said to be the synthesis of stimulated feelings that consist of multiple dimensions including light and sound.

In the present report we assume that a stronger relationship between place cognition and episodic memory will result in improved memorization for students. Herein, this learning method is referred to as Learning by Doing (LBD) and has been in use since ancient times. More specifically, in this

Fig. 1. The Learning-by-Doing system constructed in an English language school

paper, LBD refers to learning by interaction between an individual and the contents of a room monitored by sensors. As the individual reacts to the contents of the room and the spoken sounds from the LBD system, he or she associates an object or location and its appearance with the related English word. This paper describes the technical aspects of this system and the results of experimental trials in an English language school to demonstrate its practical effectiveness.

III. Experimental System in an English Language School

A. Learning-by-Doing System

Figure 1 shows the developed Learning-by-Doing system and Figure 2 shows the system configuration. The Learning-by-Doing system consists of both hardware and software components. The former is the ultrasonic 3D tag system for measuring the position of a tag, and the latter is the Learning-by-Doing content system for monitoring the foreign language content. By sending an ultrasonic wave to receivers, the ultrasonic tag system calculates the 3D position of an object or a person by an attached transmitter. Macromedia Flash™ [9] content can be changed according to this position information. Specifically, the context, in relation to the position of an object or a person, is monitored and the associated contents are displayed.

It is important for a human support system to allow an individual to concentrate on his or her primary task (e.g., learning a new language), without imposing unnecessary tasks (e.g., operation of a computer mouse, etc) [10]. A support system must have a natural interface with which a person

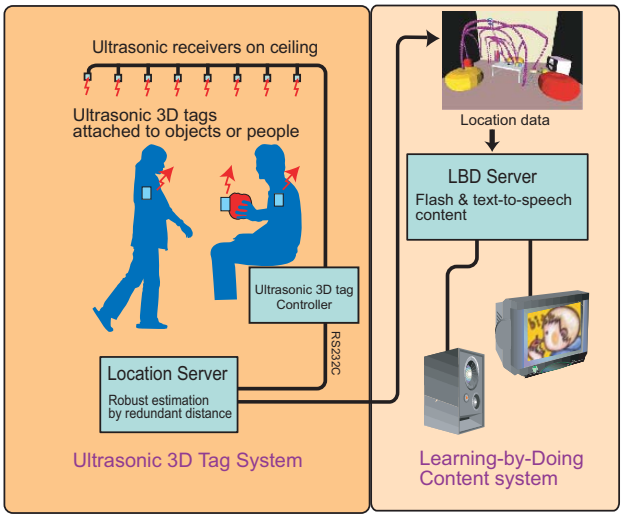

Fig. 2. Learning-by-Doing system configuration

can input their information without performing any additional conscious actions.

We developed a natural interface to facilitate the easy and efficient use of this system while not requiring initial training for the user. Users were asked to participate without any initial training or knowledge of the system. All participants indicated that they found the system easy to manipulate. In addition, the removal of any additional interfacing stages between the user and the system assured the maximum involvement and successful use of this system. The advantage of the Learning-by-Doing system is that individuals can learn a foreign language naturally by behaving in the same manner

as they do in everyday life. The system used for capturing this behavior is called the ultrasonic tag system.

B. Learning-by-Doing System Hardware

The authors have developed a three-dimensional ultrasonic tagging system [11] that can be used as a location sensor. The system calculates the position of a tag by trilateration or multilateration. Examples of the tags used in the system are shown in Fig. 3. The LBD system was constructed based on this technology.

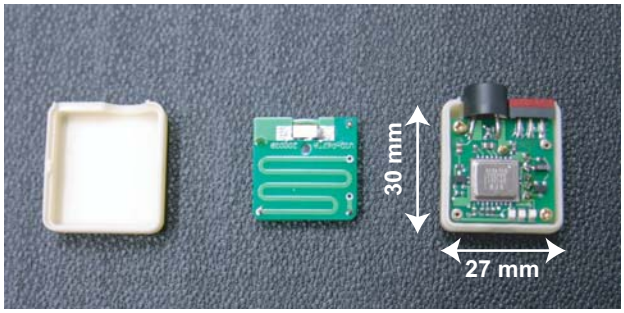

Fig. 3. Ultrasonic 3D tag

The system consists of an ultrasonic receiver, an ultrasonic transmitter, a time-of-flight gauge, a network, and a personal computer. The ultrasonic receiver receives ultrasonic pulses emitted by the ultrasonic transmitter and amplifies the received signal. The time-of-flight gauge records the travel time of the signal from transmission to reception. The network synchronizes the system and collects time-of-flight data from the ultrasonic receiver. The position of an object is calculated based on more than three time-of-flight results. The sampling frequency of the proposed system is 50 Hz.

The authors have also developed a portable ultrasonic tag system [12] facilitating the easy construction of an ultrasonic location system in a variety of environments. The developed portable system allows users to install the proposed Learning-by-Doing system quickly and cost effectively.

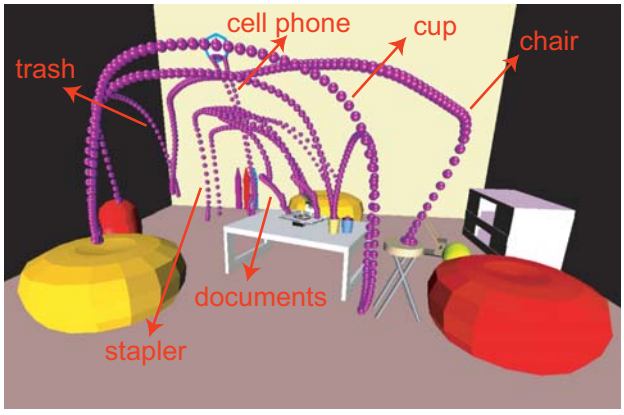

Fig. 4. Measured trajectory for the successive movement of several objects

Figure 4 shows the measured trajectory for an individual moving a cup to the chair, to the floor, and to the desk. These findings demonstrate that the system can robustly measure the positions of the objects in most locations of the room regardless of occlusion by the user's body occurring during movement [11]. It was possible to maintain a high sampling frequency by dynamically selecting which transmitters to track. For example, a transmitter can be attached to an individual's wrist, and the system can select and track transmitters in the vicinity of the wrist. The 3D position value of an object or a person is sent to the Flash content server. Flash content is commonly located on Internet websites.

The Learning-by-Doing system displays Flash content on a screen according to the position of the detected transmitter. The Flash content changes synchronously as the position of the transmitter changes. Since the software component of this system can be distributed easily, and the cost of installing the portable Learning-by-Doing system hardware in various locations is very low, the system can be easily accessed and used by different language schools.

C. Learning-by-Doing Methodology

Based on discussions with the teachers at the English language school, we devised the following standard procedure for learning a new language with the developed LBD system:

1) A student moves in the room,
2) or the student touches an object in the room.
3) The system voices a verb related to the behavior or the object.
4) Then, a teacher who is present in the room asks the student a question.
5) The student answers the question using the voiced verb.
6) The teacher corrects any mistakes the student makes at that time.

In this methodology, the student is expected to create complete sentences from sentence components such as verbs and is not expected to repeat the system's utterance.

The position of objects and users are tracked by the system, while the educational content is generated based on the detected position and is presented instantaneously. For example, when the user picks up a wine bottle, the system vocalizes "spill" or "stain". Or, when the user picks up a model of ice cream, the system voices "melt". Similarly, when the user approaches the fireplace, the system voices "burn" or "you are burned" while showing helpful visuals on a projector screen, as shown in Fig. 6.

IV. Experimental Results

A. Purpose of the Experiment

The present experiment was carried out to verify the educational effectiveness of the LBD method for foreign language study. Subjects were required to learn a total of 150 words in English using three learning methods. Following the learning task, an investigation was conducted to determine which learning method was most effective at facilitating memorization. The learning method with the highest recall rate was determined to be the most effective.

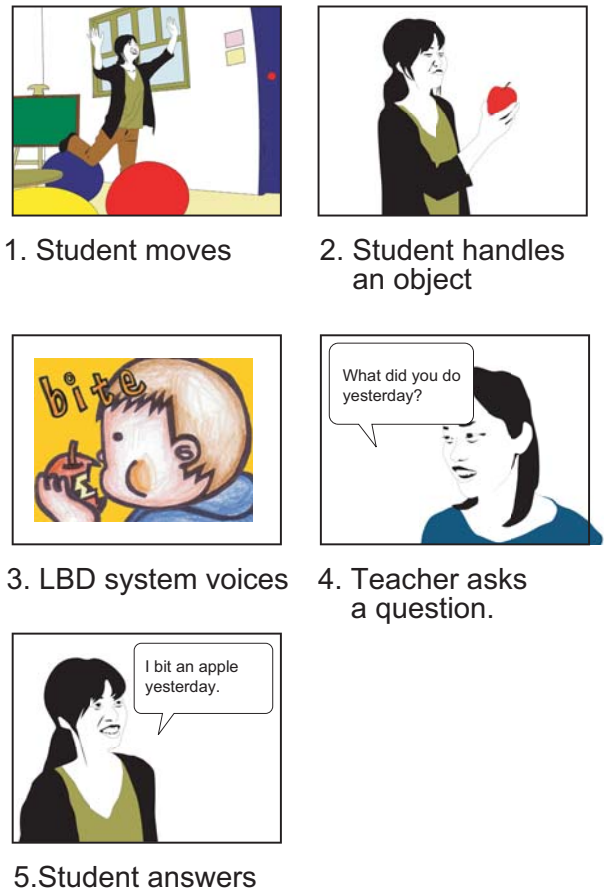

Fig. 5. LBD methodology

Fig. 6. Examples of educational visuals

B. Experiments and Learning Tasks for Students

Experiments investigating the effectiveness of the LBD system were conducted with the participation of 6 Japanese students aged 22 to 24. In the first task, each subject was required to learn 50 English verbs within 20 min using texts only, then a 10 min free recall test was administered. In the second task, each subject was required to learn 50 English verbs (different from those in the first test) using visuals and text, followed by a free recall test. In the final task, each subject was required to learn 50 verbs (again different from the previous test) using the LBD method, followed by a free recall test. The free recall test evaluated the number of verbs that the students could recall. In this experiment, when the subject approached a specific location, as shown in Figure 8, LBD contents were displayed on the screen. Fig. 7 shows the scene that the subjects actually experienced.

Fig. 7. Experiment using the Learning-by-Doing system

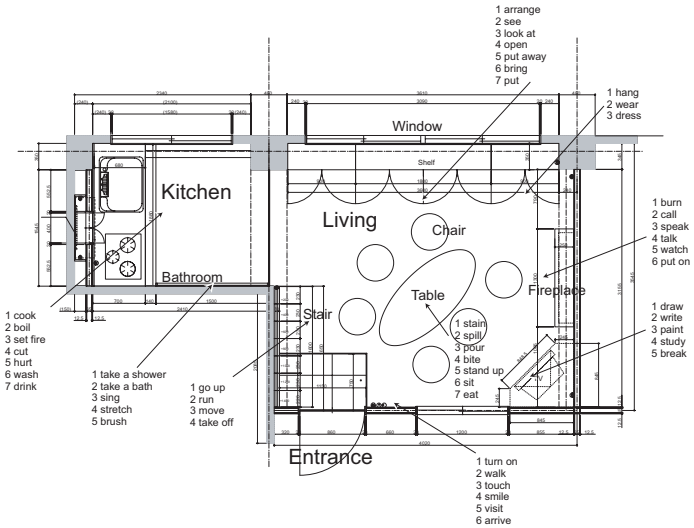

Fig. 8. Arrangement of the experimental room in the language school

C. Experimental Results and Evaluation

1) Comparison of the difficulty of verbs: The knowledge level for each of the 6 subjects was assessed by a questionnaire conducted after the experiment. Students were asked how many of the presented verbs they already knew in order

to determine the difficulty of the presented verbs. Figure 9 shows the results of the survey. Data from the "Learning by text", "Learning by pictures and text", and "Learning by doing" groups are displayed from left to right. We compared the difficulty of verbs by performing a paired Student's t-test between each group. The t-test is a commonly used method for determining whether the difference between sample means is statistically significant. In Fig. 9, p indicates the p-value (probability) associated with the t-test. These results demonstrate that there is no significant difference among the 3 groups of verbs used in the 3 learning methods. Therefore, we consider the difficulty of the verbs to be almost the same for each of the 3 groups.

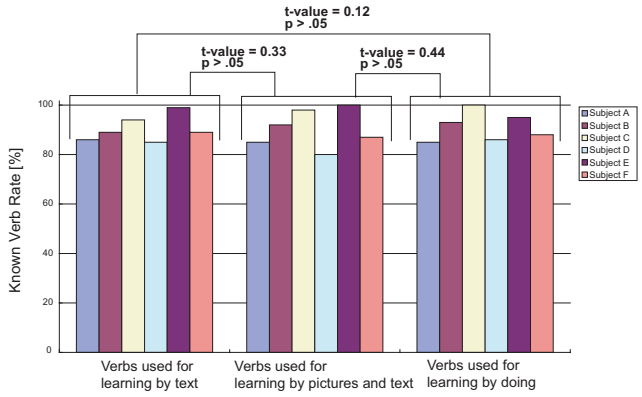

Fig. 9. Comparison of the rate of known verbs

2) Comparison of recall rates among the three methods: Figure 10, Figure 11, and Figure 12 compare the recall rates of the 6 subjects for the three learning methods after the following intervals: soon after the experiment, 7 days after the experiment, and 14 days after the experiment. The horizontal axis shows the learning method, and the results are presented in three groups, with the vertical axis indicating the recall rate (%). Five recall rates are presented for each experimental method. Figure 10 shows that there was no significant difference in recall rates among the three methods on the day of the experiment. However, in Fig. 11 and Fig. 12 we can see that there were significant differences in recall rates between the "Learning by doing" method and the other two methods at 7 and 14 days after the experiment. In Fig. 12, the average recall rate for the "Learning by text" method is 18.33%, that of "Learning by pictures and text" is 20.0%, and that of "Learning by doing" is 40.83%. This indicates the LBD method increases recall rates by 113% when compared to other methods.

3) Change in recall rates over time: We conducted a further experiment to clarify the relationship between the verbs used for the three learning methods and the change in recall rates over time. The experiment was conducted on 2 subjects, who did not participate in the above experiments. In this experiment, we used only the "Learning by text" method to eliminate the effects of the differences in learning methodology. The horizontal axis of Fig. 13 represents the

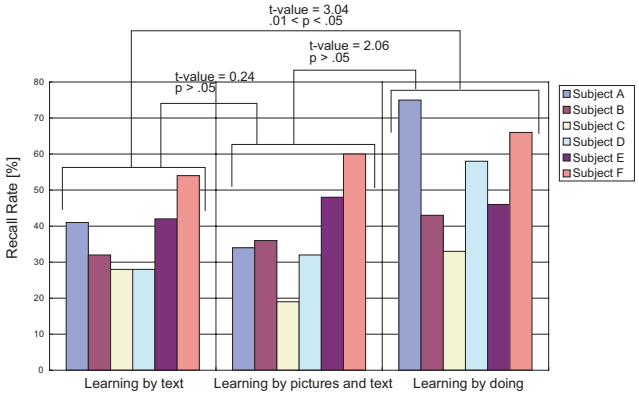

Fig. 10. Comparison of recall rates soon after the experiment

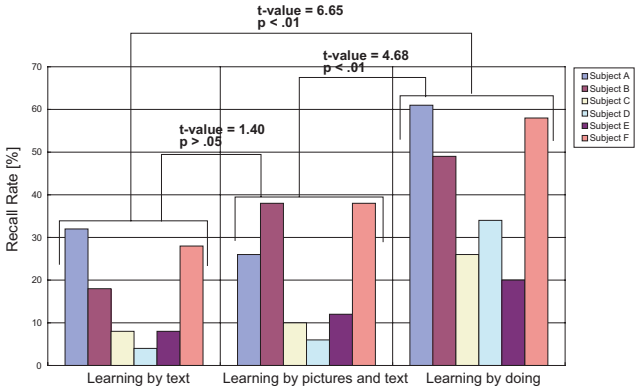

Fig. 11. Comparison of recall rates 7 days after the experiment

categories, "Verbs used for learning using text", "Verbs used for learning using picture and text", and "Verbs used for LBD", from left to right. Each bar shows the average recall rate of the 2 subjects. The results indicate that the recall rate deteriorated over time. These findings also indicate that there is no significant difference in the change of recall rates among the 3 groups of verbs. The t-value between the 1st and 2nd group = 2.93, and p = 0.042 ($p > 0.01$). The t-value between recall rates in the 2nd and 3rd group = 1.68, and p = 0.16 ($p > 0.01$). The t-value between recall rates in the 1st and 3rd

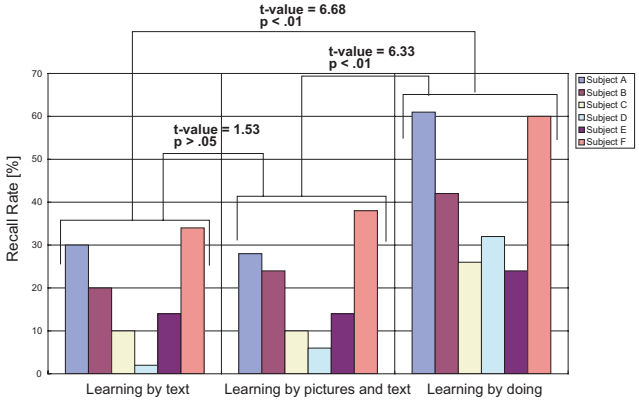

Fig. 12. Comparison of recall rates 14 days after the experiment

group = 3.03, and p = 0.038 ($p > 0.01$).

Figure 14 shows the relationship between the methodology and the change of recall rates over time. For all subjects, words learned during the LBD method were clearly more easily recalled during the free recall tests. In total the LBD increased the average recall rate by 102% in comparison to the other methodologies. The forgetting rates were 52% using text alone, 48% for visuals and text, and 24% for LBD. Specifically, the proposed system presented a harder-to-forget studying method than the other methodologies. The t-value between forgetting rates in the 1st and 2nd group = 1.38, and p = 0.23 ($p > 0.05$). Conversely, the t-value between forgetting rates in the 2nd and 3rd group = 28.55, and p = 8.94×10^{-6} ($p < 0.01$), and the t-value between forgetting rates in the 1st and 3rd group = 17.12, and p = 6.82×10^{-5} ($p < 0.01$). Since no significant difference was found for the difficulty of memorizing the three groups of verbs, it is considered that the difference shown in Fig. 14 is a result of the effects of the methodologies.

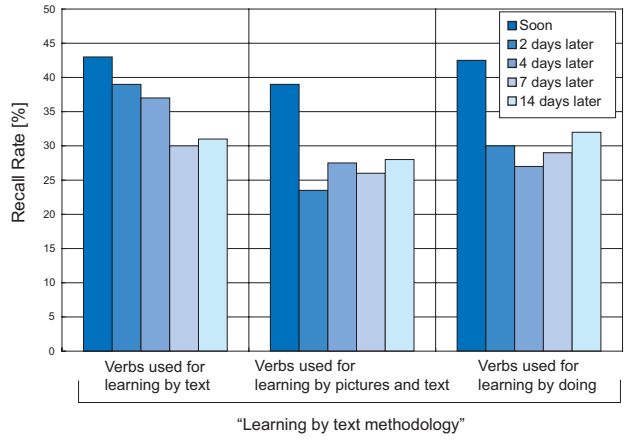

Fig. 13. Relationship between verbs and the change of recall rates over time

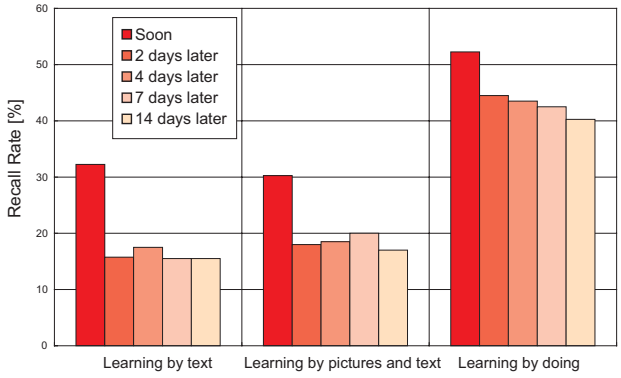

Fig. 14. Relationship between the methodology and the change of recall rates over time

V. Conclusion

This report presented a Learning-by-Doing method for teaching a foreign language. The LBD system consists of an ultrasonic location sensing system that controls the presentation of context-specific material to students. The effectiveness of the system was evaluated using an experimental LBD system installed in an English language school. Following the experiments conducted with the participation of 6 Japanese students, it was shown that the proposed system increases the recall rate of learned words by 113% 14 days after the learning task. In total, the system increased the average recall rate by 102% in comparison to the other methodologies. Therefore, the proposed system was more effective for learning verbs. The forgetting rates were 52% for using text alone, 48% for visuals and text, and 24% for LBD. Specifically, the proposed system presented a harder-to-forget studying method than the other methodologies.

Our future work also will also focus on improving the learning process using knowledge of databases and learning strategies, such as CYC [13] or WordNet [14].

Acknowledgements

The authors would like to thank Mr. Katsuhiro Shimada at TESS English Learning School for his cooperation during the experiments, and Ms. Mayuko Matsuya, an art student at Tama Art University for designing the visual educational content.

References

[1] E. Tulving, "How many memories are there?," *American Psychologist*, Vol. 40, pp. 385-398, 1985

[2] J.J. Asher, *Learning Another Language Through Actions (6th Edition)*, Sky Oaks Productions, Inc., 2003

[3] H. Rose, "Design and Construction of a Virtual Environment for Japanese Language Instruction," http://www.imprintit.com/hrose/, 1995

[4] S.S. Fels, G.E. Hinton, "Glove-Talk II: A neural network interface which maps gestures to parallel formant speech synthesizer controls," *IEEE Transactions on Neural Networks*, Vol. 9, No. 1, pp. 205-212, 1998

[5] U. Hansmann, L. Merk, M.S. Nicklous, T. Stober, *Pervasive Computing Handbook. The Mobile World*, Springer-Verlag Telos, 2001

[6] J. Fraden, *Handbook of Modern Sensors: Physics, Designs, and Applications*, AIP Press, 2003

[7] Y. Nishida, K. Kitamura, H. Aizawa, T. Hori, M. Kimura, T. Kanade, H. Mizoguchi, "Real World Sensorization for Observing Human Behavior and Its Application to Behavior-To-Speech," *Proceedings of 2004 ACM International Conference on Intelligent User Interfaces (IUI 04)*, pp. 289-291, 2004

[8] T. Kobayashi, H. Nishijo, M. Fukuda, J. Bures, and T. Ono, "Task-Dependent Representations in Rat Hippocampal Place Neurons," *J. Neurophysiol* Vol. 78, pp. 597-613, 1997

[9] http://www.macromedia.com/

[10] D.A. Norman, *Invisible Computer: Why Good Products Can Fail, the Personal Computer Is So Complex and Information Appliances Are the Solution*, The MIT Press, 1998

[11] Y. Nishida, H. Aizawa, T. Hori, N.H. Hoffman, T. Kanade, M. Kakikura, "3D Ultrasonic Tagging System for Observing Human Activity," *Proceedings of IEEE International Conference on Intelligent Robots and Systems (IROS 2004)*, pp. 785-791, 2003

[12] A. Nishitani, Y. Nishida, T. Hori, H. Mizoguchi, "Portable 3D Ultrasonic Tag System Based on a Quick Calibration Method," *Proceedings of 2004 IEEE International Conference on Systems, Man and Cybernetics (SMC2004)*, pp. 1561-1568, 2004

[13] D.B. Lenat, "Cyc: A Large-Scale Investment in Knowledge Infrastructure." *Communications of the ACM*, No. 38, no. 11, 1995

[14] C. Fellbaum, *Wordnet: An Electronic Lexical Database,* MIT Press, 1998