Mutual Learning for Second Language Education and Language Acquisition of Robots

Akihiro Yorita and Naoyuki Kubota

Abstract. Recently, language education has great demand from elementary school to adults. The robot is used as a teaching assistant in Robot-Assisted Language Learning (RALL). It is very effective to use robots for language education. But robots have some problems. One of the problems is to get bored with interacting with robots. This paper discusses the role of robots based on mutual learning in language education. Next we explain the concept of self-efficacy using evaluation for learning condition of robots. We propose a conversation system for language education. The essence of the proposed method is mutual learning of humans and robots. The experimental results show the applicability of the system used for education.

Keywords: Human-Robot Interaction, Robot Assisted Instruction, Second Language Learning, Language Acquisition, Self-efficacy.

1 Introduction

English education is being done more enthusiastically than ever. Japanese government decided to introduce English education to elementary school and some Japanese companies use English as an official language. We can talk with native speakers on the video call recently, but when we talk with them we cannot talk very well because we may be nervous. Therefore we need to practice conversation with robots.

In robot-assisted language learning (RALL) shown in Table 1 [1-5], a humanoid robot named Robovie has taught English at an elementary school in two weeks [1]. It is an effective way to motivate students learning English, although it is less effective than educational software.

Akihiro Yorita · Naoyuki Kubota
Tokyo Metropolitan University
e-mail: `yorita-akihiro@sd.tmu.ac.jp, kubota@tmu.ac.jp`

In Korea, RALL has been the major way to learn English. It is also called r-learning. The robot helps human teachers and does role-playing with students. Robot IROBI is used as a home robot and teaching assistant in a classroom [2].

The robot was used to examine the learning effect on children. Using robots is good compared with using books and tapes, or computers. Robots are also used as native teachers in rural areas [3]. As the teachers prefer not to leave big cities, the students have few opportunities to take classes by them.

Table 1 Robot-Assisted Language Learning

Type	Teaching Assistant				Learning Companion
Robots	IROBI [2]	Telepresense robot [3]	Mero and Engkey [4]	Humanoid robot [5]	Robovie [1,6]
Aim	Interest	Interest	Listening	Cheering	Motivation
	Concentration	Confidence	Speaking	Conversation practice	Long-term Interaction
	Achievement	Motivation			
Case	In class				Recreation hour
Country	Korea			Taiwan	Japan

People regard virtual agents and robots as intelligent life [19]. They appear intelligent at first, but humans discover patterns gradually. Then people may stop communication. Therefore it is difficult to realize long-term interaction.

In [6], pseudo-development and confidential personal matters enable the robot to do long-term interaction. Here the robot changed interaction patterns along with each child's experience, the robot seems as if it learns something from the interaction. In fact, the robot can learn words [7]. In order to adapt to an open environment, a robot will have to learn the language dynamically. The system for Noun Concepts Acquisition (SINCA) forms utterances about an image, but SINCA is a language acquisition system not robot [8].

We propose a method of learning words between a robot and a human. Our target is to develop the method using it in everyday life. Especially learning and growing with robot is important.

Human symbiotic robots are utilized in various fields. In welfare, Paro and ifbot are representative robots [20,21]. In entertainment, AIBO is the most famous pet robot and miuro plays music with dance adapting to human preference [22, 23].

In education, students usually build a robot. Sometimes a communication robot is used as a teacher or a teaching assistant as noted above, but not a friend who learns together. In Japanese animation "Doraemon" [27], Doraemon is sent back from future to look after Nobita. It is preferable to call them human remediation robots instead of human symbiotic robots. Humans are not good at repetitive tasks, but robots can do. Then it is useful that we give a desire to learn English to the robot and the robot encourages us to learn. Therefore the robot controls his desire by self-efficacy and changes its utterances by the value of self-efficacy so that humans continue to learn. Here, we define self-efficacy for language learning and examine the change during conversation.

First, we explain the background of robots used in education. Next, we explain the education system, and explain computational intelligence technologies. Furthermore, we discuss how to support students using robot partners. Finally, we discuss the future vision toward the realization of educational partner robotics.

2 Robot Edutainment

2.1 Various Roles of Robots in Education

Various types of robots have been applied to the fields of education with entertainment (Edutainment). Basically, there are three different aims in robot edutainment. One is to develop knowledge and skill of students through the project-based learning by the development of robots (Learning on Robots). Students can learn basic knowledge on robotics itself by the development of a robot [9,26]. Lately, low cost 3D printers have been developed and students can easily make robots that they hope [24]. The next one is to learn the interdisciplinary knowledge on mechanics, electronics, dynamics, biology, and informatics by using robots (Learning through Robots). The Local Educational Laboratory on Robotics proposes that it is good to learn about nature in primary school and to think and understand humans in secondary school with using minirobots [25]. In the Robockey Cup, the students studied how to use motors, circuits, microcontrollers, and so on [10]. Moreover, to make a humanoid robot gives opportunities to understand voice recognition and image processing for communication between humans and robots [11]. The robot is also useful for children with autism. The last is to apply human-friendly robots instead of personal computers for computer assisted instruction (Learning with Robots).

A student seldom shows physical reactions to a personal computer in the computer-assisted instruction (CAI), because the student is immersed into 2-dimensional world inside of the monitor. However, a student aggressively tries physical interactions to a robot, because the robot can express its intention through physical reactions.

We showed the effectiveness of the learning with robots in the previous works [12,13]. A robot partner in educational fields cannot be the replacement of a human teacher, but the replacement of a personal computer. Of course, the robot also should play the role of a personal computer. Therefore, we propose the concept of robot-assisted instruction (RAI) to realize the style of education based on the learning with robots. A robot can be not only an assistant, but also a partner or collaborator in RAI. A personal computer is useful to collect, access, edit, and store data, but agent-like communication capability is low in a personal computer. Therefore, a robot partner can be replaced with a personal computer.

2.2 Robot Partners for RAI

We have developed PC-type of physical robot partners called MOBiMac (Fig.1) in order to realize human-friendly communication and interaction. This robot has two CPUs and many sensors such as CCD camera, microphone, and ultrasonic sensors. Furthermore, the information perceived by a robot is shared with other robot by the wireless communication. Therefore, the robots can easily perform formation behaviors. We have applied steady-state genetic algorithm (SSGA), spiking neural networks (SNN), self-organizing map (SOM), and others for human detection, motion extraction, gesture recognition, and shape recognition based on image processing [14]. Furthermore, the robot can learn the relationship between the numerical information as a result of image processing and the symbolic information as a result of voice recognition [15]. MOBiMac can be also used as a standard personal computer.

We used Voice Elements DTalker 3.0, which was developed by EIG Co., Ltd., Japan, for voice recognition and synthesis in the robot [18]. It was able to perform voice recognition using a sound segment network that made speaker-independent recognition possible. In addition, with the number of words that are recognized dependent on the memory, it achieved a recognition rate of 96.5% (for 200 words).

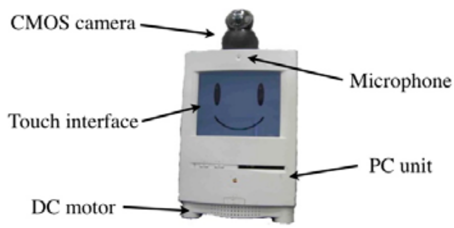


Fig. 1 Human-friendly partner robots; MOBiMac

We have used apple iPad as pocket robot partners, because we can easily use the touch interface and accelerometer in the program development. In this paper, we use iPad as a face of the robot and interaction with students. Figure 2 shows the overview of interfaces used in iPad.

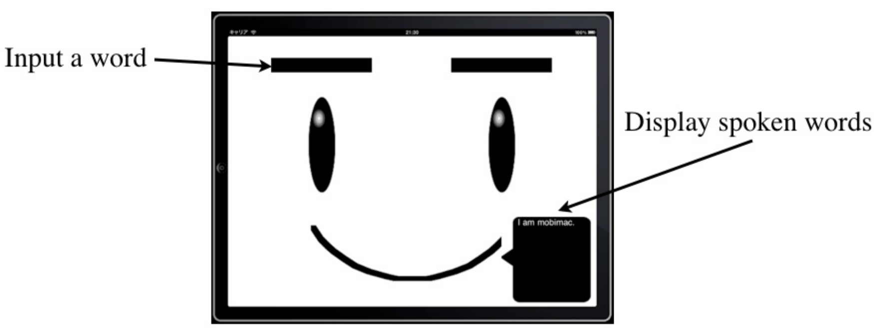


Fig. 2 iPad as a face of MOBiMac

3 Conversation System for Language Education

3.1 Learning Words and Conversation System

Figure 3 shows the total architecture of the perception, decision making, learning, and action. First, the voice recognition and image processing are performed to extract visual and verbal information through the interaction with a person. In addition, the function of word input is used for learning of the robot (Fig.4). In this paper, the robots use perceptual modules for various modes of image processing, such as differential extraction, human detection, object detection, and human hand-motion recognition (Fig.5).

After that, the robot selects the conversation mode from (1) scenario-based conversation, (2) daily conversation, and (3) learning conversation. In the scenario-based conversation mode, the robot makes utterances sequentially according to the order of utterances in a scenario. In the daily conversation, the robot uses a long-term memory based on spiking neural network. The robot selects an utterance according to the long-term memory corresponding to the internal states of spiking neurons. In the learning conversation, the robot updates the relationship between spiking neurons used in long-term memory by associative learning. Finally, the robot makes utterance. In the following sections, we explain the image processing based on steady-state genetic algorithm, and associative learning between perceptual information and verbal words (Fig.6).

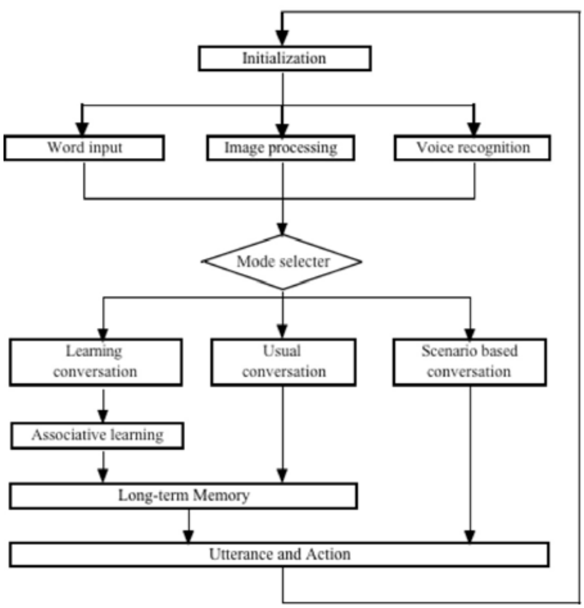


Fig. 3 Flow chart of conversation

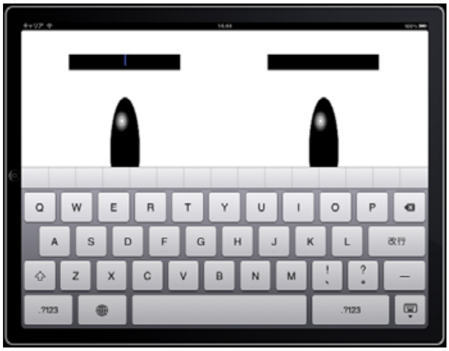


Fig. 4 The screenshot of inputting a word to the robot

We use a simple spike response model to reduce the computational cost for associative learning. First of all, the internal state $h_i(t)$ is calculated as follows:

$$h_i(t) = \tanh(h_i^{syn}(t) + h_i^{ext}(t) + h_i^{ref}(t)) \tag{1}$$

Here, a hyperbolic tangent is used to avoid the bursting of neuronal fires, $h_i^{ext}(t)$ is the input to the ith neuron from the external environment, and $h_i^{syn}(t)$, which includes the output pulses from other neurons, is calculated by

$$h_i^{syn}(t) = \gamma^{syn} \cdot h_i(t-1) + \sum_{j=1, j\neq i}^{N} w_{j,i} \cdot h_j^{EPSP}(t) \tag{2}$$

Furthermore, $h_i^{ref}(t)$ indicates the refractoriness factor of the neuron, $w_{j,i}$ is a weight coefficient from the jth to ith neuron, $h_j^{EPSP}(t)$ is the excitatory postsynaptic potential (EPSP) that is approximately transmitted from the jth neuron at the discrete time t, N is the number of neurons, and γ^{syn} is the temporal discount rate. The presynaptic spike output is transmitted to the connected neuron according to the EPSP, which is calculated as follows:

$$h_i^{EPSP}(t) = \sum_{n=0}^{T} \kappa^n p_i(t-n), \tag{3}$$

where κ is the discount rate ($0<\kappa<1.0$), $p_i(t)$ is the output of the ith neuron at the discrete time t, and T is the time sequence to be considered. If the neuron is fired, R is subtracted from the refractoriness value in the following:

$$h_i^{ref}(t) = \begin{cases} \gamma^{ref} \cdot h_i^{ref}(t-1) - R & \text{if } p_i(t-1) = 1 \\ \gamma^{ref} \cdot h_i^{ref}(t-1) & \text{otherwise} \end{cases} \tag{4}$$

where γ^{ref} is the discount rate. When the internal potential of the ith neuron is larger than the predefined threshold, a pulse is outputted as follows:

$$p_i(t) = \begin{cases} 1 & if \ h_i(t) \geq q_i \\ 0 & otherwise \end{cases}, \tag{5}$$

where q_i is the threshold for firing. The weight parameters are trained based on the temporal Hebbian learning rule as follows:

$$w_{j,i} \leftarrow \tanh(\gamma^{wgt} \cdot w_{j,i} + \xi^{wgt} \cdot h_j^{EPSP}(t-1) \cdot h_i^{EPSP}(t)), \tag{6}$$

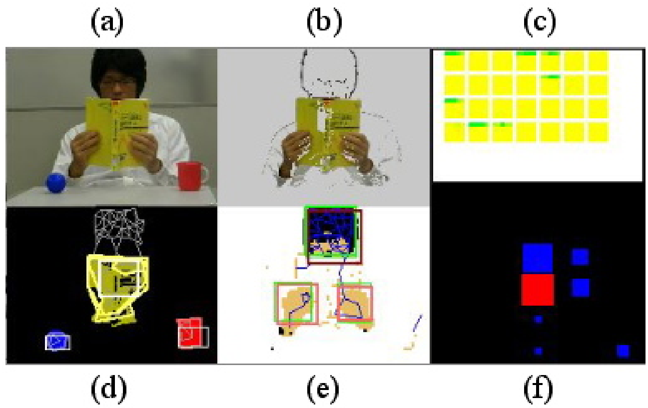


Fig. 5 The robot performs associative learning interacting with the person. (a) the original image, a photograph, (b) differential extraction, (c) the reference vectors of SOM corresponding to gestures, (d) object recognition results by SSGA-O, (e) human detection results by SSGA-H, the green box indicates the candidates for human face position produced by SSGA-H, the red box indicates the face position produced by human tracking, and the pink box indicates the hand position and (f) EPSP of the spiking neurons. which indicates the spatiotemporal pattern captured from the subject's hand motion. The red rectangle is EPSP, and it gradually diminishes, turns blue, and becomes smaller.

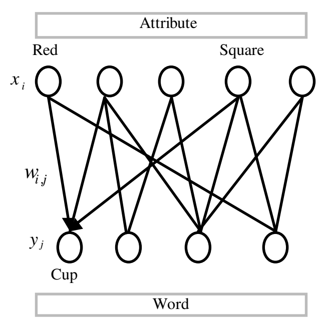


Fig. 6 Learning relationship with SNN

3.2 Self-efficacy of the Robot

When the robot learns with humans, it is desirable for them to learn similarly. Then to evaluate his own learning state, the robot uses self-efficacy proposed by A.Bandura [16].

Self-efficacy is represented by level, strength, and generality (Fig.7).

$$S = S_L + S_S + S_G \quad , \tag{7}$$

where S_L is a level of the action to put the difficulty etc. together on the achievement of the action. It shows the difficulty of speaking English. The difficulty is different because of the length of the talk e.g. in case of only one word, or sentences. The easier the talk, the higher the level.

S_S refers the strength of confidence that executes how much possibility is in each action. It is influenced by the expectation that gets replies and praises. Concretely it is determined by the number of getting replies and neglects when the robot speaks to humans.

$$S_S = \frac{\alpha n_R - \beta n_N}{n_I} \tag{8}$$

n_R is the time getting replies, and n_N is the time of neglects, n_I is the time of interaction. α and β are parameters between 0 and 1. There are 3 rules. If a person was talked in Japanese and answered, n_R increases. If a person was talked in English and answered in English, n_R increases. If answered in Japanese, n_N increases. S_G means the generality of contents adapting to similar circumstances. In comparison with Japanese conversation, the robot thinks how well it can speak English. Concretely, the number of English words is compared to the number of Japanese words.

$$S_G = \frac{n_E}{n_J} \tag{9}$$

Here n_E means the number of English words, and n_J is the number of Japanese words.

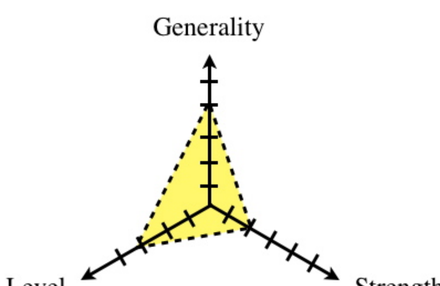


Fig. 7 The dimension of self-efficacy

Basically self-efficacy is enhanced by teaching words. If humans do not reply, it becomes weak. Therefore the conversation becomes fluid.

By enhancing self-efficacy, the robot tries to speak English actively if the robot thinks it can get replies. We think self-efficacy is high, we are willing to communicate [17]. The robot estimates the English skills of humans which will be improved in this way. Self-efficacy is used as a criterion for judgment to speak English or Japanese. Figure 8 shows the concept of self-efficacy in conversation. Outcome expectation is that the robot can get a reply and efficacy expectation decides that the robot speaks English or Japanese.

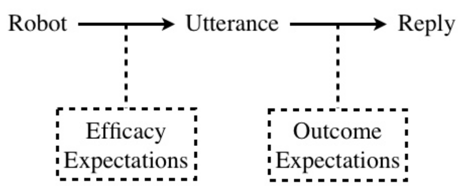


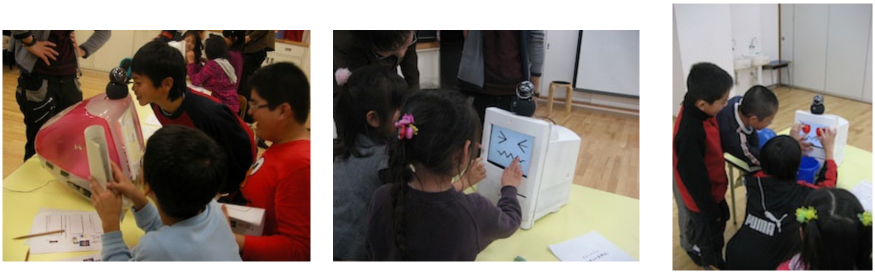
Fig. 8 Representation of self-efficacy in conversation

4 Experimental Results

This section shows experimental results of the proposed method for the language education. We did experiments in children's house. The subjects are some students in elementary school. There are some objects using playing house around the robot. The interaction starts from the conversation in daily life. The subjects merely talk to the robot and show objects then the robot responds to that. When the subjects speak in Japanese, the robot speaks Japanese as well. In learning conversation mode, the subjects teach objects to the robot, the robot learns language and speaks English about the objects. In developmental psychology, there are two kinds of child (the one who tries to remember the name of the thing and the one who tries to memorize the word concerning person's appearance) when the word is memorized. Here we developed the system making conversation by trying to learn the object name.

Next we compared the difference of interaction with the robot (Fig.9). There are 3 patterns, (a) use no touch display and objects, (b) use touch display, (c) use both touch display and objects. The conversation contents are a lot of varieties in pattern (c). Therefore we concluded the pattern (c) is the most appropriate style of robot-assisted language learning. We did not do questionnaire because the children could not listen English words the robot spoke. This was attributable to children's capacity.

Next we show the value of self-efficacy of the robot (Fig.10). We set n_J=30, α=1.0, β=0, threshold for self-efficacy θ=0.5. It was changed by the contents of conversation of a person.



(a) No touch display and objects (b) Using touch display (c) Using touch display and objects

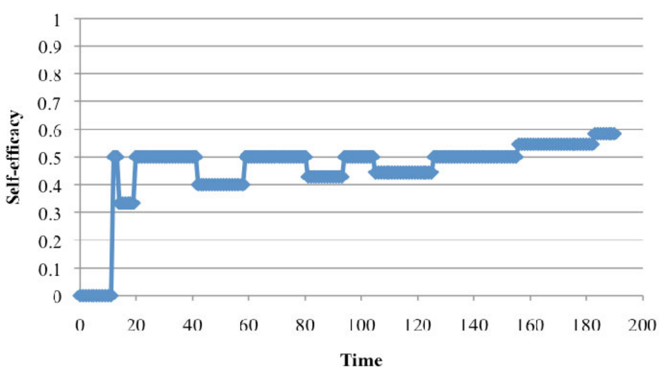
Fig. 9 Learning with the robot

The condition of case 1 are n_I=12, n_R=7 in the final state. It is difficult to talk in English if we did not teach words at first. In the talk of the robot, it is also difficult to talk in Japanese. Therefore it is important that the interaction to teach words to the robot.

In the case 2, we talked to the robot with teaching words up to n_E=10. The value of self-efficacy was monotonically increasing and became high. In this case, the robot always spoke English because the value was high. If we want to make a conversation fluid, we would need to set the value of ß high. And we also need to think the way of communication after self-efficacy is high.

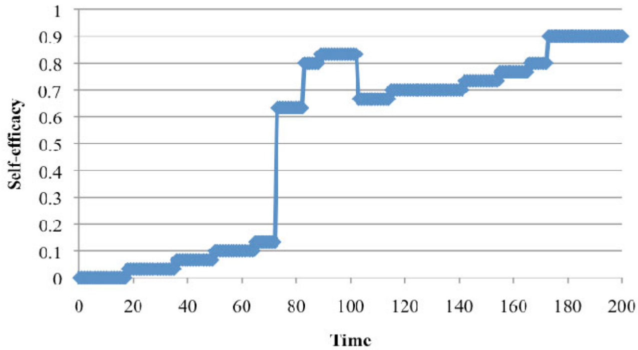
In the case 3, we had talked a little bit longer term than other cases. When the value was low, the robot had spoken only Japanese. Then a person taught English words to the robot and did conversation with the robot, the value was rising. After that the robot had spoken English.

Therefore the robot can learn English as the same pace with the person. If the value were too high, the value was down when the person did not answer the robot's question. As a result of conversation, self-efficacy is effective for robot-assisted language learning.

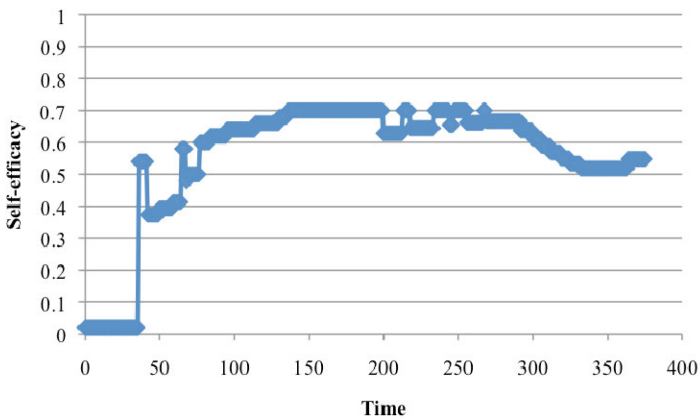


(a) Case 1

Fig. 10 The value of self-efficacy



(b) Case 2



(c) Case 3

Fig. 10 (*continued*)

5 Summary

In this paper, we discussed the applicability of robots in language learning. First, we explained the robot-assisted language learning. Next, we discussed how to interact and communicate with students in the language education. We proposed learning conversation system of physical robot partners. The essence of the proposed method is how humans and robots will improve each other.

As future work, we will inspire the students to learn English. We need to clarify how long the robot can interact with humans. To do long-term communication, we will take gaming element to our system, and enable it to do long-term communication. For example, the value of self-efficacy decides victory or defeat.

Also we will examine whether children can enhance conversation and vocabulary capacity. This time we tried to make elementary school students use this system though, more than junior high school students would be able to use it effectively.

References

1. Kanda, T., Hirano, T., Eaton, D., Ishiguro, H.: Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. Human-Computer Interaction 19, 61–84 (2004)
2. Han, J., Jo, M., Jones, V., Jo, J.H.: Comparative Study on the Educational Use of Home Robots for Children. Journal of Information Processing Systems 4(4), 159–168 (2008)
3. Kwon, O.-H., Koo, S.-Y., Kim, Y.-G., Kwon, D.-S.: Telepresence robot system for English tutoring. IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO), 152–155 (2010)
4. Lee, S., Noh, H., Lee, J., Lee, K., Lee, G.G.: Cognitive effects of robot-assisted language learning on oral skills. In: Proceedings of the Interspeech 2010 Workshop on Second Language Studies: Acquisition, Learning, Education and Technology (2010)
5. Chang, C.-W., Lee, J.-H., Chao, P.-Y., Wang, C.-Y., Chen, G.-D.: Exploring the Possibility of Using Humanoid Robots as Instructional Tools for Teaching a Second Language in Primary School. Educational Technology & Society 13(2), 13–24 (2010)
6. Kanda, T., Sato, R., Saiwaki, N., Ishiguro, H.: A two-month Field Trial in an Elementary School for Long-term Human-robot Interaction. IEEE Transactions on Robotics (Special Issue on Human-Robot Interaction) 23(5), 962–971 (2007)
7. Nakano, M., et al.: Grounding New Words on the Physical World in Multi-Domain Human-Robot Dialogues. In: Dialog with Robots: Papers from the AAAI Fall Symposium (FS-10-05), pp. 74–79 (2010)
8. Uchida, Y., Araki, K.: Evaluation of a System for Noun Concepts Acquisition from Utterances about Images (SINCA) Using Daily Conversation Data. In: Proceeding of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT) 2009 Conference, pp. 65–68 (2009)
9. Kubota, N., Wagatsuma, Y., Ozawa, S.: Intelligent Technologies for Edutainment using Multiple Robots. In: Proc. (CD-ROM) of the 5th International Symposium on Autonomous Minirobots for Research and Educatinment (AMiRE 2009), pp. 195–203 (2009)
10. Fiene, J.: The Robockey Cup. IEEE USA Robotics & Automation Magazine 17(3), 78–82 (2010)
11. Billard, A.: Robota: Clever Toy and Educational Tool. Robotics & Autonomous Systems 42, 259–269 (2003)
12. Yorita, A., Hashimoto, T., Kobayashi, H., Kubota, N.: Remote Education based on Robot Edutainment. In: Proc (CD-ROM) of the 5th International Symposium on Autonomous Minirobots for Research and Educatinment (AMiRE 2009), pp. 204–213 (2009)
13. Yorita, A., Kubota, N.: Essential Technology for Team Teaching Using Robot Partners. In: Liu, H., Ding, H., Xiong, Z., Zhu, X. (eds.) ICIRA 2010, Part II. LNCS, vol. 6425, pp. 517–528. Springer, Heidelberg (2010)

14. Kubota, N., Nishida, K.: Cooperative Perceptual Systems for Partner Robots Based on Sensor Network. International Journal of Computer Science and Network Security (IJCSNS) 6(11), 19–28 (2006)
15. Kubota, N., Yorita, A.: Structured Learning for Partner Robots based on Natural Communication. In: Proc. (CD-ROM) of 2008 IEEE Conference on Soft Computing in Industrial Applications (SMCia), pp. 303–308 (2008)
16. Bandura, A.: Self-efficacy: Toward a Unifying Theory of Behavioral Change. Psychological Review 84(2), 191–215 (1977)
17. Yashima, T.: Willingness to communicate in a second language: The Japanese EFL context. The Modern Language Journal 86(1), 54–66 (2002)
18. DTalker for Mac OSX Ver3.0, http://www.createsystem.co.jp/dtalkerMacOSX.html
19. Hofstadter, Douglas: Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought. Basic Books, New York (1995)
20. Wada, K., Shibata, T., Saito, T., Tanie, K.: Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged. In: Proceedings of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, pp. 2847–2852 (2003)
21. Kanoh, M., Kato, S., Itoh, H.: Facial Expressions Using Emotional Space in Sensitivity Communication Robot "ifbot". In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1586–1591 (2004)
22. Fujita, M.: AIBO: towards the era of digital creatures. International Journal of Robotics Research 20(10), 781–794 (2001)
23. Yaguchi, A., Kubota, N.: The Style of Information Service by Robot Partners. In: Liu, H., Ding, H., Xiong, Z., Zhu, X. (eds.) ICIRA 2010, Part II. LNCS, vol. 6425, pp. 529–540. Springer, Heidelberg (2010)
24. Gonzalez-Gomez, J., Valero-Gomez, A., Prieto-Moreno, A., Abderrahim, M.: A New Open Source 3D-printable Mobile Robotic Platform for Education. In: Proceedings of the 6th International Symposium on Autonomous Minirobots for Research and Edutainment, p. S22 (2011)
25. Riedo, F., Rétornaz, P., Bergeron, L., Nyffeler, N., Mondada, F.: A two years informal learning experience using the Thymio robot. In: Proceedings of the 6th International Symposium on Autonomous Minirobots for Research and Edutainment, p. S11 (2011)
26. Salvini, P., Macrì, G., Cecchi, F., Orofino, S., Coppedè, S., Sacchini, S., Guiggi, P., Spadoni, E., Dario, P.: Teaching with minirobots: The Local Educational Laboratory on Robotics. In: Proceedings of the 6th International Symposium on Autonomous Minirobots for Research and Edutainment, p. S12 (2011)
27. http://en.wikipedia.org/wiki/Doraemon