Interactive e-Hon as Parent-Child Communication Tool

Kaoru Sumi[1] and Mizue Nagata[2]

[1] Hitotsubashi University, Information and Communication Technology Center,
2-1 Naka, Kunitachi, Tokyo 186-8601 Japan
[2] Jumonji University, Sugasawa, Niiza-shi, Saitama 352-8510 Japan
kaoru.sumi@acm.org

Abstract. In this paper, we described a media for helping children understand content, called Interactive e-Hon. It works by transforming text into an easily understandable storybook style with animation and dialogue. In this system, easy-to-understand content is created by a semantic tag generator through natural language processing, an animation generator using an animation archive and animation tables, a dialogue generator using semantic tag information, and a story generator. Through our experiment, we have shown that this method of transmitting visual images with verbal information is effective for promoting understanding.

Keywords: Understanding, animation.

1 Introduction

When providing information to children in words, we must carefully choose words or concepts that the children know, or we must include explanations of the words or concepts themselves. Explanations with visual information, however, are more intuitively understandable than verbal explanations. Children's picture books include both verbal and visual information so that children can easily understand the content. If we could dynamically generate a picture book based on words, it would improve children's understanding of content.

We think that image media such as pictures or animation can efficiently support understanding, particularly in the case of children. Simultaneously presenting verbal information with visual information should support communication and understanding. Communication via a picture book broadens children's vocabulary and helps them learn about unknown concepts.

We have developed a system, Interactive e-Hon [1], for helping children understand content. Interactive e-Hon transforms text from electronic content into an easily understandable "storybook world." The Japanese word hon means "book," while ehon means "picture book." By transforming text into animation, Interactive e-Hon can generate a dynamic picture book.

Attempts to transform natural language into animation began in the 1970s with SHRDLU [2], which represents a building-block world and shows animations of adding or removing blocks. In the 1980s and 1990s, more applications [3][4][5] appeared, in which users operate human agents or other animated entities derived from natural

G. Salvendy, M.J. Smith (Eds.): Human Interface, Part II, HCII 2011, LNCS 6772, pp. 199–206, 2011.

language understanding. Recently, there has been research on the natural behavior of life-like agents in interactions with users [6][7][8]. The main theme in this line of inquiry is the question of how to make these agents as human-like as possible in terms of dialogicality, believability, and reliability. WordsEye [9] is a text-to-scene generation system that includes spatial data. In contrast, our system generates animations but not scenes.

In this paper, we describe the effect of Interactive e-Hon using the system as a dynamic picture book based on existing digital text.

2 Interactive e-Hon

Interactive e-Hon helps children understand difficult content through the use of animation. Our idea is that visual data attracts a child's interest, and that the use of actual examples, like metaphors, facilitates understanding because each person learns according to her own unique mental model [10][11], formed based on her background.

Interactive e-Hon is a fully automatic word translation medium that provides expression through the use of 3D animation and dialog explanation to help users understand Web content or any other electronic resources, such as news, novels, and essays. For given content, animation and a dialog explanation spoken by a voice synthesizer are synchronized.

Figure 1 shows the system framework of Interactive e-Hon. The system generates documents with semantic tags (.tag files), morphological and dependency structure information (.morph files), and animation files (.ehon files), based on the .x file format of DirectX. We use Japanese for text.

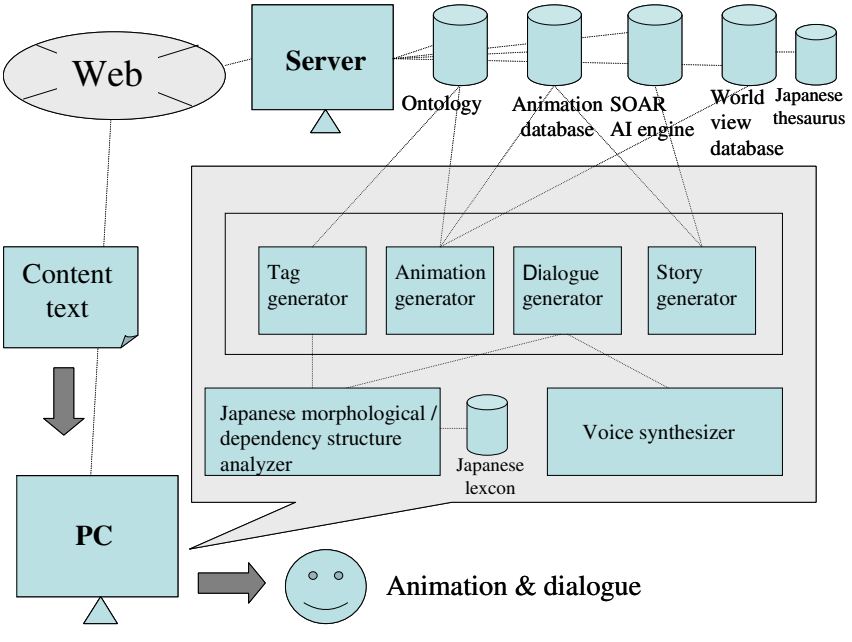

Fig. 1. System framework of Interactive e-Hon

3 Experiment of Parent-Child Communication

The users of Interactive e-Hon are assumed to be a pair consisting of a parent and a child. By observing real users, we could evaluate the effect of using the system through their interactions with it. Therefore, we conducted experiments using real subjects to examine whether Interactive e-Hon was helpful for the users' understanding. We used pairs consisting of a teacher and a child, instead of a parent and a child.

The subjects were two preschool teachers and four children. Teacher A was in her fifties and had been teaching for 25 years. Teacher B was in her forties and had been teaching for 3 years. Children S and H were both boys, approximately 5 years and 6 months old. Child C was a girl, also 5 years and 6 months old. Child M was a girl, 4 years and 3 months old.

In the experiment, the subjects viewed the content of "the origin of the teddy bear" via a dialogue and an animation generated for the text from the Web. The content was presented on the screen with explanation via the dialogue of the parent agent. Each pair of users was asked to sit in front of the display and talk freely while viewing the system. Their interactions were recorded on video. Each teacher was then asked to respond to a questionnaire afterward. Figure 2 shows a screen shot of the resulting content.

3.1 Confirming of Children's Understanding

According to the responses, the teachers and children have previously been unaware of "the origin of the teddy bear's name." We concluded that the concept was not very easy for the children, because both teachers said that the content included some difficult words and concepts. Consider the following examples of a teacher asking about a child's understanding.

<Example 1>
Teacher A: Do you know about the United States of America?
Children S: I don't know. (He shakes his head.)
Teacher A: You don't know?

<Example 2>
Teacher A: Do you know about bear hunting? It means catching a bear.
Children S: (He nods.)
According to their responses, both teachers reported that visualization was an advantage of the system. Because of the visualization, the teachers could easily explain concepts even to the small children, who had poor vocabularies. Regarding the effectiveness of visualization, we observed that both teachers repeatedly pointed to the display during their interactions with the children.

<Example 3>
Teacher A: The Washington Post is a newspaper, you know. (She points to the display.)
Children S: (He nods.)
Teacher A: That story was published in a newspaper in the United States of America. (She points to the display.)

<Example 4>
Teacher B: Then, the company made and sold the stuffed bears. They sold a lot of them. (She points to the display.)

The mother and child agents talk about the content. The original text information can be seen in the text box above the animation.

The following is a dialogue explanation for this example:

Parent Agent: President Roosevelt went bear hunting. Then, he met a small, dying bear.
Child Agent: The President met a small bear who was likely to die.
Parent Agent: But, what do you think happens after that?
Child Agent: I can't guess. Tell me the story.
Parent Agent: The President refused to shoot and kill the bear. And, he helped it instead.
Child Agent: The President assisted the small bear.
Parent Agent: The occurrence was carried by the Washington Post as a heartwarming story,with a caricature by Clifford Berryman.
Child Agent: The episode was carried by the newspaper as a good story

Fig. 2. Sample view from Interactive e-Hon

The next example shows the possibility of showing text on the screen to support the understanding of children who can read.

<Example 5>
Teacher A: Then, what it is called in the text?
Child H: I don't know.
Teacher A: Here it is. (She pointed to the display.)
Child H: Teddy.
Teacher A: Yes. Oh, Yes.

3.2 Attracting Children's Attention

Teacher A reported that the animation attracted the children's attention, indicating another advantage of the system. Child H was very interested in the animation on the display, from the beginning to the middle.

<Example 6>
Child H: What's this? What?
Teacher A: (She nods.)
(The story starts.)
Child H: Ooooh! Ooooh!

Explanation using animated representation can thus facilitate children's understanding and enable easier explanation. It also attracts a child's attention, as shown by the above example.

3.3 Combining Content and Existing Knowledge

Teacher A pointed out the possibility for content to combine children's experience and existing knowledge with their imaginations. The next example illustrates this type of interaction.

<Example 7>
Child S: In my house……
Teacher A: Yes?
Child S: I have a stuffed bear… A big one…. I have it during sleeping time….
Teacher A: Oh. That's nice.
Child S: Such a big one.
Teacher A: You have a bear in your house.
Child S: Yes.

<Example 8>
Teacher B: Do you have a stuffed bear in your house?
Child M: (She nods.) Yes. A blue ribbon one.
Teacher B: It has a blue ribbon? Like this? (She pointed to the display.)
Child M: (She nods.) I always take care of it.
Teacher B: You always take care of it.
Child M: (She nods.)

3.4 Promoting Children's Question

We also observed the children asking the teachers for explanations.

<Example 9>
(Voice: The company exhibited the stuffed bear at an expo.)
Child M: What does "exhibit" mean?
Teacher A: It means "bring and show." He is bringing it, see? (She points to the display.)

This example illustrates the possibility of a child working with an adult and actively acquiring knowledge through their interaction.

3.5 Acceleration of Children's Understanding

The next example demonstrates the acceleration of a child's understanding as a result of interaction with the teacher.

<Example 10>
(Voice: Then, 3000 bears were ordered and there was a teddy bear boom in America. So the name "teddy bear" became established.)
Teacher A: Ah… Americans thought the teddy bear was cute, and it attracted their attention. Then, everyone said, "I want to buy a bear." So the company made a lot of them, like this. (She points to the display while explaining.) 3000 bears were ordered, you know? That's so many, isn't it?
Child S: Yes it is.
Teacher A: You understand? 3000 bears is a lot.
Child S: Yes.
Teacher A: Then, so many bears were ordered. All these people in America said, "I want to buy a teddy bear," and they bought them. (She points to the display.)
Child S: Now, you know what?
Teacher A: Yes?
Child S: Ah, the teddy bear was everybody's favorite?
Teacher A: Yes. It was everybody's favorite.

From the expressions, "3000 bears were ordered" and "there was a teddy bear boom in America," it would be difficult for a child to understand the consequences of the expression, "the teddy bear was everybody's favorite." Additionally, the concepts of "3000 bears", "order," and "boom" are not easy, and the inference of why ordering 3000 bears led to a boom is also difficult. As a result of interaction via the Interactive e-Hon system, however, Child S understood and translated his own word for "favorite."

Consequently, through this experiment, we demonstrated the possibility of actively supporting children's understanding by having them use our system to interact with an

adult. The experiment also showed the advantages of visualization and of explanation by showing a related concept.

4 Discussion

Through our experiment using subject pairs consisting of a teacher and a child, instead of a parent and a child, we have observed some effects of parent-child communication via Interactive e-Hon. In our experiment, we observed many instances of a teacher's pointing to the display to explain concepts. We observed that pointing is effective for explaining, so the interval of time in the content presentation for explanation is important. Interactive e-Hon shows content for each paragraph of the original text, so that there is an interval between paragraphs. When the presented content from a paragraph ends, users have to click the "next" button. Then, the next paragraph is shown as content. We need more verification of the appropriate interval time unit for not interrupting the users' communication.

We think that visual information plays an important role in intuitive understanding of content. Visualization is always more understandable than explanation using many words. In the experiment, we observed a child's paraphrasing unknown words into his own words. There were inferences, such as this: When many bears were ordered, many people said they wanted to buy a teddy bear (teacher). If many people wanted to buy a teddy bear, the teddy bear was everybody's favorite (child). This interaction leads to children's understanding and paraphrasing. We also observed that this visualization via Interactive e-Hon attracted the children's attention and encouraged their questions. By enabling the children to combine content with their existing knowledge, Interactive e-Hon facilitated communication.

Interactive e-Hon is not a passive medium like television but a medium for mediating users' communication. Parent-child interaction via Interactive e-Hon was facilitated like interaction via a picture book through the experiment. This communication style leads to correct, improved understanding through the users' discussion. It also leads to further interaction between users.

We could have changed the original text to make it easier, as there were some difficult words for the children in this experiment. If a parent changed the words or edited the content, it would provide better content. If the interface were enriched in this manner, the system could also be used as a tool for making educational material for parents or teachers and creating content for children.

5 Conclusion

We have discussed the effect of using Interactive e-Hon, a system for facilitating children's understanding of electronic content by transforming it into animation and dialogue. The system corresponds to understanding of the external world using a visual and verbal media. Through our experiment, we have shown that this method of transmitting visual images with verbal information is effective for promoting understanding.

References

1. Sumi, K., Tanaka, K.: Automatic conversion from E-content into virtual storytelling. In: Subsol, G. (ed.) ICVS-VirtStory 2005. LNCS, vol. 3805, pp. 260–269. Springer, Heidelberg (2005)
2. Winograd, T.: Understanding Natural Language. Academic Press, London (1972)
3. Vere, S., Bickmore, T.: A basic agent. Computational Intelligence 6, 41–60 (1990)
4. Bolt, R.A.: "Put-that-there": Voice and gesture at the graphics interface. In: International Conference on Computer Graphics and Interactive Techniques archive, Proceedings of the 7th annual conference on Computer graphics and interactive techniques. ACM Press, New York (1980)
5. Badler, N., Phillips, C., Webber, B.: Simulating Humans: Computer Graphics, Animation and Control. Oxford University Press, Oxford (1993)
6. Cassel, J., Vilhjalmsson, H.H., Bickmore, T.: BEAT: the Behavior Expression Animation Toolkit. In: Prendinger, H., Ishizuka, M. (eds.) Life-Like Characters, pp. 163–187. Springer, Heidelberg (2004)
7. Tanaka, H., et al.: Animated Agents Capable of Understanding Natural Language and Performing Actions. In: Prendinger, H., Ishizuka, M. (eds.) Life-Like Characters, pp. 163–187. Springer, Heidelberg (2004)
8. Marsella, S., Gratch, J., Rickel, J.: Expressive Behaviors for Virtual World. In: Prendinger, H., Ishizuka, M. (eds.) Life-Like Characters pp. 163–187. Springer, Heidelberg (2004)
9. Coyne, B., Sproat, R.: WordsEye: An Automatic Text-to-Scene Conversion System. In: SIGGRAPH 2001, Proceedings of the 28th Annual Conference on Computer Graphics. ACM, Los Angeles (2001)
10. Johnson-Laird, P.N.: Mental Models. Cambridge University Press/Harvard University Press, Cambridge/Mass. (1983)
11. Norman, D.A.: The Psychology of Everyday Things. Basic Books (1988)
12. Sumi, K.: Anime Blog for collecting Animation Data, Virtual Storytelling. In: LNCS. Springer, Heidelberg (2007)
13. Nishida, T., Kinoshita, T., Kitamura, Y., Mase, K.: Agent Technology, Omu Sya (2002) (in Japanese)
14. Liu, H., Singh, P.: Commonsense reasoning in and over natural language. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3215, pp. 293–306. Springer, Heidelberg (2004)