

Videophone Conversation of Two Individuals with Dementia Using an Anime Agent System

Kiyoshi Yasuda^{1,2(✉)}, Masao Fuketa³, Kazuhiro Morita³,
Jyun-ich Aoe³, and Noriaki Kuwahara⁴

¹ Department of Rehabilitation, Chiba Rosai Hospital Tatsumidai-Higashi 2-16,
Ichiharasi 290-0003, Japan

fwkk5911@mb.infoweb.ne.jp

² Augmented Community AID Research Center,
Kyoto Institute of Technology, Kyoto, Japan

³ Faculty of Engineering, University of Tokushima, Tokushima, Japan
{fuketa, kam, aoe}@is.tokushima-u.ac.jp

⁴ Graduate School of Engineering and Science,
Kyoto Institute of Technology, Kyoto, Japan
nkuwahar@kit.ac.jp

Abstract. We reported that videophone conversation was effective in increasing psychological stability of individuals with dementia. We also developed an anime agent system to serve as a conversation partner for individuals with dementia. The computer screen showed an animated face resembling “a 5-year-old grandchild”. The agent was programmed to ask any of 120 pre-set reminiscence questions, automatically detect the ends of replies, and follow with new questions. In the third experiment, the remote multi-party conversation system Skype™ was integrated with the agent system. The agent participated as a presenter of conversation topics for the multi-party interaction. Three pairs conversed under two conditions: conversation of two subjects and a chairperson (human condition) and conversation in which the agent participated as a topic presenter to the above groups (agent condition). The quality of the conversation was scored by three evaluators. The average score of the evaluation was 3.9 (78 %) in the agent condition and 4.9 (100 %) in the human condition. With further improvement of this agent system, it will become practical as a topic presenter for multi-party conversation of individuals with dementia.

Keywords: Agent · Dementia · Conversation · Videophone · Intervention

1 Introduction

The progression of dementia often begins with amnesia and involves behavioral disturbances such as agitation and incontinence [1]. To reduce the stress of individuals with dementia and the burden on their caregivers, various therapeutic approaches have been proposed, such as validation and music therapy [2]. Reminiscence intervention aims to increase self-esteem and psychosocial well-being [3] and to decrease behavioral disturbances [4]. A conversation is a common and enjoyable activity for most individuals [5]. Individuals with dementia, however, tend to be isolated, with few

opportunities to converse [5]. Therefore, providing them with opportunities to converse with people is an important intervention. As the number of individuals with dementia is rapidly increasing, it is becoming difficult to find sufficient numbers of talking partners. In the last decades, many talking dolls and toys for the elderly have become available in the market. Recently, several talking robots have been developed.

One such alternative intervention is conversation with animated agents on a computer screen. Previous studies have investigated the acceptance of such agents by the elderly. They suggested that it was important for the agents to display social signals such as smiling and head nods [6–8]. Sakai et al. [9] developed a computer agent system that could serve as a talking partner for individuals with dementia in clinical settings such as hospitals and clinics. The results revealed that all the individuals with dementia replied and were satisfied with the conversation with the agent. However, the above agent systems could not conduct long conversations such as a 30-min reminiscence or review of life [10]. Short conversations may not be enough to satisfy and stabilize individuals with dementia. To conduct long conversations, we have developed another computer agent system for individuals with dementia [11, 12]. We evaluated the efficacy of this agent system in the three experiments discussed below.

2 An Anime Agent System for Conversation of Individuals with Dementia (Yasuda et al. 2013)

In our agent system [11, 12], the computer screen shows an animated face of a child agent resembling “a 5-year-old grandchild”. When the subject speaks, the agent reacts to them, automatically showing nods, mouth movements, and acknowledgement.

2.1 Methods

We prepared eight categorical sets of approximately 15 (total 120) reminiscence questions, such as those about parents, hometowns, and school life. These were spoken

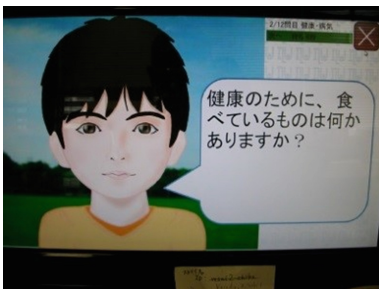


Fig. 1. Face of the anime agent. Note. The agent is asking: What are you eating for your health?



Fig. 2. Multi-party conversation in the agent condition. Note. Agent is on the left. The upper right shows two subjects. The lower right is the chairperson.

aloud through the synthesized voice of the agent. In the preliminary trial, the continual questioning of the agent yielded an atmosphere like “a police interrogation”. To improve this atmosphere, each question was composed of two parts. The first was composed of introductory comments by the agent. The agent introduced his own reminiscences, e.g., “I used to eat watermelon in the summer”. The second part was a question for the subjects, e.g., “what kind of fruits do you like?” Introductory comments and questions were also shown in the written form on the screen for the visual confirmation of questions and compensation for hearing difficulty (Fig. 1).

According to the results of Sakai et al. [9], the waiting time was very important for smooth conversation with this type of agent system. Because of the analysis of Sakai et al. [9], the wait time was set to 3.5 s for this experiment. If speech sounds were not detected during the 3.5 s waiting time, the agent moved to the next question.

Eight subjects with mild Alzheimer’s disease participated in the first experiment. The average age was 78.5 years, and the mean Mini-Mental State Examination (MMSE) [13] score was 22.2. To evaluate the effectiveness of this system, subjects replied to questions by the agent (agent condition) and by a chairperson (human condition). In both conditions, almost the same 15 questions were asked. Each conversation took approximately 20 min.

2.2 Results and Discussion

Recently, several robots and smart phones have speech recognition systems installed to enable them to converse with people. However, the robustness of such systems for speech recognition remains unstable. Furthermore, elderly individuals do not always speak clearly. Because robustness is quite important for the practical use of this system, we employed a sound-detection system.

In this experiment, the syllable number in the subjects’ replies was calculated for the two conditions. The subjects uttered a total of 5494 (74 %) syllables in the agent condition compared with 7406 (100 %) syllables in the human condition. Although the number of syllables was lesser in the agent condition, the system could elicit 74 % syllables from subjects with dementia. This percentage means that this system may become a valuable alternative method to elicit conversation when no human conversation partner exists.

The interview after the experiment was very impressive. A female subject was moved to tears while conversing with the agent because it was “so enjoyable for” her. This suggests that she may have felt the agent was a real boy. Another male subject said “With this system, I can speak freely without any hesitation or anxiety.” Conversations with normal people were stressful to him because of their difficult questions, to which he could not reply. Therefore, we consider that this type of artificial conversation system is required to provide enough chances to speak without hesitation and worry for subjects like him.

3 Multi-party Conversation Between the Agent and Two Subjects (Yasuda et al. 2014)

The number of individuals with dementia is increasing. Because of the shortage of talking partners, even one-(agent)-to-one-(individual with dementia) conversations such as those in experiment 1 are becoming difficult. Therefore, a one-(agent)-to-multi-party conversation will be required. In our second experiment, we performed a multi-party conversation between the agent and two subjects with mild dementia [14].

3.1 Methods

Five pairs (10 subjects in total) conversed with the agent (agent condition) or without it (human condition). Their average age was 75.9 years, and their mean MMSE score was 24. In the human condition, they spoke freely with each other for 10 min. In the agent condition, the agent participated with each pair as a topic presenter. One categorical set, namely *health and disease*, was selected from among eight categorical sets of reminiscence questions. This category includes 12 questions (see the Appendix). This waiting time was fixed to 3.5 s, as in experiment 1. If the conversation in each pair stopped or speech sounds were not detected during the waiting time, the agent uttered the next question. After the question, they talked freely. The 10 conversations (five pairs \times two conditions) were videotaped. We evaluated the influence of the agent participation on the quality of their conversation using original five-scale ratings in five categories (Table 3). In this evaluation, higher scores mean better-quality conversation.

3.2 Results and Discussion

The three evaluators scored the quality of the 10 conversations independently. The total average score was 2.7 (77 %) for the agent conditions and 3.54 (100 %) for the human conditions. Among the five categories, the average score for the category *Flow of conversation* was the worst in the agent condition. The 3.5 s waiting time was too short: the agent often interrupted the speech flow of subjects. On the other hand, this agent system succeeded in raising the interactivity of conversation; in other words, it prevented monopolization by one subject. One participant said “This agent was good at providing topics for the conversation”. Although further improvements are required, this system will be practical as a presenter of topics.

4 Videophone Conversation of Two Subjects with the Participation of an Anime Agent System

Over the last decade, videophone conversations have been proposed for the assistance of individuals with dementia [15]. Kuwahara et al. [16] and Yasuda et al. [17] demonstrated that a combination of videophones and reminiscence interventions was more effective for psychological stability. Some subjects remained psychologically stable for more than 3 h after the conversation session ended [16, 17]. The number of people with

dementia is increasing. Because of the deficit of talking partners, even agent-to-multi party conversations [14] such as those in experiment 2 will become difficult in the future. Therefore, an agent-to-multi-party conversation via a videophone will be required. To determine the effectiveness of anime agent participation in remote conversation, we created two types of multi-party conversation for our third experiment.

One is conversation between the two subjects and a chairperson (human condition). In the other conversation, the agent participated as a topic presenter for the above groups (agent condition). By evaluating the quality of these conversations, we discussed the effectiveness of agent participation in multi-party conversations via a videophone and necessary revisions for further improvement.

4.1 Methods

Subjects. The participants were six subjects with mild and moderate dementia. Their average age was 75.0 years, and their mean MMSE score was 19.4. They were classified into three pairs (Table 1).

Table 1. The Participants

Pairs	Subjects	Sex	Age	MMSE	Etiology
A	Ik	f	78	21	Alz
	Mu	f	78	19	Alz
B	Ok	m	82	24	Alz
	Mu	m	77	20	VD
C	Sh	f	61	14	Alz
	Fu	f	74	20	Alz
	<i>Mean</i>		<i>75.0</i>	<i>19.4</i>	

Note. Alz, Alzheimer's dementia VD, Vascular Dementia MMSE, Mini-Mental State Examination

Procedure. A volunteer engineer set up hardware such as a PC and a webcam in the home of each subject. The engineer connected the PCs to the Internet via fiber optic cables. Skype™ was set up so that it automatically launched when a chairperson (a speech-language pathologist, one of authors of this study) clicked the subject's Skype™ name. The caregivers were asked not to turn off or unplug the PC. Recently, Skype™ has acquired the capability to conduct a multi-party videophone conversation with up to 10 persons simultaneously without charge (Fig. 2).

Each pair conversed in the following two conditions: conversation between two subjects and a chairperson with the participation of an agent on the PC monitor (agent condition) and conversation without the agent (human condition). The order of the two conditions was randomized. In both conditions, one categorical set, namely *health and diseases*, was used. The 12 questions used were the same as those in experiment 2 (see the Appendix). The supposed age of the agent was 5 years. The chairperson's age was 62 years. Therefore, the introductory comments were slightly modified to suit the age of the speaker (see the Appendix).

In experiment 2, the main waiting times were set to 3.5 s, which interrupted the flow of the conversations. Because two subjects and a chairperson talked in this experiment, the main waiting time was extended to 10 s. These six conversations (three pairs \times two conditions) were videotaped for the following evaluation.

In both conditions, the chairperson encouraged the subjects when another subject did not reply or when their replies were short, with the following prompts: “Mr./Ms. __, let’s talk now,” or “Mr./Ms. __, let’s talk about it in detail”.

We evaluated the effectiveness of these conversations using a psychological five-point scale rating (Table 3) with three evaluators. They independently evaluated the quality of these six conversation videos in random order. Three evaluators, with an average age of 70 years, were talking volunteers for elderly patients at a hospital.

4.2 Results

The average conversation time was 13.72 min for the agent condition and 12.28 min for the human condition. Table 2 shows the average quality scores for the six conversations by the three evaluators. A higher score means better conversation quality. The total average score was 3.9 (72 %) for the agent condition and 4.9 (100 %) for the human condition.

Table 2. Average scores by three evaluators

Pair	Human	Agent
Pair A		
Naturalness of conversation	4.3	3.3
Liveliness of conversation	4.3	3.7
Flow of conversation	4.3	3.3
Interactivity of conversation	4.7	3.6
<i>Mean</i>	<i>5.0</i>	<i>4.6</i>
Pair B		
Naturalness of conversation	4	3.7
Liveliness of conversation	4	3.3
Flow of conversation	4	3.3
Interactivity of conversation	4	4
<i>Mean</i>	<i>4</i>	<i>3.6</i>
Pair C		
Naturalness of conversation	5	3.3
Liveliness of conversation	5	3.7
Flow of conversation	4.7	3.7
Interactivity of conversation	4.3	3.7
<i>Mean</i>	<i>4.8</i>	<i>3.6</i>
<i>Total average</i>	<i>4.9</i>	<i>3.9</i>

Note. Another: encourage another subject to talk; more: encourage subject to explain more.

The number of times the chairperson encouraged another subject to talk or to explain more was calculated. The total number of encouragements was 47 for the agent condition and 38 for the human condition. The number of encouragements by the chairperson, particularly encouragements for another subject to talk, was greater in the agent condition than in the human condition (Table 3).

Table 3. Number of encouragements for subjects

		Human		Agent	
		another	more	another	more
Pair	A	5	8	16	6
	B	10	3	6	6
	C	4	8	12	1
<i>Subtotal</i>		<i>19</i>	<i>19</i>	<i>34</i>	<i>13</i>
<i>Total</i>		<i>38</i>		<i>47</i>	

Observations by evaluators. In the agent conditions, some subjects unconsciously made noises such as table tapping. The computer of one subject emitted an electronic whine. These events prevented the agent system from uttering the next question. Direct conversation between subjects did not occur, except in the human condition of pair C. In pair C, the two subjects asked questions to each other and cracked jokes. In most cases, the turns to reply was naturally fixed. Some subjects did not start replying until the chairperson urged them, in both conditions.

4.3 Discussion

The videophone is an appropriate communication tool for individuals with dementia to understand what is said in a conversation with gestures and body language. Some individuals were even stable for more than 3 h after the end of a videophone conversation session [16, 17]. Insufficient communication was considered to be a cause of stress, worry, or instability, which may have been reduced by the videophone conversation. These studies suggest that conversation itself has the potential to prevent individuals with dementia from showing anticipated behavioral disturbances such as “evening syndrome” [16, 17].

However, the number of individuals with dementia is rapidly increasing; it is very difficult for them to engage in conversation at all times. Frequent and regular videophone conversation is becoming difficult to perform. As a possible resolution of these situations, we incorporated an agent as topic presenter in the videophone conversation. Although a prototype has been proposed [8], this is the first clinical trial of the participation of an agent in a videophone conversation.

We observed a multi-party conversation with an agent (agent condition) and without an agent (human condition). The time required to conduct these conversations was almost the same. However, conversations in this experiment were, strictly speaking, not normal conversations. The topics of the conversations were preset. Flexible reactive comments to the replies of subjects were impossible in the agent conditions (system) and were restrained by the chairperson in the human conditions to balance the conversation style in the two conditions. Although these operational procedures prevent intrinsic comparison on the quality of conversations between two conditions, all average scores of the quality of conversations were better in the human conditions. However, the percentage of the scores was 72 % for the agent condition compared with 100 % for the human condition. We consider that this percentage means it is worth applying this system in supporting group conversations via a videophone.

From the number of encouragements, conversations in the agent conditions seemed to need more prompts, particularly encouragements for another subject to talk, than those in the human conditions. In the agent conditions, some subjects may have felt hesitation from speaking at will. In future revisions, encouraging words should be used to prompt more reserved participants to talk, such as “how about another person?” as well as “please explain in detail”.

Direct conversation between subjects occurred naturally in the human condition of pair C. To compensate the technical insufficiency of the agent system, direct talking between participating subjects should be augmented by prompts such as “Let’s talk to each other”. Furthermore, to increase the benefit of the agent system or of ICT interventions, use of various internet resources such as pictures, music, and short movies will be greatly beneficial.

Smooth transfer to the next question was sometimes disturbed by the subject’s unconscious noises and electronic whines. Although this system may work well under quiet circumstances, some methods of coping are required in these cases, such as a forced transfer function to next question. Future revision will incorporate the above prompts and functions in the agent system to increase the usability of this system.

Most individuals with mild or moderate dementia still have the ability to talk to each other. They also say that “I would like to make a social contribution, even though I have dementia”. Talking volunteers are one of the few remaining employments for them. Indeed, they are even more suited to be talking volunteers for other individuals with dementia. They easily forget what has been already said; therefore, they are not annoyed by repetition by other individuals with dementia. However, it is often difficult to recollect topics because of their degraded recall abilities. This agent system can work as a topic-providing system for them. Most families in advanced nations have computers and access to the internet. Younger seniors with dementia who are accustomed with the operation of PCs and smart phone are increasing. The operation of Skype™ will not be difficult for such individuals.

In this field study, the quality of conversation was based on the evaluator’s impressions. In the future, the use of artificial intelligence tools such as facial expression analysis and laughing voices analysis is desirable for the objective proof of efficient and enjoyable conversation. Procedural limitations and the small number of subjects require caution in the interpretation of the results in this experiment. However,

agent participation in multi-party videophone conversations showed the possibility of supporting individuals with dementia.

4.4 Conclusion

An anime agent participated in multi-party videophone conversations as a conversation topic presenter. Although further improvements are required, this agent system may become a promising intervention for assisted conversation of individuals with dementia.

Acknowledgments. The authors wish to thank the cooperation of the six subjects and their caregivers, the three evaluators, Hachiro Kiyoto, and Shingo Aoe.

A Appendix: List of Questions

The underlined are introductory comments. In the human condition, the words before the parentheses were changed with the words within them.

My father (I) has (have) back pain. Do you have any chronic pain?

My older sister (I) eats (eat) fruit (yogurt) every morning for maintaining health.
What are you eating for health?

Last night, I had a dream in which I was flying in the sky. Do you have dreams while you are sleeping?

I am going to bed as early as possible for my health. Are you doing anything to maintain your health?

My father (I) is (am) wearing eyeglasses. How is the condition of your eyes?

I am enjoying vigorous health (I go to swimming once a week). My mother says not to run indoors. How is your health recently?

My grandmamma (I) is (am) taking medicine to reduce high blood pressure. Is your blood pressure high?

Usually, I go to bed at nine o'clock. What time do you go to bed?

My father (I) smokes (smoke) a cigarette after supper. Do you smoke?

I was hospitalized due to bone fracture. Have you been hospitalized?

Recently, I fell and cut my knee. Have you have experienced serious disease or injuries?

If I had a lot of money, I would like to go traveling. What would you want to do if you were 30 years younger?

References

1. Davis, R.N., Massman, P.J., Doody, R.S.: Cognitive intervention in Alzheimer disease: A randomised placebo-controlled study. *Alzheimer Dis. Assoc. Disord.* **15**, 1–9 (2001)
2. De Vreese, L.P., Neri, M., Fioravanti, M., Belloi, L., Zanetti, O.: Memory rehabilitation in Alzheimer's disease: A review of progress. *Int. J. Geriatr. Psychiatry* **16**, 794–809 (2001)
3. Lai, C.K., Chi, I., Kayser-Jones, J.: A randomised controlled trial of a specific reminiscence approach to promote the well-being of nursing home residents with dementia. *Int. Psychogeriatr.* **16**, 33–49 (2004)
4. Finnema, E., Droes, R.M., Ribbe, M., Van, W.: The effects of emotion-oriented approaches in the care for persons suffering from dementia: A review of the literature. *Int. J. Geriatr. Psychiatry* **15**, 141–161 (2000)
5. Yasuda, K., Kuwabara, K., Kuwahara, N., Abe, S., Tetsutani, N.: Effectiveness of personalised reminiscence photo videos for individuals with dementia. *Neuropsychological Rehabil.* **19**, 603–619 (2009)
6. Cassll, J.: Embodied conversational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (eds.) *EC Agents*, pp. 1–27. The MIT Press, Cambridge (2000)
7. Heerink, M., Krose, B., Evers, V., Wielinga, B.: Studying the acceptance of a robotic agent by elderly users. *Int. J. ARM* **7**, 33–43 (2006)
8. Smith, C., Crook, N., Boye, J., Charlton, D., Dobnik, S., Pizzi, D., Cavazza, M., Pulman, S., de la Camara, R.S., Turunen, M.: Interaction strategies for an affective conversational agent. In: Safonova, A. (ed.) *IVA 2010. LNCS*, vol. 6356, pp. 301–314. Springer, Heidelberg (2010)
9. Sakai, Y., Nonaka, Y., Yasuda, K., Nakano, Y.: Listener agent for elderly people with dementia. In: *Proceedings of the Seventh annual ACM/IEEE International Conference on Human-Robot Interaction (HRI2012)*, pp. 199–200 (2012)
10. Butler, R., Butler, N.: Successful aging and the role of the life review. *J. Am. Geriatr. Soc.* **22**, 529–535 (1974)
11. Yasuda, K., Aoe, J., Fuketa, M.: Development of an agent system for conversing with individuals with dementia. In: *The 27th Annual Conference of the Japanese Society for Artificial Intelligence* (2013). <https://kaigi.org/jsai/webprogram/2013/pdf/979.pdf>
12. Fuketa, M., Morita, K., Aoe, J.: Agent-based communication system for elders using a reminiscence therapy. *Int. J. Intell. Syst. Technol. Appl.* **12**, 3–4 (2013). doi:10.1504/IJISTA.2013.056533
13. Folstein, M.F., Folstein, S.E., McHugh, P.R.: “mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975)
14. Yasuda, K., Fuketa, M., Aoe, J.: An anime agent system for reminiscence therapy. *Gerontechnology* **13**, 118–119 (2014)
15. Sävenstedt, S., Brulin, C., Sandman, P.O.: Family members' narrated experiences of communicating via video-phone with patients with dementia staying at a nursing home. *J. Telemedicine Telecare* **9**, 216–220 (2003)
16. Kuwahara, N., Yasuda, K., Tetsutani, N., Morimoto, K.: Remote assistance for people with dementia at home using reminiscence systems and a schedule prompter. *Int. J. Comput. Healthc.* **1**, 126–143 (2010)
17. Yasuda, K., Kuwahara, N., Kuwabara, K., Morimoto, K., Tetsutani, N.: Daily assistance for individuals with dementia via videophone. *Am. J. Alzheimer's Dis. Other Dementias* **28**, 508–516 (2013)