

高齢者の発話を促す会話システム

金星一、中島一樹、中村加奈子*、東祐二**、南部雅幸、藤元登四郎**、田村俊世
国立長寿医療研究センター、八日会グリーンホーム*、八日会藤元早鈴病院**

国立長寿医療研究センター老人支援機器開発部
474-8522 愛知県大府市森岡町源吾 36-3

kim@nils.go.jp

あらまし

本研究では、高齢者と介護者の生活の質を向上させることを目的とする。それを実現する方法として、高齢者の音声を認識し、介護者の録音音声を応答する会話システムを開発した。そして、老人保健施設でデイケアを受けている痴呆高齢者を対象として、システムを導入する場合としない場合の発話頻度を分析した。その結果、本システムの導入により被験者の発話頻度が増加した。

キーワード 会話システム、音声認識、老人、福祉、デイケア

Interactive System for Encouraging Utterance of the Elderly

Sung-Il Kim, Kazuki Nakajima, Kanako Nakamura*,
Masayuki Nambu, Yuji Higashi**, Toshiro Fujimoto**, Toshiyo Tamura

Department of Gerontechnology, National Institute for Longevity Sciences
36-3 Gengo, Morioka, Obu, Aichi 474-8522, Japan

*Greenhome, **Fujimoto Hayasuzu Hospital, Miyakonojo, Miyazaki, Japan

kim@nils.go.jp

Abstract

The goal of the present research is to improve the quality of life of both the elderly and caregivers. The aim of this study is to develop a simple interactive system that recognizes the utterances of the elderly and then replies by the recorded voices of caregivers. The system was evaluated in a nursing home by introducing the system to an older person with dementia. The comparison study was then carried out when the system was given and not, respectively. The evaluation results showed that the frequency of encouraging the utterances of the subject increased by introducing the system.

key words Interactive System, Speech Recognition, Elderly, Welfare, Day Care

1. Introduction

As it has become an aging society both in developed and in developing countries, the overall quality of life(QOL) of the elderly is getting more and more important issue in reality. Particularly, keeping a good health physically and mentally is a very important issue for all elderly and increasingly so with progressing age. Therefore, such related technologies and secure environments have been designed for independent living and social participation of older persons. For one of the methods realizing them, mutual interaction based on a conversation has been studied for increase of mental health-related QOL and rehabilitation.

The goal of the present study is to improve the QOL of the elderly. As the preliminary research, we confine a subject to an aged person with dementia, with a focus on improvement of the QOL of dementia elderly as well as his or her family caregivers in long-term care [1]. Dementia possesses unique features that make caregiving particularly burdensome. The QOL of the family caregivers is, therefore, inevitably affected. Namely, the family caregivers experience mental stress, burden, and depression as outcomes of caregiving, throughout the changing disease process. The caregiving stress often leads to problems in caregivers' mental and physical health because of the behavioral problems of their family member with dementia. Thus, they usually use a social support system such as day care centers, or skilled nursing facilities. In the case, health professionals can play a vital role by providing advice, support, and help to assist the family caregivers in the difficult task of tending their dementia patient during the long and stressful illness.

On the basis of these social backgrounds, we developed an interactive system of recognizing and understanding demands of subject depending on the speech recognition techniques, which is then applied to him at nursing facility. We aim that the proposed system will be a conversational partner to patient whenever he or she wants. Therefore, an emotional stability may be recovered by mutual communication, thus having effects of rehabilitation. Furthermore, the system might also be helpful for the nursing works of family caregivers at home or professional caregivers at nursing facilities. In order to realize it, our approach is focused on a natural interaction with the system by using only spontaneous speech, without using any other interfaces, for instance, keyboard or mouse, etc.. The wireless microphone is also used between the dementia patient and the system. In addition, the system should have functions of speech recognition, time recognition, and agreeable response for a natural conversation.

2. Subject Description

2.1 Observation of Behavioral Patterns of Subject

Subject was 72 years male patient with vascular dementia as well as visual handicapped. He has received day care services at nursing home, which is performed from 8:30 a.m. to 3:30 p.m. on weekdays. For the study, the formal ethics committee approval of Fujimoto Hayasuzu Hospital was obtained in advance and the written informed consent was also obtained from his spouse.

Nursing procedures at nursing facility are particularly disruptive and can lead to sudden agitation or confusion in a dementia patient like the subject who was calm at home. Although he has relatively stable behavioral patterns before lunch break at the nursing facility, he shows an unstable status after lunch. He increasingly continues to complain to allow him to go back home in time. It is because he is very sensitive to the environment of the nursing home, so that changes in resident environment or life schedule can cause extreme disappointment, frustration, or fear. Therefore, those environmental stresses are much harder for caregivers to care for him at nursing home.

As preliminary investigation to cope with the behavioral problems of the subject at nursing home, behaviors and utterances of the subject during the day care service were first taken by video camera. Fig. 1 shows the typical observation results at the time of most unstable status of the subject from 2:00 p.m. to 3:30 p.m. before returning to home. It illustrates average frequencies of behavioral patterns, such as complaint, knocking on a door, and wandering, which were observed during each 90 minutes of the day care service for 3 days.

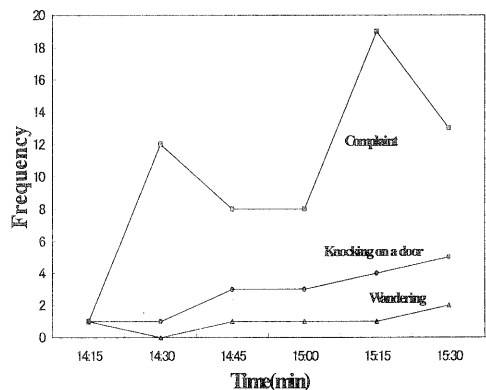


Fig.1. Average frequencies of behavioral patterns of subject during each 90 minutes of the day care service for 3 days, that is, frequencies of complaint, knocking on a door, and wandering

It was found that the subject had the most frequent behavioral pattern in complaint among three kinds of the patterns, because agitation or uncomfortable emotional status persists and grows worse over time. Since dementia has a disorder of memory, subject affected by it occasionally develops agitation or mental instability, making it much harder to care for him. Accordingly, the interactive system should be designed to be a conversation partner as a good listener, sometimes leading him to talk, so that it might help him to maintain a stable emotional status.

2.2 Criteria for Selection of Dialog

The elderly with severe dementia often finds it hard to remember the meaning of words that we are using, or to think of words he wants to say. In this study, therefore, we selected a dialog concerning how to talk with dementia patients according to the following criteria [2] on the basis of the experiences of the professional caregivers or the occupational therapists.

First of all, caregivers or families should try to use a dialog that does not make agitation or frustration worse for the elderly with dementia as their patient or family member. Secondly, they try to call by name or intimate nickname. Thirdly, they try to speak slowly and distinctly, and try to use familiar words and short sentences mixed with regional dialects. Fourthly, they try to keep things positive. For example, offer positive choices like "Let's wait until your family get to here," or "After lunch, let's go back home". Finally, they try to ask simple questions that can be answered with yes or no or one-word answers, when it is difficult to know what he wants.

Under these criteria, we collected voice data from two female caregivers at nursing home, whose voices were particularly more responsive for subject than the voices of others. We then constructed several kinds of dialog database, considering the change of necessary information for subject, such as the time information returning to home. The interactive system, therefore, needs to have functions of announcing current time and related information on time schedule in order to maintain stable emotional status for him during day care service of nursing home.

3. Outline of Conversation System

In order to realize communication between the dementia patient and the interactive system, the frequently used words of the subject at nursing home were first surveyed. The utterances were then registered in the recognition candidate lists. The techniques of command word recognition were used

for recognition module. It is due to the limited speech sources of the subject with severe dementia, on the basis of observation of his utterance patterns. For responses of the system, it is organized to give relevant vocal responses by speech synthesis or recorded replies when the subject complains or demands. Therefore, the system interacts with him through integration of human speech interaction and graphical user interface. It also provides functions of making agreeable responses to his utterances and recognizing current time as well. Integration of all those modes enables the system to interact with the dementia subject smoothly and naturally.

Fig. 2 shows the main frame of the interactive system with recognition candidate lists and graphical face image of virtual caregiver. In the right side of the main frame, the frequently used words of the subject at nursing home were registered for recognition.



Fig.2. Main frame of the interactive system with speech recognition candidate lists and graphical agent

Basically, the proposed system is based on the speech recognition for understanding of utterances of subject. However, the high performance of recognition of a spontaneous speech is not always guaranteed owing to an ambient noisy environment and slurred speech with dialects of aged subject. In this study, therefore, additional functions supplement those shortcomings of speech recognition to maintain natural interaction. The proposed system has two available functions, such as time recognition and agreeable response. For instance, when subject asks current time, the system informs him of necessary information on time schedule of day care service as well as the present time, sometimes making timely agreeable responses. Subject has a tendency of asking about current time more frequently for about two hours before returning home, because of an agitation or emotional instability(see Fig.1). When a necessary time information of home-coming or agreeable response is given to him, he relatively maintains calm emotional status. Even when the

system fails in recognizing utterances of subject, or when the incoming speech of subject have not registered in speech recognition candidates, the system is designed to keep those relevant performances to an incoming speech of the subject on time schedule. Consequently, those functions are particularly helpful for promoting a natural interaction between the system and the subject. Therefore, the function of both time recognition and agreeable response compensates for a weakness of speech recognition. As a result, subject might feel as if the system hears and understands his words attentively, just like the familiar caregivers at nursing home or members of his family. Recursively, the system is ready for detecting an input speech.

Fig. 3 shows basic concept of virtual conversation between severe dementia patient like the subject and the system.

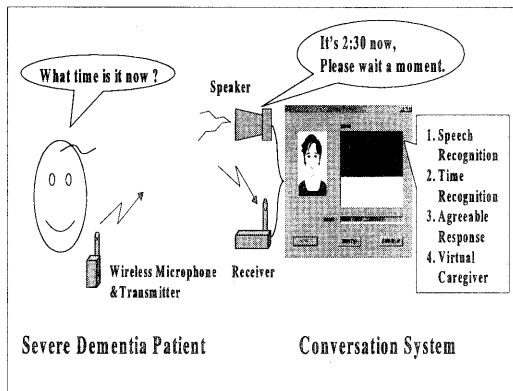


Fig.3. Basic concept of virtual conversation between the severe dementia patient and the system

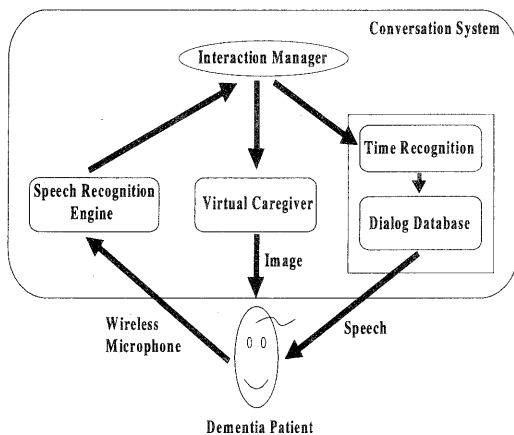


Fig.4. Schematic diagram of interaction between the system and the subject with dementia

As mentioned above, the system has four major functions, such as speech recognition, time recognition, agreeable response, and virtual caregiver. For means of interaction, a wireless microphone (WT-1110, TOA, Japan) was used not to disturb a free behavior of subject.

Fig. 4 illustrates a schematic diagram to form conversation between the system and the subject. The interaction manager controls three kinds of modules such as speech recognition engine [3,4], graphical agent, and time schedule recognition module. In case the subject talks or demands to the system, his voices are first recognized by speech recognition engine module. In the module, the feature vectors of incoming speech signals are compared with standard models registered in the recognition candidate lists. The recognized result is then given to the interaction manager for necessary responses. The manager memorizes current time of acquiring recognition result to offer the most relevant information that the subject wish to obtain at the time interval, by automatically searching for dialog pattern database. It thus chooses the most likely and suitable response and then makes a response with recorded voices, simultaneously making a face image of virtual caregiver move just like speaking to the subject.

4. Experiments

As shown in behavioral patterns illustrated in Fig. 1, he frequently complains to have necessary information returning to home after lunch at nursing facility. In experiments, therefore, the system is required to introduce especially when he continues to complain because he has the most unstable mental or emotional status. For mutual communication, in addition, recorded speech of caregivers was used instead of synthesized speech that was not responsive to the subject. The system set in a relatively noisy environment of main hall of nursing facility where nursing home residents have rests or meals. For evaluation of the system, comparison study was also carried out during each 90 minutes in the afternoon of day care service for 5 days with and without the system, respectively. The occupational therapists then observed his reactions to handling of the system while recording his behaviors using a video camera at nursing home.

5. Results

We evaluated the frequencies of both intervention of caregivers and corresponding reaction of the subject without system. On the other hand, we also evaluated the same performance with system to see how effectively the system carries out instead of

caregivers. Fig. 5 shows normalized frequencies of handling of both the caregiver and the system to complaints of the subject when the system was introduced and not, for 5 days. When the system was not given, it was found that the caregivers had 21% responses (n=82) to complaints (n=389) of the subject. When the system was given, on the other hand, the system had 79% (n=468) to complaints (n=593) of the subject.

These comparison results include the following meaning. Namely, in case the system was not introduced, although caregivers handled complaints of the subject by 21%, they did not still handle relevantly by the other 79%. In case the system was introduced, on the other hand, handling of the system comes to 79%. The system had the other 21% of no responses, owing to the main cause of sensitivity of detecting an incoming speech. In comparison of the both graphs, consequently, we could see that the system was more responsive by 58% than the caregivers of busy nursing schedules.

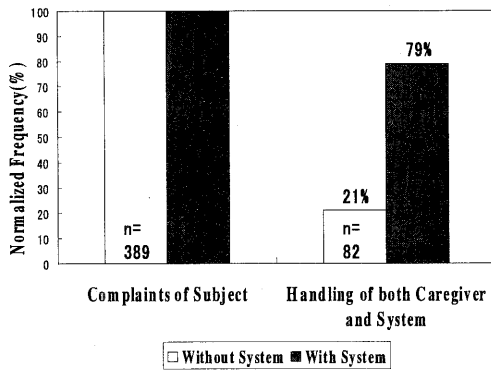


Fig.5. Normalized frequencies of handling of both the caregiver and the system to complaints of the subject

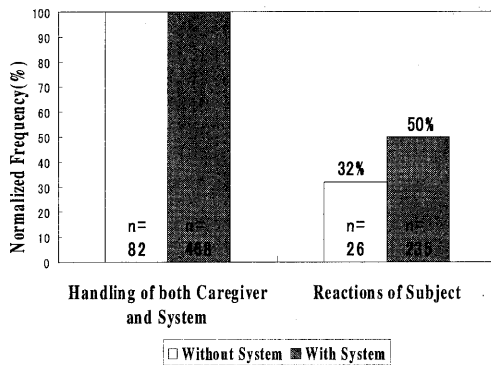


Fig.6. Normalized frequencies of reactions of the subject to each handling of both the caregiver and the

system

Fig. 6 shows normalized frequencies of reaction of subject to each handling of both the caregiver and the system. When the system was not given, the subject had 32% reactions (n=26) to handling frequencies (n=82) of the caregivers. When the system was given, on the other hand, the subject had 50% (n=235) to handling frequencies (n=468) of the system. It means that the subject has relatively more active conversation with the system than with the caregivers. Thus, It was shown that the subject was more interactive with the system by 18% than with the caregivers.

Although the above results showed that the subject had more frequent interaction with the system, it was not shown whether a natural or smooth conversation was built or not. To evaluate how effectively the system maintains an interaction with the subject, therefore, it is important to study what emotional status he preserves during conversation. Accordingly, this study is useful for evaluating whether the system provides the subject some emotional burden or peace. We first analyzed his reactions to the handling of both the caregiver and the system, respectively. We then divided his reactions with several kinds of components. As a first component, subject showed a reaction of asking back to get or confirm agreement, particularly when he heard necessary information from the caregiver or the system as information provider. As a second component, subject showed a reaction of asking back twice when the replies of caregiver or system to his question were quite different from his expectation. Finally, subject showed an affirmative reaction with simple answers such as 'yes' or 'I see'.

Fig. 7 shows normalized frequencies of reactions of the subject, such as asking back once and twice, and affirmation, as a further analysis of Fig.6.

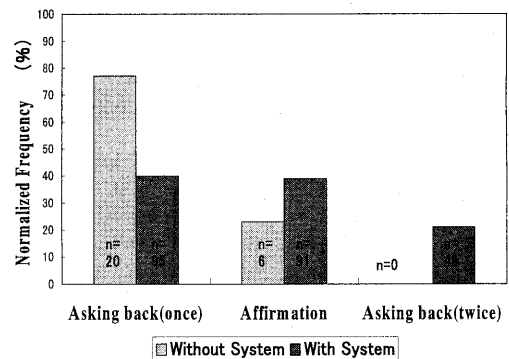


Fig.7. Normalized frequencies of reactions of the subject, such as asking back once and twice, and affirmation in each case

The notable differences were not shown both in the component of asking back once and in the component of affirmation. However, in the component of asking back twice to handling of the caregiver, the subject had no any reactions of asking back twice. To handling of the system, on the other hand, the subject showed reactions of asking back more than two times. The reactions were shown particularly when conversation was not built smoothly and naturally, resulting in the uneasiness or emotional instability of the subject. In the case, therefore, caregivers are obliged to intervene the conversation between the subject and the system.

6. Discussion

This study aims to improve QOL of the elderly. In order to achieve the goal, the preliminary research has been initiated to apply the proposed conversation system to dementia elderly, with a focus on not only increasing QOL of them but also decreasing the nursing burden of family caregivers.

As results of the experimental evaluation, speech recognition accuracies were not satisfactory owing to the slurred speech with regional dialects and characteristic accents of the subject. Nevertheless, addressing or making agreeable responses enabled him to smoothly communicate with the system through a spontaneous speech. Accordingly, it could be seen that the subject felt as if the system heard and understood his complaints curiously and sometimes seriously. As shown in comparison results in Fig.5, the system was able to provide more responsive handling to complaints of subject than caregivers. In addition, the subject had more active conversation with the system than with the caregivers as shown in Fig. 6. Therefore, the system might be expected to lighten the burden of nursing works of caregivers particularly in the busiest time interval, by letting the system converse with their patients who wish to have some conversation or consultation. Moreover, emotional stability or such rehabilitative effect was also expected through a mutual communication whenever their patients want.

However, we still have essential issues to solve. As illustrated in Fig.7, when the system was not introduced, there were no any reactions of asking back more than two times in interaction with caregivers. It means that the subject had natural conversation with the caregiver. Therefore, we could see that the natural and timely conversation never gives some kinds of emotional stresses or such burdens to the subject. When the system was introduced, on the other hand, it was found that there were occasionally reactions of asking back more than two times. It was mainly due to the failure of

recognizing speech of the subject, resulting in unnatural conversation between them. It eventually developed an agitation or uneasiness, so that he had some kinds of stresses or lack of emotional stability during interaction with the system.

One of the evaluation results obtained in this study showed that the system should be designed to keep a natural conversation lest stressful burden was given to the subject in mutual interaction. Therefore, the future works will be considered to deal with an improvement of speech understanding based on the speaker adaptation techniques for providing the most proper answer to demands of individual dementia patient. Moreover, there is one thing to remember from the evaluation results. It is about roles of nonverbal communication such as emotional sympathy, body touch, etc., when talking with each other in daily human conversation. Actually, caregivers talk with their patients while grasping their hands. Those actions also build a new type of mutual communication through a nonverbal interaction, in addition to a human speech. It can be implemented, for example, by using a graphic agent with human emotional expressions in conversation.

7. Conclusion

In this paper, we developed an interactive system and applied it to an elderly patient with severe dementia. As a result of survey, we could see that our proposed system encouraged the utterances of the subject depending on mutual conversation, resulting in more responsive and more frequent to complaints of subject than caregivers.

Acknowledgement

This study was partly supported by Comprehensive Research on Aging and Health (H11-CRAH-039, H11-CRAH-042, H12-CRAH-041) from the Ministry of Health and Welfare; and the Research Grant for Longevity Sciences (10C-02, 11C-05) from the Ministry of Health and Welfare and the Creative, and Industrial Technology Research Grant Program in '00 from the New Energy and Industrial Technology Development Organization (NEDO) of Japan (00B59002d).

References

- [1] George S. Alexopoulos, Jonathan M. Silver, David A. Kahn, Allen Frances, Daniel Carpenter, "Treatment of Agitation in Older Persons with Dementia (Editors for the Guidelines)", The Expert Consensus Guideline Series, <http://www.psychguides.com/gagl.html>
- [2] Internet Mental Health : Disorder Research, <http://www.mentalhealth.com/dis-rs/frs-or05.html>
- [3] IBM ViaVoice Software Development Kit 1.5
- [4] Microsoft Speech API 4.0