Title: Natural language-based presentation of cognitive stimulation to people with dementia in assistive technology: a pilot study

Running head: Natural language-based cognitive stimulation

Keywords: assistive technology, natural language, dementia, cognitive stimulation

Authors and contact details:

Dr Nina Dethlefs
University of Hull
School of Engineering and Computer Science
Cottingham Road, Hull HU6 7RX
United Kingdom

Dr Maarten Milders
Vrije Universiteit Amsterdam, Depart-
ment of Clinical Neuropsychology Van
der Boechorststraat 1
1081 BT Amsterdam
The Netherlands

Dr Heriberto Cuayáhuitl
University of Lincoln
School of Computer Science
Brayford Pool, Lincoln, LN6 7T
United Kingdom

Turkey Al-Salkini
Heriot-Watt University
School of Mathematical and Computer Sciences
Edinburgh, Riccarton, EH14 4AS
United Kingdom

Lorraine Douglas Heri-
ot-Watt University
School of Life Sciences
Edinburgh, Riccarton, EH14 4AS
United Kingdom

Corresponding author: n.dethlefs@hull.ac.uk

Abstract: Currently, an estimated 36 million people worldwide are affected by Alzheimer's disease or related dementias. In the absence of a cure, non- pharmacological interventions, such as cognitive stimulation, which slow down the rate of deterioration can benefit people with dementia and their caregivers. Such interventions have shown to improve well-being and slow down the rate of cognitive decline. It has further been shown that cognitive stimulation in interaction with a computer is as effective as with a human. However, the need to operate a computer often represents a difficulty for the elderly and stands in the way of widespread adoption. A possible solution to this obstacle is to provide a spoken natural language interface that allows people with dementia to interact with the cognitive stimulation software in the same way as they would interact with a human caregiver. This makes the assistive technology accessible to users regardless of their technical skills and provides a fully intuitive user experience. This article describes a pilot study that evaluated the feasibility of computer-based cognitive stimulation through a spoken natural language interface. A prototype software was evaluated with 23 users, including healthy elderly people and people with dementia. Feedback was overwhelmingly positive.

Natural language-based presentation of cognitive stimulation to people with dementia in assistive technology: a pilot study

Introduction

Alzheimer's disease and related dementias are progressive degenerative conditions with no known cure. Important symptoms include memory loss, cognitive impairment, difficulty communicating and changes in mood and these symptoms become gradually worse over time. Age is the main risk factor for dementia and with an ageing population, Alzheimer's disease and related dementias are expected to become a major health burden in the coming decades [1] . Given the increasing costs of Alzheimer's disease and related dementias, the World Alzheimer Report 2013 emphazised the urgency of developing cost-effective care for people with dementia.

In the absence of an effective cure for dementia, interventions that slow down progression of the disease can benefit people with dementia and their caregivers, and reduce the pressure on health and social care services (e.g. hospital admission, day care, professional home care or admission to nursing home). A small number of drugs (e.g. donepezil, rivastigmine, galantamine, memantine) can be prescribed to alleviate the symptoms or slow down progression of dementia. However, these effects are only temporally and medication is only one part of the dementia care. Non-pharmacological interventions (NPIs) are recognised as valuable alternatives or complements to medication [1]. Recent reviews indicated that NPIs containing cognitive stimulation (CS) were particularly promising for maintaining cognitive functioning and well-being in people with dementia [2, 3]. In CS people with dementia engage in activities aimed at stimulating cognitive and social functioning, e.g. quizzes, memory activities or conversations about childhood. CS is typically provided by a health professional. The largest randomised controlled trial to date into CS by a professional therapist found improvement in cognitive functioning (memory, attention, language) and quality-of-life in people with dementia in the intervention group compared to the no-treatment control group [4]. The effectiveness of the intervention, expressed in the number needed to treat, was comparable to that of the commonly used medication for people with dementia. Separate studies showed that CS in combination with medication was more effective in slowing cognitive decline in people with dementia than medication alone [5, 6].

For a chronic condition like dementia, effective interventions should have long-term effects. The positive effects of CS tend to disappear within weeks after cessation of the intervention [7], and to maintain the effects continuous stimulation would be required [8, 9]. Long-term intervention by professional therapists is prohibitively expensive, but CS provided by relatives or other family caregivers is a potential alternative, which could also improve access to the intervention. To date, a number of studies has evaluated CS by caregivers and generally found positive effects on cognitive functioning [5, 6, 13] and no evidence for negative effects on well-being in the person with dementia [14]. However, while involving caregivers might be a sustainable option for long-term intervention, not all people with dementia are supported by a caregiver and providing long-term CS can further increase the burden on caregivers [14]. Computer-assisted NPIs for people with dementia provide an alternative method to present long-term stimulation and training at home without greatly increasing the burden on caregivers.

Computer-based cognitive interventions are based on computer software that aims to provide cognitive interventions, e.g. cognitive stimulation therapy, to people with dementia in a way that is similar to how it is often provided by a human caregiver. Such computer-based programmes can come in different forms, e.g. running on a desktop computer or on a mobile phone or tablet. Particularly the latter form of cognitive intervention presentation has received increased attention of late. There appear to be two main types of commercial software: (1) those that provide cognitive stimulation, and (2) those that focus on the user experience and apply reminiscence therapy. The first category includes systems such as Smartbrain[1] or Captain's log[2], which have both been clinically tested. In an attempt to make the application engaging, the latter offers users to take breaks and engage in a video game—-which is unrelated to their therapy, however. The second category includes software such as iReminisce[3] and Memory Box[4]. These apps provide multi-media materials, including photos, videos and music that can be combined manually with private material to create a collection of memories that patients can browse through, extend or talk about. Other software packages, such as Circa Connect[5] or MyLife[6] aim to help pationts engage in conversation with other patients, their relatives or carers rather than doing their activities alone. Again, a multimedia archive of materials is included for users to converse about. While this idea is similar to our goal of embedding activities within a spoken natural language context, contrary to our approach, it also relies on an active conversational partner to be physically present.

In terms of effectiveness to improve cognition, computer-based cognitive interventions that rely on cognitive stimulation have made good progress in offering interventions for healthy older people and people with mild cognitive impairments or dementia. Computer-based interventions can typically address different cognitive func-

---

[1] http://www.smartbrain.net/sb_new/

[2] http://www.braintrain.com/captains-log-mindpower-builder/

[3] http://ireminisce.co.uk/

[4] http://memoryboxnetwork.org/

[5] http://www.circaconnect.co.uk/

[6] http://www.mylifesoftware.com/

tions depending on specific individual needs. Cipriani et al. [15] presented a study using software for cognitive stimulation that targeted both the retrieval of known information as well as learning new information. In an experimental study they demonstrated that daily cognitive training with the software led to significant improvements in people with dementia, particularly in the areas of memory (long term, short term, verbal and visual), perception (recognition and identification) and attention. While the results reported in the study were still preliminary, positive effects on the global cognitive functioning of the participants with dementia were observed [15]. A related computer- based study on cognitive training in older participants with mild cognitive impairment observed similar positive effects on cognitive functioning and memory [16]. As much as the cognitive functions targeted and input modalities can vary, activities can also be adapted to suit a therapy's needs or a user's interests. While quiz activities are common, other activities are possible that users are familiar with from their daily lives. For example, Hoffmann et al. [17] present cognitive training software that simulates shopping on the high street experiences with people with dementia. The study found that participants enjoyed the activity and that their task performance— finding a predefined shopping route, buying three items and answering 10 multiple choice questions relating to the route—had improved. To date there has been little research into the effectiveness of commercial reminiscence software to improve cognitive functioning. However, reminiscence presented by a person in small groups has shown to be effective for improving cognition and daily functioning in people with dementia [3]. Furthermore, cognitive stimulation therapy can also comprise a reminiscence component ([8, 11].

Technological interventions such as described above reduce the risk of burdening caregivers and have been shown to be comparably effective to interventions presented by human caregivers [18, 19]. Advantages of computerised interventions over standard mode are the potential cost effectiveness, the possibility of individualisation and improved accessibility. However, currently existing assistive technologies that support cognitive stimulation all require their users to physically manipulate a computer, e.g. through mouse clicks, keyboard commands or touch screen. While the technologies have shown to be clinically effective [10, 11], the absence of an intuitive and accessible interface for the less technically skilled represents a major obstacle to wide-spread adoption [12]. This may be partially because communication through a keyboard or mouse provides a very different user experience than the social interactions normally achieved through spoken language. The natural spoken language interface that we propose provides a natural means of interaction between users and system: users can speak to the assistive technology similarly as they would normally speak to a human.

We describe an experimental setup based on a Wizard-of-Oz interface that aims to investigate whether the spoken language interface is feasible for interactions with elderly people and people with dementia in particular.A Wizard-of-Oz interface can be used to simulate a set of targeted computer behaviours and make the computer appear to behave autonomously, when it is in fact controlled by a human confederate behind the scenes. This setup can be useful to explore an initial set of potential computer behaviours, one or several of which are then implemented in a later version of the software. The main aim of this pilot study was to investigate the general acceptability of computer-based cognitive stimulation in people with dementia and their attitudes towards such software.

Using a Wizard-of-Oz setup, we report a study involving healthy elderly people and elderly people with mild to moderate dementia, which shows that participants found it easy and natural to interact with the system and overall enjoyed interacting with the system. The long-term aim of our research is to establish whether providing a spoken natural language interface for users to access the technology can make a difference towards wide-spread adoption of computer-based cognitive stimulation and thus have positive effects on the health of elderly people. The current study presents initial results on the feasibility of spoken language technology for elderly people with and without dementia.

Method and materials

Participants

Twenty-three participants took part in our study, of which 13 were healthy elderly people (9 men, 4 women) and 10 were people with dementia (8 men, 2 women). Mean age in the healthy group was 84.33 years (range 81 – 89). Mean age in the group of people with dementia was 78.20 years (range 57 – 92). The age difference between the groups was significant ($t=2.31$, $p<.05$).

Healthy participants were included in the study to ensure that the setup was suitable for elderly persons before it was presented to elderly, participants with dementia. The healthy participants were recruited from a subject panel of healthy elderly people who are interested to take part in research which is held by the Department of Geriatric Medicine (University of Edinburgh). The participants with dementia were recruited through the Scottish Dementia Clinical Research Network (SDCRN). The SDCRN maintains a research register containing names and contact details of people with dementia and their main caregiver who can be approached with requests to participate in research. Only participants with mild to moderate dementia (Clinical Dementia Rating < 3) were invited to participate in this study. The 5-point CDR scale has been developed especially for rating dementia severity, has high validity and reliability and is commonly used in clinical research and practice [31]. Mild to moderate severity on the CDR scale can be mapped onto Mini Mental State Examination (MMSE) severity levels [30, 31], corresponding with the mild (21-25) and moderate (11-20) range on the MMSE.

The mean Clinical Dementia Rating (CDR) of the 10 participants with dementia was 0.67 (SD 0.25). Seven of the people with dementia had been diagnosed with Alzheimer's disease, two with early onset dementia, and 1 with dementia with Lewy bodies.

For both the participants without dementia and the participants with dementia, the presence of major additional diseases (e.g. stroke, Parkinson's disease, depression) or inability to give informed consent were exclusion criteria. For the persons with dementia, type of dementia (e.g. Alzheimer's disease, vascular dementia, dementia with Lewy bodies) was not an exclusion criterion. All participants from the healthy group attended the test session alone. The participants with dementia were all accompanied by a relative. Since our study involved vulnerable human participants, it was approved by the South East Scotland Research Ethics Committee prior to its undertaking.

Research design

To establish the feasibility of our technology, we used a Wizard-of-Oz (WoZ) interface. This is a computer programme, which creates the impression of behaving autonomously even though it is actually controlled by a human experimenter (the wizard) behind the scenes. A WoZ interface is often used in experimental settings in linguistics, psychology or usability engineering to find out how a particular system behaviour is perceived by human users [e.g. 20, 21]. To this end, unknown to the participant, the wizard controls all communication between the user and the system by choosing buttons that will lead to a specific system behaviour being displayed. In this way, different system behaviours can be explored that are later implemented in the system. In our specific scenario, e.g., a human wizard would operate a graphical user interface (GUI) behind the scenes of the experiment. The human wizard would listen to a participant's spoken utterances, interpret them and then click a button in the GUI to prompt a system action. The system would then display the behaviour chosen and respond back to the participant using spoken language. In this way, the participant receives the impression of interacting with an autonomous system.

Figure 1 shows an architecture of a conventional spoken interactive system, which interacts with human users in five main steps: (1) processing incoming user speech by mapping sound signals onto a string of words (Automatic Speech Recognition), (2) mapping the words onto a meaning representation understood by the system (Natural Language Understanding), (3) choosing actions to take, e.g. cognitive stimulation activities to present from an existing database (interaction management), (4) formulating a string of words to accompany the chosen action to the user (Natural Language Generation) and, finally, (5) presenting the string of words to the user by mapping the words back onto sound signals (Speech Synthesis). While in conventional spoken interactive systems all five stages are performed autonomously by different system components, in our Wizard-of-Oz interface the wizard took charge of the first three stages, Automatic Speech Recognition, Natural Language Understanding and Interaction Management. Natural Language Generation was handled through manually pre-specified text templates, and Speech Synthesis was performed through the off-the-shelf Mary TTS software [22]. Later versions of our software will explore a fully automatic spoken interactive system. Please see [23] for an overview of spoken interactive systems.

[FIGURE 1 ABOUT HERE]

For our pilot study, we chose a WoZ setting because it gave us the opportunity to react promptly to changing behaviour of participants, e.g. if they seem to struggle with a specific stimulation item, appear confused or wish to end the interaction. Provided that the interactions between participants and system lead to positive feedback, the wizarded interactions can also inform the design and functionality of an autonomous assistive technology as a next step. This way we could motivate system behaviour directly based on a human expert performing the same task. To ensure that the wizard made appropriate choices for the elderly participants, some of whom were affected by dementia, human wizards had previous experience in providing cognitive stimulation on a human-to-human basis. Participants in our study were initially kept unaware of the fact that the system did not actually behave autonomously. This was done in order to elicit natural user behaviour. All participants were informed of the wizard on conclusion of their session with the system.

[FIGURE 2 ABOUT HERE]

Figure 2 provides an illustration of the WoZ interface used. The left-hand side of the interface showed video input of the participant's face (top) and the screen that the participant saw (bottom). In this way, the wizard could see the participant at all times and react to changing situations promptly. The right-hand side shows the actions that the system was able to perform. The wizard chose from these during interactions. Example actions are "greeting", "present activity", "acknowledge response", etc. All of these prompted the system to perform an action and / or say something to the participant using synthesised speech. In addition to these pre-specified actions, the wizard also had a text option available. Choosing this would pop up a text window through which the wizard could send messages (to be spoken/synthesised by the system) to the participant. This option could be used to give hints on activities, e.g. when participants appear to have difficulties in finding a response. The text option turned out to be an important feature in our WoZ interface because it gave the human wizard the flexibility to make interactions as stimulating and natural as possible.

Procedure

The experimental setup is shown graphically in Figure 3. The participant and the wizard were located in separate but adjacent rooms, separated by a one-way mirror. Both were facing a computer screen. The participant did not need to touch the computer but controlled all interaction via speech. Visual stimuli, such as pictures, were presented to the participant on the computer screen. First, all the healthy participants completed the tasks, followed by the participants with dementia.

[FIGURE 3 ABOUT HERE]

Experimental interventions

Cognitive stimulation (CS) is based on stimulation of cognitive functions through selected exercises and practice. In CS the emphasis is on engaging participants in activities that are cognitively stimulating and on processing information. Providing the correct answer is less important in stimulating activities than engaging in the task. Therefore, participants never received negative feedback, regardless whether their answers were correct or not.
CS can target different areas of cognition, including memory, communication, attention, and logical thinking. In this pilot study, we focused on memory and communication presented in four categories of activities that targeted established knowledge: sorting, name recall, quiz and proverbs. All activities were chosen to require some effort but be achievable, even for most people with mild to moderate dementia.
All activities were presented to participants in spoken language. The questions read out by the computer voice were presented visually on the computer screen, together with any pictures that were part of the activity. The activities were based on a manual for caregivers to present CS activities to their relative with dementia. This manual had been developed in collaboration with health care professionals as well as people with dementia and their caregivers. The activities chosen for this study had been rated as enjoyable by people with dementia in a previous study [14].

In the *sorting activity*, a range of different objects was displayed as pictures on the screen. Two objects belonging to different categories, e.g. kitchen, bathroom, clothing, were presented on the screen and a computer voice asked the participant where a specific item belonged, for example in the kitchen or in the bathroom. The participants responded by saying out loud the name of one of the categories.

In the *name recall activity*, the participant was shown a photograph of a famous person and was asked to recall the name of the person. The wizard could use the text option of the WoZ interface to assist participants by asking questions or giving hints. For example, “This man used to be an American president.”

In the *quiz activity*, participants were presented with general knowledge questions alongside a set of possible answers and were asked to choose the correct one. An example is “What language is spoken in Brazil?” with possible answers “English”, “Portuguese” and “Spanish”. Again, the wizard could use the text option to provide hints.

In the *proverb activity*, the participant was presented with proverbs in which one or several words were missing. The participant was asked to complete the proverb and, as before, the wizard could choose to help using the text option. An example is “The early bird catches the …” with participants expected to say the word “worm”.

The order of presentation of the different activities was varied across the participants. Participants could stop the activities at any time, but all chose to continue and complete the full set, which took about 20 minutes. Following each category of activities participants were asked (1) whether they had enjoyed interacting with the system, and (2) whether they would like to use the system again in the future.

Results

All participants interacted with the system for about 20 minutes each, engaging in all four categories of cognitive stimulation activities. Following each category participants were asked whether they had enjoyed the activity and would like do it again in the future. The results of these ratings are summarised in Table I for the healthy participants and in Table II for the participants with dementia. The ratings are separated by activity category. One participant with dementia was not able to choose between the four ratings and no results are available for this participant.

Overall the large majority of the participants rated the activities and the interaction with the system positively; at least 11 out of 13 healthy participants and at least 8 out of 10 participants with dementia indicated that they had enjoyed doing the activities. The activities that received the highest preference of participants for enjoyment in both the healthy group and the people with dementia were *Quiz* and *Proverbs*, i.e. the communication activities (see Tables I and II). When asked whether they would like to do the activities again, for most activities a majority of the participants indicated that they would do the task again. In particular, *Proverbs* were popular in both groups. By contrast, a majority of the healthy participants preferred not to do the sorting tasks again, possibly because it was too easy. Among the participants with dementia, a majority indicated for each activity that they would like to do the activity again.

In further analysis of the video recordings collected during sessions, three researchers watched the videos independently and made a number of shared observations. It was noted that healthy participants were all quick and accurate to respond to all types of questions. For sorting questions, it was in some cases apparent that the participant found the task too easy. An example of this was the system repeatedly asking a user whether an item belonged on the fruit or vegetable shelf in a supermarket. For some healthy participants the answer was obvious and they seemed to enjoy these parts of the session less. However, reactions were positive for more challenging tasks such as the quiz or proverb activities.

In contrast, most participants with dementia also responded quickly and mostly accurately (although accuracy is less important in cognitive stimulation), and made the impression to enjoy answering the questions on the computer. It was again the case that the quiz and proverb activities were perceived as more challenging (and enjoyable) than the sorting the name recall activities. Comments made after the task confirmed these impressions for both groups of participants.

[TABLES I AND II ABOUT HERE]

All participants who were asked found it natural to speak to a computer and were able to understand the computer voice easily. None of the participants reported any adverse effects of interacting with the computer. In fact, one of the participants with dementia remarked that he preferred doing the activities with a computer, as the computer would not judge him when he made a mistake or did not know an answer. All participants with dementia came with a relative and the relatives of three people with dementia were surprised how well their relative with dementia performed on the activities, much better than they had expected.

Discussion

The aim of this study was to investigate the general feasibility of presenting computer-based cognitive stimulation to elderly people and people with dementia through the use of spoken natural language. Initial results show that this is possible. The majority of the participants confirmed that they enjoyed interacting with our system and might like to use it again in the future. These are promising results that confirm that spoken language may present an appealing medium for communication between computers and elderly people. It represents a more intuitive interface between users and technology than other forms of input, e.g. mouse clicks, keyboard manipulation or touch screens, as previous studies have relied on. None of the participants in our study experienced problems with the technology, e.g. not being able to understand the computer voice, or not being able to identify pictures of objects on the screen, and feedback on the medium of communication was very positive. Computer-based cognitive stimulation software that is based on spoken language has rarely been used before, and while initial observations are encouraging further research and experimentation is needed to confirm our results on a larger scale and to confirm that the positive evaluation of the system was based on its usability and content of the activities, rather than other factors such as the participants just enjoying the general experience of taking part in an experiment or interacting with a computer.

Particularly, further research will require the implementation of cognitive stimulation through spoken dialogue in an autonomous system, which is not controlled by a wizard behind the scenes. Such software should be easy and intuitive to use, preferably be started through a single click with all following interaction being controlled through spoken language. The system can be implemented both as a desktop application or as a mobile app to be used on mobile phones or tablets. The latter would directly link in with an already existing and developing market for assistive technology for people with dementia. As outlined above, commercially available technology on mobile phones or tablets focuses predominantly on reminiscence in order to appeal to patients through a variety of multimodal materials, including music, photos or video clips. A shortcoming of these applications is that reminiscence has not been shown to be clinically effective in contrast to cognitive stimulation. Our technology has the opportunity to close this gap and offer therapy that can be both engaging and effective. Spoken language can play a crucial role in the former aspect, engagement. While the application on mobile phones and tablets assumes that users are able to start the application, spoken language might provide a more socially-engaging user experience during the sessions that button clicks or screen touches alone. The current article represents only a first step in this direction. Future research will need to investigate the usability of an application in detail. Nonetheless, the fact the similar applications are already in use by people with dementia, makes us optimistic towards the usability of our proposed approach.

Particular challenges that we envisage in the development of a fully autonomous system are in interaction management and personalisation of system activities. While earlier results on automatic speech recognition have shown that elderly voices can be more difficult to recognise than middle aged voices, e.g. [24], this no longer represents a bottleneck. The reason for the earlier results was that many automatic speech recognisers were trained from voices of middle aged people, so that elderly voices displayed several characteristics, which the automatic speech recogniser had not encountered during training. Recent advances in commercial speech recognisers, such as e.g. available from Google or Microsoft, largely circumvent this problem, however, by training from very large datasets. In practice, automatic speech recognition should therefore be stable enough for the use with elderly users. In addition to training from larger datasets than before, research on automatic speech recognition has also benefitted from the exploration of deep neural network architectures that learn abstract hierarchical feature representations from large amounts of varied input [27, 28, 29]. Such abstract features allow the computational model to generalise over the speech features of particular individuals (who might be represented in the training set) and provide better recognition results across the board. This has recently been shown to work well for children voices [29], which similarly to elderly voices differ in certain characteristics from the standard speakers, who have traditionally been used for the training of speech recognition systems. The performance of speech recognition systems is typically measured by their Word-Error Rate (WER), which Google has recently reported to be as low as 13.5% in an open dictionary task on a mobile phone.

A follow-up study to the present article will need to investigate the particular difficulties that elderly patients might experience when interacting with our technology. All results reported above were obtained with a Wizard-of-Oz scenario, in which no serious problems were simulated or occurred. This was because we were predominantly interested in investigating the general acceptability of computer-based cognitive stimulation. Future research will need to take this investigation further and probe into peoples' attitudes in more realistic use scenarios in which errors or misrecognitions can occur. Such studies could also confirm whether the current positive evaluation of the system was indeed based on its usability and content of the activities. Other factors such as the safe setting of the study, a positive attitude towards the study in the participants who volunteered to take part could also have contributed to the positive evaluation in the current study.

A further aspect that future work should investigate is the potential for personalisation of system behaviour and activities. In terms of interaction management, an autonomous system would not only have to decide which action to take at any point during a conversation, but would ideally also be sensitive to the user's mood and sentiment. For example, when a user shows signs of distress, maybe wishing to end an interaction, a system that interacts with that user would need to be able to identify such cues and react accordingly. This problem can be challenging to address, but research on sentiment analysis from linguistic or visual information can provide a first approach [25, 26]. Such research also draws on data-driven techniques and trains algorithms to automatically detect signs of agitation or stress from a dataset of human recordings. A pre-trained algorithm could be integrated into our system in the future.

Finally, cognitive stimulation would ideally be personalised towards individual users. Cognitive stimulation can target various cognitive functions, including long-term and short-term memory, recognition and identification, attention and verbal capabilities. Decline of any of these areas can lead people with dementia to experience negative effects or isolation, which is why it is important to support users' in exactly the area that is most relevant to them. Such adaptation constitutes a further point for future research. When the system detects that a user consistently performs less well on activities targeting a specific cognitive function, then these activities could increasingly be presented in order to stimulate the user's cognitive function in a targeted way. This incidentally can also include language skills, as some patients face difficulties in the retrieval of words or semantic fields as their disease progresses. Future work can investigate whether positive contributions can be made in this area through the use of a spoken natural language interface to the assistive technology. The application could be extended to deliberate help patients to keep their language skills active, and spot any changes in linguistic behaviour.

A further aspect to explore is adaptation towards users' individual interests. It appears from our study that participants enjoyed particularly the activities that rely on well-established knowledge and semantic memory, such as the quiz and the proverb activities. In contrast, name recall and sorting were perceived less favourable. This was true for both the healthy participants and participants with dementia. Interestingly, most participants in our study were men so it is possible that the type of intervention we propose is more appealing to men than women. Aspects relating to personal interest and gender could be investigated in future research. These aspects could be partially important in order to make the system enjoyable to use. Most existing applications similar to ours focus on reminiscence therapy in order to appeal to patients. Cognitive stimulation, however, is the non-pharmaceutical intervention that has most consistently shown to be effective. We therefore hope to bridge an existing gap with our software, which is both effective and appealing through its use of language.

Limitations of the current study are that we only tested the system with a small number of participants. Future evaluations should be based on a larger sample. The number of activities presented to participants during their session was also limited. As a result, proper matching of the activities to the participants' interests and abilities was not possible in this pilot study. User evaluation of the setup and activities was based on two questions only and one question (whether they would like to use the system again in the future) may not have been suitable for some of the participants with dementia in being too abstract. Finally, while the WoZ scenario gave us initial insights into the feasibility of a spoken language interface in general, the system was still controlled by a human. In an autonomous system, it is more likely that speech recognition errors can occur or that the system chooses actions than a human wizard would never choose. Therefore, once an autonomous system has been developed, users' perception of that system would need to be established again.

In summary, this article has presented initial results on the feasibility of using spoken natural language to provide cognitive stimulation to healthy elderly people and people with dementia through a computer-based assistive technology. A very important aspect for future research is the investigation of more personalised strategies to provide CS. Personalisation can take two forms: (a) the system can adapt to a participant's cognitive needs and provide stimulation for the area most relevant to an individual; and (b) activities can be tailored to allow participants to speak more freely and relate activities to their personal memories and lives. Previous studies involving human-provided cognitive stimulation have reported that participants particularly enjoyed activities with a personal perspective [7, 8, 9, 11]. Current computer-based cognitive stimulation systems, however, do not offer personalisation to individual needs and interests.

Acknowledgements

This research was partially funded by the Royal Society of Edinburgh while all of the authors were based at Heriot-Watt University, Edinburgh, UK. We thank the Department of Geriatric Medicine at the University of Edinburgh and the Scottish Dementia Clinical Research Network for their assistance with participant recruitment.

References

[1] M. Prince, M. Prina, and M. Guerchet. Alzheimer's Disease International. World Alzheimer Report 2013. London: Alzheimer's Disease International.

[2] G. Livingston, K. Johnston, C. Katona, J. Paton, C. Lyketsos. Systematic review of psychological approaches to the management of neuropsychiatric symptoms of dementia. *American Journal of Psychiatry* 2005; **162**: 1996-2021.

[3] J. Olazarán, B. Reisberg, L. Clare *et al.*, Nonpharmacological therapies in Alzheimer's Disease: A systematic review of efficacy. *Dementia and Geriatric Cognitive Disorders* 2010; **30**: 161–178.

[4] L. Robinson, S. Iliffe, C. Brayne, et al. Primary care and dementia: 2. long-term care at home: psychosocial interventions, information provision, carer support and case management. International Journal of Geriatric Psychiatry 2010; 25: 657–664.

[5] M.P. Quayhagen and M Quayhagen. Testing of a cognitive intervention stimulation for demenia care-giving dyads. Neuropsychologial Rehabilitation 2001; 11:319-332.

[6] Moniz-Cook E, Agar S, Gibson G, et al. A preliminary study of the effects of early intervention with people with dementia and their families in a memory clinic. Aging & Mental Health 1998; 2: 199 – 211.

[7] D. Sitzer, E. Twamley, D. Jeste. Cognitive training in Alzheimer's disease: a meta-analysis of the literature. Acta Psychiatrica Scandinavica 2006; 114: 75-90.

[8] E. Aguirre, A. Spector, J. Hoe, et al. Maintenance Cognitive Stimulation Therapy (CST) for dementia: A single-blind, multi-centre, randomized controlled trial of Maintenance CST vs. CST for dementia. Trials 2010; 11: 46.

[9] M. Orrell, A. Spector, L. Thorgrimsen, et al. A pilot study examining the effectiveness of maintenance cognitive stimulation therapy (MCST) for people with dementia. International Journal of Geriatric Psychiatry 2005; 20: 446-451.

[10] E. Aguire, R.T. Woods, A. Spector and M. Orrell. Cognitive stimulation for dementia: a systematic review of the evidence of effectiveness from randomized controlled trials. Ageing Research Reviews 2013; 12(1):253-262.

[11] Spector A, Orrell M and Woods B. Cognitive Stimulation for the treatment of Alzheimer's Disease. Expert Review of Neurotherapeutics 2008; 8(5):751-757.

[12] Agudo-Prado S, Pascual-Sevillano A, and Fombona-Cadavieco J. Uses of Digital Tools among the Elderly. *Comunicar. Scientific Journal of Media Communication* 2012; 39:193–20.

[13] G. Onder, O. Zanetti, E. Giacobini, G.B. Frisoni, L Bartorelli, G. Carbone, P. Lambertucci, M.C. Silveri and R. Bernabei. Reality orientation therapy combined with cholinesterase inhibitors in Alzheimer's disease: randomised controlled trial. British Journal of Psychiatry 2005; 187:450-455.

[14] M. Milders, S. Bell, A. Lorimer, T. MacEwan and A. McBain. Cognitive stimulation by caregivers for people with dementia. Geriatric Nursing 2013; 34(4): 267-273.

[15] Cipriani, G., Bianchetti, A., Trabucchi, M. (2006). Outcomes of a computer-based cognitive rehabilitation program on Alzheimer's disease patients compared with those on patients affected by mild cognitive impairment. Archives of Gerontology and Geriatrics 43 (2006) 327–335.

[16] González-Palau, F., Franco, M., Bamidis, P., Losada, R., Parra, E., Papageorgiou, S. & Vivas, A. (2014) The effects of a computer-based cognitive and physical training program in a healthy and mildly cognitive impaired aging sample, Aging & Mental Health, 18:7, 838-846.

[17] Hofmann, M., Rosler, A., Schwarz, W., Muller-Spahn, F., Krauchi, K., Hock, C., Seifritz, E., 2003. Interactive computer-training as a therapeutic tool in Alzheimer's disease. Compr. Psychiatry 44, 213–219.

[18] A. Kueider, J. Parisi, A. Gross and G. Rebok. Computerized Cognitive Training with Older Adults: A Systematic Review. PloS ONE 2012, 7(7).

[19] V. Faucounau, Y.H. Wu, M. Boulay, J. de Rotrou, and S. Rigaud. Cognitive Intervention Programmes on Patients Affected by Mild Cognitive Impairment: Promising Intervention Tool for MCI? The Journal of Nutrition, Health and Aging, 14(1), 2010.

[20] M. Nalin, I. Baroni, I. Kruijff-Korbayova, L. Canamero, M. Lewis, A. Beck, H. Cuayahuitl, and A. Sanna. Children's Adaptation in Multi-Session Interaction with a Humanoid Robot. In Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN); 2012; Paris, France.

[21] V. Rieser and O. Lemon. Learning Effective Multimodal Dialogue Strategies from Wizard-of-Oz data: Bootstrapping and Evaluation. In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACT-HLT); 2008; Columbus, Ohio, USA.

[22] M. Schröder, M. Charfuelan, S. Pammi and I. Steiner. Open source voice creation toolkit for the MARY TTS Platform. Proceedings of INTERSPEECH 2011. International Speech Communication Association (ISCA).

[23] V. Zue and J. Glass. Conversational Interfaces: Advances and Challenges. Proceedings of the IEEE 2000. Vol. 80, No. 8, pp. 1166-1180.

[24] S. Anderson, N. Liberman, E. Bernstein, S. Foster, E. Cate, B. Levin and R. Hudson. Recognition of elderly speech and voice-driven document retrieval. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 1999. Phoenix, Arizona, USA.

[25] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs Up? Sentiment Classification using Machine Learning Techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP) 2002. Pennsylvania, Philadelphia, USA.

[26] V. Perez-Rosas, R. Mihalcea and L.P. Morency. Utterance-Level Multimodal Sentiment Analysis. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL) 2013. Sofia, Bulgaria.

[27] Ian McGraw, Rohit Prabhavalkar, Raziel Alvarez, Montse Gonzalez Arenas, Kanishka Rao, David Rybach, Ouais Alsharif, Hasim Sak, Alexander Gruenstein, Françoise Beaufays, Carolina Parada (to appear) Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE (2016).

[28] X. Lei, H. Lin, and G. Heigold, "Deep Neural Networks with Auxiliary Gaussian Mixture Models for Real-Time Speech Recognition," in Proc. of the IEEE In- ternational Conference on Acoustics, Speech, and Signal Processing (ICASSP), Van- couver, Canada, 2013.

[29] Hank Liao, Golan Pundak, Olivier Siohan, Melissa Carroll, Noah Coccaro, Qi-Ming Jiang, Tara N. Sainath, Andrew Senior, Françoise Beaufays, Michiel Bacchiani (2015) Large Vocabulary Automatic Speech Recognition for Children. INTERSPEECH. Dresden, Germany.

[30] Balsis, S., Benge, J., Lowe, D., Geraci, L. & Doody, R. (2015). How do scores on the ADAS-Cog, MMSE, and CDR-SOB correspond? The Clinical Neuropsychologist, 29, 1002-1009.

[31] Perneczky, R. Wagenpfeil, S., Komossa, K., Grimmer, T., Diehl, J., & Kurz, A. (2006). Mapping scoresonto stages: mini-mental state examination and clinical de mentia rating. American Journal of Geriatric Psychiatry, 14, 139-144.

FIGURES:

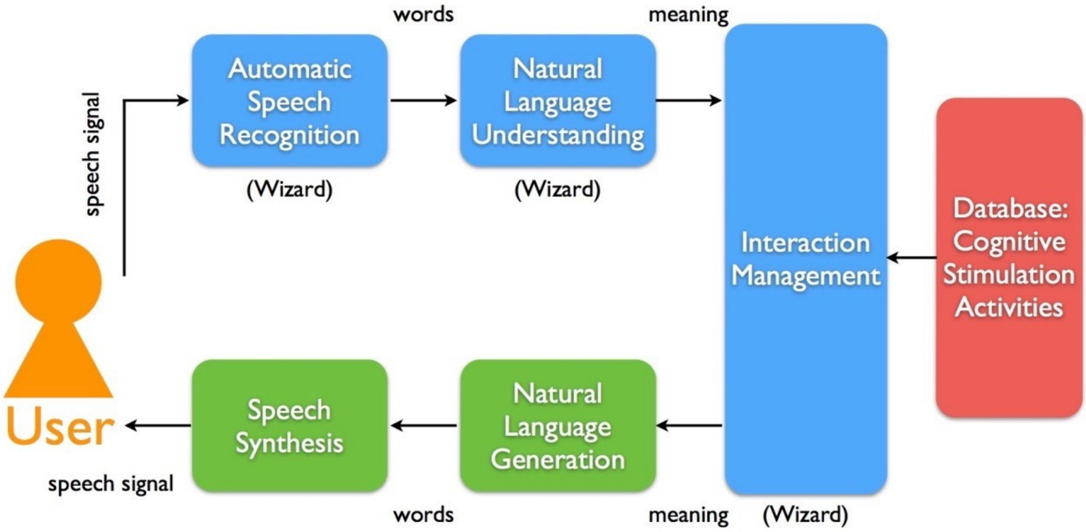

Figure 1: Architecture of a spoken interactive system involving stages for Automatic Speech Recognition, Natural Language Understanding, Interaction Management, Natural Language Generation and Speech Synthesis. In our Wizard-of-Oz scenario, the first three stages were controlled by the wizard.

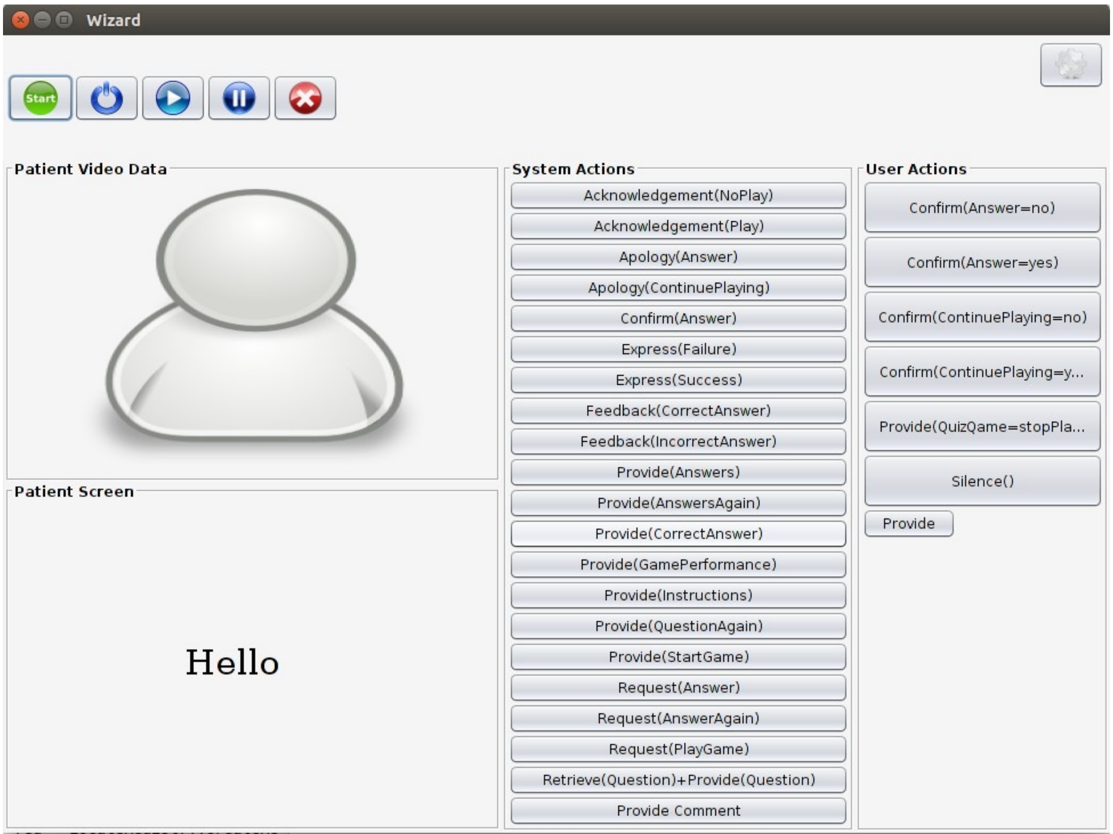

Figure 2: Screenshot of the WoZ interface which allows the wizard to observe the participant through a video stream during the experiment (top-left window), keep track of the activities displayed to the participant through a mirrored screen (bottom-left window) as well as control the interaction using the buttons on the right of the interface.

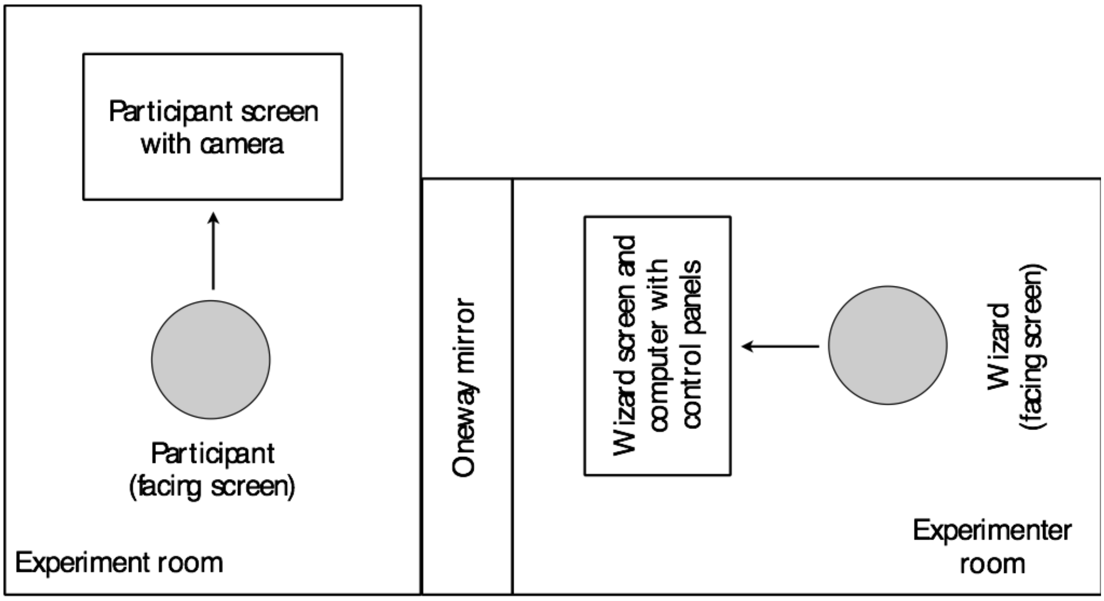

Figure 3: Experimental setup. Participant and wizard are seated at computers in different rooms that are separated by a mirror wall, so that the wizard can see the participant. The participant's screen displays visual material integrated in some of the activities, the wizard's screen shows the control panels to control the interaction and a video stream of the participant.

Activity	Enjoys using the system	Enjoys using the system a little	Does not enjoy using the system	Would use system in future	Would perhaps use system in future	Would not use system in future
Sorting	11	1	1	4	4	5
Name Recall	12	1	0	6	4	3
Quiz	13	0	0	8	2	3
Proverb	13	0	0	9	3	1

Table I: Results summarising the ratings of healthy elderly participants on whether they enjoyed using the system and would use it again the future per activity type. Results are based on 13 healthy elderly people.

Activity	Enjoys using the system	Enjoys using the system a little	Does not enjoy using the system	Would use system in future	Would perhaps use system in future	Would not use system in future
Sorting	9	0	0	5	2	2
Name Recall	8	1	0	4	3	2
Quiz	9	0	0	6	1	2
Proverb	9	0	0	8	1	0

Table 2: Results summarising the ratings of elderly people with dementia on whether they enjoyed using the system and would use it again the future per activity type. Results are based on 9 participants. One participant with moderate dementia was unable to choose a rating. excluded from the group.