Towards Attention Monitoring of Older Adults with Cognitive Impairment During Interaction with an Embodied Conversational Agent

Pierre Wargnier*
MINES ParisTech, PSL Research University

Adrien Malaisé
CEN Stimco

Julien Jacquemot
Broca hospital, Université Paris Descartes

Samuel Benveniste
MINES ParisTech, PSL Research University, CEN Stimco

Pierre Jouvelot
MINES ParisTech, PSL Research University

Maribel Pino
Broca hospital, Université Paris Descartes

Anne-Sophie Rigaud
Broca hospital, Université Paris Descartes

ABSTRACT

Embodied conversational agents (ECAs) are virtual characters using verbal and non-verbal communication for Human-machine interaction. The aim of our research is to create an ECA-based user interface for assistive technologies targeting older adults with cognitive impairment. Our design methodology is a co-design *living lab* approach, collecting design guidelines through questionnaires, focus groups and user trials.

In this paper, we report on the results of the first phase of this iterative design process. We developed Louise, a semi-automatic ECA prototype that aims to compensate, through attention monitoring, for a user's attentional disorders by performing autonomous *prompting*, i.e., calling the user to regain his or her attention in case he or she got distracted. We evaluated the performance of Louise with a group of experts in assistive technologies and collected their feedback. Louise's simple attention estimator is more than 80% accurate. The system got quite positive reviews from users.

Keywords: Embodied conversational agents, assistive technologies, attention estimation, dementia

Index Terms: I.5.5 [Implementation]: Interactive systems; K.4.2 [Social Issues]: Assistive technologies for persons with disabilities—Handicapped persons/special needs

1 INTRODUCTION

Due to the increase in the number of people living with dementia, usually caused by Alzheimer's disease, and growing caregiver shortage, assistive technologies will become critical in dementia care. Effective assistive devices will be key to help people living with dementia stay at home for as long as possible, and to enable caregivers to deliver higher quality care to an ever larger target population with highly constrained resources. The World Health Organization predicted that the number of people living with dementia, worldwide, will exceed 100 million by 2050 [24], making it a public health priority.

Assistive technologies for older adults, in particular those living with dementia, are often computer-based. But while their functional capabilities are getting better thanks to very fast technological improvement, people with dementia and their caregivers still do not use them very much. Indeed, the two main factors influencing the real-world effectiveness of these systems are their usability and acceptance, the later of which is highly related to perceived usefulness. The user interface should thus be adapted to the person's special needs for the product to be both easy to use and enjoyable.

---

*e-mail:pierre.wargnier@mines-paristech.fr

The overall goal of our project is to develop a computer user interface for older adults with cognitive impairment based on an embodied conversational agent (ECA), called Louise, in order to make assistive technology's use more accessible to them. If dementias are usually associated with memory impairment, they also cause executive and attentional disorders [12], which come in addition to the diminution of attentional capabilities observed in normal aging [9]. These specific constraints call for a user-centric design approach.

To develop Louise, we adopted a co-design methodology, involving the target users and other stakeholders (mostly care professionals and the patients' helpers), following a living lab methodology [5, 1]. It is based on five basic principles:

Openness. The conception process is public and anyone considering himself or herself impacted can get involved.

Influence. All participants should have a balanced influence on the design. Once people were involved their opinion has to impact the design based on their importance regarding the product's deployment.

Reality. The conception and evaluation of the solution should be performed as ecologically as possible.

Value creation. The design process should create value for everyone involved, not only from an economic perspective but also at the social level.

Sustainability. Once developed, the solution should keep evolving and being re-evaluated thanks to new data from the field.

In sync with the principles of the co-design methodology, an evaluation must be performed at every stage of the life-cycle of the ECA-based solution.

Based on the experience of our team[1] in the evaluation of assistive technologies, we identified that maintaining the user's attention during the interaction with any assistive device is critical. By "attention", we mean the ability one has to focus on a specific stimulus or task. In prior assistive technology trials conducted with patients with dementia, we noticed that the experimenters kept redirecting the participants' attention towards the task that they had to complete.

In this paper, we thus focus on the first step of the iterative design process of Louise, in which we address the issue of attention monitoring. At this point, Louise is able to compensate, through an attention monitoring algorithm we developed, for a user's attentional disorders by performing autonomous *prompting*, i.e., using its synthetic voice to call the user to regain and redirect his or her

---

[1]The Louise project is hosted at the Broca Hospital in Paris, in the LUSAGE Living Lab [18].

attention in case he or she got distracted. Our prototype is a semi-automatic ECA that uses (1) a Microsoft Kinect sensor and the associated software development kit (SDK) for user tracking, (2) the Unreal Engine game engine for ECA character animation and (3) the Cerevoice text-to-speech synthesizer. Louise's simple attention estimation implementation turns out to be more than 80% accurate, and the overall system got quite positive reviews from users.

The paper is organized as follows. In Section 2, we present similar attempts and results from the literature that suggest the use of an ECA is a promising solution for assistive device-user interaction; in addition, we survey some of the literature on automatic monitoring of attention. Section 3 contains design and implementation details regarding the attention monitoring algorithm we developed for our prototype, Louise. In Section 4, we present an experimental evaluation of our attention management system with a group of experts in assistive technologies. Section 5 focuses on the design recommendations collected from assistive technology experts. In the last section, a conclusion and future work perspectives are presented.

2 RELATED WORK

Though the field of social computing devices has grown very fast in the past decade, a lot of research remains to be done about their use with cognitively impaired older adults. Attention monitoring, on the other hand, is linked to the wider field of user engagement in Human-machine interaction. In this related work section, we first review the potential applications of ECAs in elderly care and show why researchers think they are promising. Then, we present the research related to attention monitoring.

2.1 Embodied Conversational Agents in Elderly Care

Several authors have studied the use of ECAs with cognitively impaired older adults. Ortiz *et al.* [17] have performed an empirical study involving 3 groups, of 5 people each, with different levels of cognitive capabilities: normally aging, mild cognitive impairment (MCI) and moderate Alzheimer's disease (AD). They compared a virtual character with text accompanied with speech and text only. They found that the users in all groups performed significantly better when guided by the virtual character instead of other prompting modalities. They also observed that the virtual character was judged pleasant by the study participants.

Morandell *et al.* [14] evaluated the use of photographs on which the lips are animated with 10 cognitively impaired older adults. They tested it for likability, task performance, and attention-catching, comparing with a speech and text prompt. They found that the interface with the talking face was preferred, that it was linked to slightly better task performance for the most cognitively impaired participant, and that the users' attention was better maintained with the talking face than without it. In addition, the authors observed better understanding of the talking face than the disembodied voice. Similar results were observed in another study by the same team, involving 12 cognitively healthy older adults and 12 patients with MCI [15]. This later study was more focused on usability. MCI users performed as well as non-MCI users when interacting with a conversational agent, which shows the promise of ECAs for cognitively impaired older adults.

Carrasco *et al.* [6] proposed a system in which a female virtual character is rendered on a TV set. The interaction with the user is done by hitting keys on the remote control, a familiar object for the elderly. They conducted a validation trial with 21 persons with AD. They showed that all participants understood well the character and successfully interacted with her. They also found that all participants engaged well and that the human-like character led them to interact naturally (using verbal and non-verbal communication) even though the conversational agent was unable to process such inputs. In a study using the same interactive system, Diaz-Orueta *et al.* studied the influence of the user's level of cognitive impairment,

called cognitive status, on user interaction with the device. They concluded that cognitive assessment tests could guide the assistive device design, by taking into account the user's cognitive status [8]. More specifically, the measure of attentional capabilities and processing speed through cognitive assessment tests can bring useful information for a personalization of such a user interface. In [12], the authors have proposed to adapt the prompting modalities to the person's cognitive and perceptive abilities.

Recently, some ECAs using natural verbal and non-verbal interaction have been proposed for the assistance of older adults with cognitive impairment. In [19], Sakai *et al.* present a listener ECA that includes speech input processing and non-verbal feedback for people with dementia. In [10], Huang *et al.* propose to enrich the interaction with this ECA by using a life-logging system. This software was used in [16] for a preliminary study towards automatic monitoring of patients' with dementia evolution by assessing their conversational responsiveness to the ECA. An ECA for people living with dementia has also been developed by Yasuda *et al.* [25]. Their system differs by the appearance of the virtual character, which is based on a five-year-old child. An evaluation was conducted with 8 participants, all having Alzheimer's disease. The results show that the participants uttered 74% as many syllables when conversing with the virtual kid as when conversing with a human partner (26% less). In [23], Wilkis *et al.* have tested a prototype ECA, which includes speech recognition and emotional state estimation, for smart home applications with a brain-injured veteran. The test resulted in good communication with the test patient.

All these studies all show good appreciation of ECAs by AD or MCI patients. In addition, encouraging results in terms of task performance were found. This, as well as the fact that the human appearance of the virtual agents arouses natural communication in the user, makes us think that ECAs are a good solution for user interaction with assistive devices and cognitive prosthesis. However, there still are a lot of open questions regarding the designs of adapted ECAs for AD or MCI patients. Namely, the behavior of the character, the input and output modalities, the presentation of information and the automatic management of user engagement with the ECA still have to be investigated in depth. On this last subject, to the best of our knowledge, no work has specifically focused on attention management or user engagement in the specific case of people with dementia in interaction with ECAs. In the following section, we present some of the literature on engagement estimation in human-machine interaction.

2.2 Engagement in Human-Machine Interaction

Engagement was defined by Sidner *et al.* as *"the process by which interactors start, maintain and end their perceived connection to each other during an interaction."* [20]. Thus, engagement includes the attention and interest that a person pays to the interaction in which he or she participates. It is important during the communication to observe and to analyze the behavior of other participants in order to assess their engagement. By vocal or visual prompting, it is possible to bring back the attention of an interlocutor in the conversation [11]. The engagement recognition, in the context of an interaction with an ECA, allows to show to the user that the system analyses and understands his behavior. Sidner *et al.* showed in [20] that a robot or an ECA performing engagement behaviors could more easily mobilize the user's attention.

The visual focus of attention is an important source of information regarding engagement [2]. This information can be measured by the estimation of someone's gaze or face orientation. Several studies show how to determine the face's orientation to estimate the focus of attention in several contexts: in front of a computer screen [3], in interaction with an ECA [11, 7] or a mobile robot [21, 4]. In [13], the user's attention is estimated using several features: po-

sition and posture, face orientation, proximity from the system and smile detection. These features are combined to compute his or her attention level. In [21], the focus of attention is estimated using the mean face orientation over the t last images. Engagement is evaluated by analyzing backchannel behaviors, which correspond to actions realized during an interaction, like head nods or non-verbal vocal behaviors. In [22], the authors also stress that such backchannels are an important source of information to estimate user engagement. In [4], engagement, that the authors define as the intention to interact, is estimated with the help of a Microsoft Kinect sensor. They show that only 7 features are sufficient: the relative orientation of the body and shoulders, distance from the sensor's plane, the speed on the axis parallel to the sensor plane, the size of the head seen by the sensor, the x position of the head in the color image produced by the sensor, the activated sound beams and the sound source localization angle.

3 ATTENTION MONITORING USING THE LOUISE ECA

In this section, the implementation of attention monitoring in our ECA is presented. The current prototype of Louise is a semi-automatic system that allows to animate a female cartoon-like character called Louise, represented on Figure 1, and synthesize voice from text. The character only moves in an idle pose and moves its lips when speaking. The system includes attention monitoring and interaction management based on a pre-defined script and keyboard inputs of a (presumably hidden) operator. Such a semi-automatic approach is well suited for rapid design prototyping and scenario screening, before embarking on the development of an automatic full-fledged system.

Figure 1: Louise, our virtual character

3.1 System Description

The current Louise prototype is built from three software modules:

Attention estimator. A Microsoft Kinect sensor (version 1) and the Kinect for Windows software development kit (SDK) version 1.8 are used to track a user, compute an attention estimation value from the tracking data, and decide the user's attentional state, as detailed in Section 3.2.

Interaction manager. An interaction management module has been developed. It takes a dialog script as an input and uses the attention estimation states as well as keyboard inputs to determine the next utterance to be produced in a semi-automatic way. The scenario is composed mostly of yes/no questions and an operator presses the corresponding key depending on the user's answer to trigger the next utterance. When the person interacting with the ECA is considered distracted by the attention estimator, a prompt to regain his or her attention is triggered. The character stays in prompting mode until the attention of the user is detected again.

Behavior realizer. A behavior realizer has also been implemented, using the Unreal Engine game engine, coupled with the Cerevoice text-to-speech voice synthesizer. This module animates the virtual female character and produces vocal messages with synchronized lips animations.

The modules communicate as illustrated on Figure 2. The arrows stand for data exchanges. Note that the interaction manager requires feedback from the behavior realizer, mostly to know when the previous utterance is over.
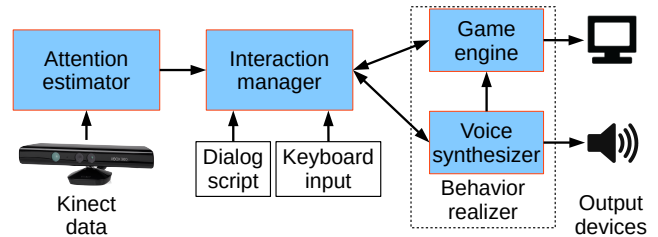

Figure 2: Functional diagram of the Louise ECA prototype

The system is, for the time being, semi-automatic. It takes, as input, a restrictive interaction scenario, composed of an introduction, several yes/no questions, backchannel utterances, prompting sentences, a transition question and a conclusion. The interaction starts by the introduction. Then, for each question, a key has to be pressed (Y for yes, N for no) to trigger a backchannel behavior and go on to the next question. The backchannel and prompting utterances are randomly selected from the corresponding lists. If the attentional state changes from *focused* to *distracted*, the ongoing behavior is automatically interrupted and a prompt behavior is produced. Once the user is considered by the attention estimator as attentive again, the transition question is asked. If the person wishes to continue the interaction, the last unanswered question is asked (again). If not, it jumps directly to the conclusion. When the character has gone through all the questions, the conclusion is triggered and the interaction ends. Lastly, if no user is tracked, the interaction cannot start and, if the user leaves, the character stops talking.

3.2 Attention Monitoring Algorithm

Older adults with cognitive impairment can be easily distracted by an external stimulus. Thus, the interaction between them and the ECA could be interrupted. We want to determine automatically if a person wants to communicate, and also if he or she remains attentive during the interaction. We implemented a method to estimate a person's attention, which is an important source of information in engagement (see Section 2.2). During an interaction, a person naturally faces his or her interlocutor to communicate with him or her. This is why we extract the posture of the user, and his or her face orientation, to estimate the level of attention he or she pays to the ECA. Our method is inspired by the works of Stricker *et al.* [21], Benkaouar *et al.* [4] and Li *et al.* [13]. We use the Microsoft's Kinect sensor and Kinect for Windows SDK to extract tracking data of the user's posture and position.

With the Kinect, a 20-point skeleton tracking is obtained, which allows to identify the user's upper-body pose. We extract the 3D position of the right and left shoulders to estimate if the person faces the ECA (see Figure 3 for notation). This method assumes that the sensor is placed in alignment with the ECA's display.

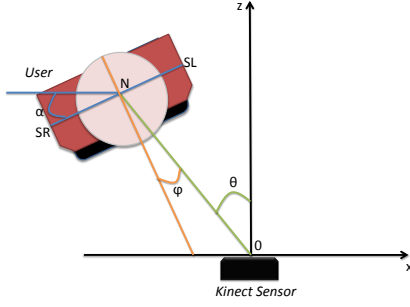The azimuth of the user is defined as $\theta = arctan(N_x/N_z)$ and the angle of the upper-body α as

Figure 3: Estimation of the angle φ between the shoulder line and the Kinect sensor. θ is the angular position of the user, SR the position of the right shoulder, SL the position of the left shoulder, and N the position of the neck (calculated as the center of the vector $[SR, SL]$).

$$\alpha = arcos(\frac{N_z - SL_z}{\sqrt{(SL_x - N_x)^2 + (SL_z - N_z)^2}}) - 90. \tag{1}$$

The posture feature f_1 is thus, at each time t:

$$f_1 = \varphi = \alpha - \theta. \tag{2}$$

In addition to the posture, we use the face tracking data of the Kinect for Windows SDK to obtain the face orientation of the user. The Kinect outputs, in real time t, the three rotations of the head: pitch $\gamma_{pitch}(t)$, yaw $\gamma_{yaw}(t)$ and roll. Only the two first parameters are used in our system. In order to improve reliability of the face orientation estimation from the sensor, we compute the mean of each measure over the last $T = 30$ frames, which correspond to about one second, given the sensor's sample frequency (see Equations 3 and 4); such a value is relevant in the context of natural interaction. This approach produces stable input values f_2 and f_3 for the attention monitoring algorithm at each time t, even if the face tracker fails for a few frames:

$$f_2 = yawMean = \frac{1}{T} \sum_{k=t-T+1}^{t} \gamma_{yaw}(k), \tag{3}$$

$$f_3 = pitchMean = \frac{1}{T} \sum_{k=t-T+1}^{t} \gamma_{pitch}(k). \tag{4}$$

The three features f_j ($j = 1, 2, 3$) are normalized as $\overline{f_j}$ in the same way:

$$\overline{f_j} = \frac{cos(f_j) - cos(Max_j)}{1 - cos(Max_j)}, \tag{5}$$

where Max_j represents the maximum value for each feature f_j. They correspond to the Kinect's tracking limitations (30° for yaw, 20° for pitch and 60° for upper-body pose).

A sum of the $n = 3$ normalized features $\overline{f_j}$, weighted by coefficients ω_j (see Equation 6), is computed to assess the *AttentionLevel* at each time. The face's horizontal rotation has the most important weight to account for the importance of the face orientation in the attention estimation. This corresponds to using this information as a proxy to the user's gaze direction. For normalization purposes, the sum of the weights is equal to 10:

$$AttentionLevel = \sum_{j=0}^{n} \omega_j \overline{f_j}. \tag{6}$$

The obtained attention level values range from 0 to 10, 10 being the maximum level, when the user's body and face are directly oriented towards the sensor. These values are then used to decide the user's attentional state.

There are four attentional states: *NoUser*, when no user is detected; *UserDetected*, when a user is detected for the first time and not considered attentive; *AttentionAcquired*, when the user is focused on the screen; *AttentionLost*, when the user is not focused anymore. Only some transitions between states are allowed:

- from *NoUser* to *UserDetected*;
- from *UserDetected* to *AttentionAcquired*;
- between *AttentionAcquired* and *AttentionLost*;
- from any other state to *NoUser*.

The transition rule uses two thresholds, a high and a low threshold, like a hysteresis system. The *AttentionLevel* value has to reach the high threshold for current state to change from *UserDetected* or *AttentionLost* to *AttentionAcquired*. The transition from *AttentionAcquired* to *AttentionLost* occurs when the low threshold is reached. This choice avoids the occurrence of some fluctuations between two states, which could possibly occur if a single threshold was used. Moreover, it allows considering as engaged a person who pays real attention to the ECA and considering the user disengaged in case of an important decrease of his or her measured *AttentionLevel*. During our experiments we set the high threshold to 8 and the low one to 6.

4 ATTENTION MONITORING ALGORITHM VALIDATION

To validate our attention estimation algorithm before its integration in an automatic system and according to co-design principles, we conducted a semi-automatic "Wizard-of-Oz" experiment at the Broca hospital in Paris with members of our research team and outside experts. The sample was composed of 14 participants (4 men and 10 women), naive to the goal of the experiment, specialists in assistive technologies for older adults or care professionals (medical doctors and neuropsychologists, mostly). In this section, we first describe the test protocol and then present the evaluation results. We discuss experts' recommendations in Section 5.

4.1 Experimental Protocol

In this first experiment, the participants were seated about two meters from a 46-inch screen on which Louise, our virtual character, presented on Figure 1, was displayed in full screen, standing in the middle of the display. The Kinect sensor was placed, centered, on top of the screen. An experimenter was in the room with the participant, outside of the field of view of the Kinect sensor, behind a control screen, to discreetly press the keys corresponding to the answers given by the participant (Y key for yes, N key for no). All experiment sessions were recorded by logging the tracking data, the computed features, the attentional states, and the color image produced by the sensor. In addition, a camera was placed behind the participant, filming the screen and recording the sound. A sample was recorded every second.

The interaction consisted of 3 utterances for the introduction, 7 questions, 4 backchannel utterances, 3 prompting utterances and 2 utterances for the conclusion. During the interaction, two distractions were voluntarily introduced at fixed moments. The first one was introduced by the experimenter in the room at the beginning of the third introduction utterance: he asked the participant if the sound was loud enough. The second distraction was introduced by another experimenter, opening the room's door and asking the participant if everything was fine, during the fifth question.

All recorded data was, later on, manually annotated by two specialists, using the videos taken from the Kinect and the camera. The

annotations consisted of the observed attentional states of the participants. They were then used as ground truths to compare with the decisions taken by the system. When the two annotations did not match, a third annotator decided which was right.

After the interaction with the ECA, a satisfaction questionnaire was administered to collect the participants' expert opinions about Louise's design. Their feedbacks are discussed in Section 5. The following subsection presents a performance evaluation of the attention estimator.

4.2 Results

All participants interacted naturally with the ECA. Most displayed high levels of attention. For some users the experimenter introducing the second distraction had to insist to distract the participant. This shows that the ECA is significantly engaging. The interaction lasted between 2 minutes 18 seconds and 5 minutes 16 seconds.

To evaluate the performance of the attention estimator, a receiver operating characteristic (ROC) curve was computed for several threshold values, ranging from 0 to 10 with a step of 0.5. The obtained curve is presented on Figure 4. The area under the ROC curve is about 0.8, which indicates that our algorithm is a quite good classifier. With the two attention level thresholds set to 8 for the high value and 6 for the low one, we obtained an accuracy of 82.78%. This could be slightly improved by setting the high threshold to 8.5. As already mentioned, using two decision thresholds is good to limit the amount of false positives when the user is not attentive to the ECA and reduces the number of false negatives when the user is considered engaged in the interaction, which could produce undesirable interruptions.
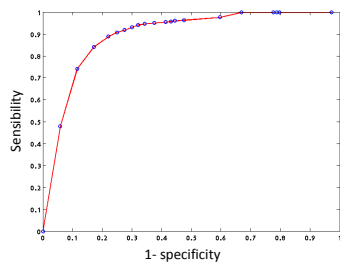

Figure 4: ROC curve of the attention estimator.

Overall, whenever people disengaged from the interaction, the avatar stopped talking with little latency. The system failed intermittently with only one participant. In this particular instance, the Kinect's face tracker was unable to recognize her face properly and tracked instead her hair, which was curly and tied on top of her head in a big bun. The Kinect's face tracker is one of the main limitations of our system as it fails when the user turns his or her head more than $60°$ from the normal to the sensor.

5 Discussion

During the experiment described in Section 4, we asked our participants, who were assistive technology experts, for feedback. After they had interacted with the virtual agent, we gave them a 5-questions questionnaire with lots of freedom for observations during an open interview. The questions were about how clear the questions asked by the character during the interaction were, how much (on a 1-to-4 scale) they liked the character's behavior, how they felt during the interaction, if they think the character reacted well to their behavior and whether the length of the interaction was appropriate.

For the first question about understandability, 10 out of 14 participants thought it was clear enough, 2 did not answer and 2 said no. One of them commented that it would turn out to be probably too complicated for cognitively impaired older adults.

Regarding likability, 5 participants appreciated the character a lot, 6 thought it was alright and 2 said it was not good enough. Most of the ones who said it was alright or not good enough commented that the lack of facial expression and blinking made them feel uneasy. In addition, one participant advised us to add directed gaze to show engagement on the character's side. Finally, several participants thought that Louise should smile.

As far as the third question about mood is concerned, all participants felt the interaction interesting (6) or very interesting (7). No additional comments were given.

Question 4, regarding how the character adapted her behavior to their's, got more divergent opinions. 6 of them judged it reacted very well, while 3 said it was fine and 4 considered it a bit poor.

Lastly, 7 participants thought the duration of the interaction was adapted, 4 said it was too short and the last 2 didn't have an opinion.

Globally, as can be seen, the feedbacks were positive. The main improvement points that were raised were about the animation of the character's face and reactivity to their behaviors.

6 Conclusion and future work

In this article, we presented a first prototype of Louise, an ECA targeting older adults with dementia. The proposed system is currently centered on the key issue of attention management. A new attention estimator was developed, tested with assistive technology experts and validated; we found good accuracy in attention estimation. In addition, we gathered feedbacks from experts that turned out to be rather positive. The main limitations we could identify are linked to the system tracking capabilities, to improve the attention estimator, and the character's animation, linked to user experience.

As future work, we are currently in the planning phase for running the same experiment with dementia patients to further validate the attention estimation method and obtain useful data on how they interact with the ECA. So far, we have performed an informal trial with one patient from the Broca hospital; she seemed to engage well with the character, but we could not record the data. Thus, experimentations with the target user group are clearly required.

With the gathered data, we will be able to work on the two axes of improvement we identified: character animation and robustness of attention estimation to tracking failures. A new prototype with a better animation system is already under development. Lastly, we have to collect information about older adults' preferences regarding the appearance and animation of the ECA, what applications they would like it to be included into and under what conditions they would accept such a system in their homes.

References

[1] M. Agogué, G. Comtet, P. Le Masson, J.-F. Menudet, and R. Picard. Managing innovative design within the health ecosystem: the living lab as an architect of the unknown. *Revue Management et Avenir Santé*, 1(1):17–32, Oct 2013.

[2] S. Asteriadis, K. Karpouzis, and S. Kollias. The importance of eye gaze and head pose to estimating levels of attention. In *Third International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, pages 186–191, May 2011.

[3] S. Asteriadis, K. Karpouzis, and S. Kollias. Visual focus of attention in non-calibrated environments using gaze estimation. *International Journal of Computer Vision*, page 124, Dec. 2013.

[4] W. Benkaouar and D. Vaufreydaz. Multi-sensors engagement detection with a robot companion in a home environment. In *Workshop on Assistance and Service robotics in a human environment*

at IEEE International Conference on Intelligent Robots and Systems (IROS2012), page 4552, Oct 2012.

[5] B. Bergvall-Kareborn and A. Stahlbrost. Living lab: an open and citizen-centric approach for innovation. *International Journal of Innovation and Regional Development*, 1(4):356–370, Jan. 2009.

[6] E. Carrasco, G. Epelde, A. Moreno, A. Ortiz, I. Garcia, C. Buiza, E. Urdaneta, A. Etxaniz, M. F. Gonzlez, and A. Arruti. Natural interaction between avatars and persons with alzheimer's disease. In K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, editors, *Computers Helping People with Special Needs*, number 5105 in Lecture Notes in Computer Science, pages 38–45. Springer Berlin Heidelberg, Jan. 2008.

[7] J. Chen, D. Chen, X. Li, and K. Zhang. Towards improving social communication skills with multimodal sensory information. *IEEE Transactions on Industrial Informatics*, 10(1):323–330, Feb 2014.

[8] U. Diaz-Orueta, A. Etxaniz, M. F. Gonzalez, C. Buiza, E. Urdaneta, and J. Yanguas. Role of cognitive and functional performance in the interactions between elderly people with cognitive decline and an avatar on TV. *Universal Access in the Information Society*, 13(1):89–97, Mar. 2014.

[9] A. D. Fisk, W. A. Rogers, N. Charness, S. J. Czaja, and J. Sharit. *Designing for Older Adults: Principles and Creative Human Factors Approaches*. Human Factors & Aging Series. CRC Press, second edition, 2009.

[10] H.-H. Huang, H. Matsushita, K. Kawagoe, Y. Sakai, Y. Nonaka, Y. Nakano, and K. Yasuda. Toward a memory assistant companion for the individuals with mild memory impairment. In *IEEE 11th International Conference on Cognitive Informatics and Cognitive Computing (ICCI*CC)*, pages 295–299, Aug. 2012.

[11] R. Ishii, Y. I. Nakano, and T. Nishida. Gaze awareness in conversational agents: Estimating a user's conversational engagement from eye gaze. *ACM Transactions on Interactive Intelligent Systems*, 3(2):11:111:25, July 2013.

[12] J. Lapointe, A. Verreault, B. Bouchard, J. Bouchard, A. Potvin, and A. Bouzouane. How to maximize the effectiveness of prompts in assistive technologies according to the particular cognitive profile of people with alzheimer? *International Journal of Smart Home*, 7(5):19–38, Sept. 2013.

[13] L. Li, Q. Xu, and Y. K. Tan. Attention-based addressee selection for service and social robots to interact with multiple persons. In *Proceedings of the Workshop at SIGGRAPH Asia*, WASA '12, pages 131–136, New York, NY, USA, Nov 2012. ACM.

[14] M. M. Morandell, A. Hochgatterer, S. Fagel, and S. Wassertheurer. Avatars in assistive homes for the elderly. In A. Holzinger, editor, *HCI and Usability for Education and Work*, pages 391–402. Springer Berlin Heidelberg, Jan. 2008.

[15] M. M. Morandell, A. Hochgatterer, B. Wöckl, S. Dittenberger, and S. Fagel. Avatars@home. In A. Holzinger and K. Miesenberger, editors, *HCI and Usability for e-Inclusion*, number 5889 in Lecture Notes in Computer Science, pages 353–365. Springer Berlin Heidelberg, Jan. 2009.

[16] Y. Nonaka, Y. Sakai, K. Yasuda, and Y. Nakano. Towards assessing the communication responsiveness of people with dementia. In *Proceedings of the 12th International Conference on Intelligent Virtual Agents*, IVA'12, pages 496–498. Springer-Verlag, Sep 2012.

[17] A. Ortiz, M. del Puy Carretero, D. Oyarzun, J. J. Yanguas, C. Buiza, M. F. Gonzalez, and I. Etxeberria. Elderly users in ambient intelligence: Does an avatar improve the interaction? In C. Stephanidis and M. Pieper, editors, *Universal Access in Ambient Intelligence Environments*, number 4397 in Lecture Notes in Computer Science, pages 99–114. Springer Berlin Heidelberg, January 2007.

[18] M. Pino, S. Benveniste, R. Picard, and A.-S. Rigaud. User-driven innovation for dementia care in france: The lusage living lab case study. *Interdisciplinary Studies Journal*, 3(4):251–268, 2014.

[19] Y. Sakai, Y. Nonaka, K. Yasuda, and Y. Nakano. Listener agent for elderly people with dementia. In *7th IEEE/ACM International Conference on Human-Robot Interaction (HRI)*, pages 199–200, Mar. 2012.

[20] C. L. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich. Explorations in engagement for humans and robots. *Artificial Intelligence*, 166(12):140–164, Aug. 2005.

[21] R. Stricker, S. Hommel, C. Martin, and H.-M. Gross. Realtime user attention and emotion estimation on a mobile robot. In *Proceedings of 55th Ilmenau International Scientific Colloquium*, page 1317, Sep 2010.

[22] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder. Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Transactions on Affective Computing*, 3(1):69–87, Aug 2012.

[23] Y. Wilks, J. M. Jasiewicz, R. Catizone, L. Galescu, K. M. Martinez, and D. Rugs. CALONIS: An artificial companion within a smart home for the care of cognitively impaired patients. In C. Bodine, S. Helal, T. Gu, and M. Mokhtari, editors, *Smart Homes and Health Telematics*, Lecture Notes in Computer Science, pages 255–260. Springer International Publishing, Jan. 2015.

[24] World Health Organization and Alzheimer's Disease International. *Dementia: a public health priority*. 2012. [Online] available at: http://www.who.int/mental_health/publications/dementia_report_2012/en/.

[25] K. Yasuda, J.-I. Aoe, and M. Fuketa. Development of an agent system for conversing with individuals with dementia. In *The 27th Annual Conference of the Japanese Society for Artificial Intelligence*, Jun 2013.