Towards a Natural Interaction Concept for SUI in AAL

Kizito Ssamula Mukasa
Fraunhofer IESE
Fraunhofer-Platz 1, D-67663 Kaiserslautern, Germany
kizito.mukasa@iese.fraunhofer.de

Christian Graf
CINACS Graduate Research Group
University of Hamburg
Vogt-Kölln-Str. 30, D-22527 Hamburg, Germany
graf@informatik.uni-hamburg.de

ABSTRACT
One of the important requirements for the user interaction in the Ambient Assisted Living (AAL) environment is to be intuitive and natural. This contributes to their acceptance especially for elderly users. Providing intelligent Speech User Interfaces (SUI) is one solution, since voice is a natural way of communication. However a proper interaction concept for SUI has not yet been defined. In this paper, we propose a concept based on two aspects: 1) Dialog initiation and control based on the shopping metaphor. This means that the system should be proactive but still leave the dialog control to the user. 2) Designing voice dialogs by reducing user input but maximizing system output. We also define canonical tasks and introduce a unified digital interaction interface to simplify dialog descriptions and interaction with heterogeneous devices in AAL.

Author Keywords
intelligent user interfaces, ambient assisted living, interaction concept, universal control, canonical task

ACM Classification Keywords
H5.2. Information interfaces and presentation: User interfaces. K4.2. Computers and Society: Social issues

INTRODUCTION
Ambient Intelligence (AmI) refers to electronic environments that are sensitive and responsive to the presence of people. In an AmI world, devices operate collectively using information and intelligence that is hidden in the network connecting the devices [10]. Current technological developments enable the integration of electronics into the environment, thus enabling the actors, i.e., people and objects, to interact within their environment in a seamless, trustworthy, and natural manner. The miniaturized devices as well as the user interface (UI) become invisible to the user.

One important AmI-based application is Ambient Assisted Living (AAL). The vision of AAL is to ease people's daily life by providing personal support in the immediate personal proximity through technology. For example, supporting elderly people in their daily lives allows them to stay longer within their own home to live longer independently and self-determined. This is motivated by the observed demographical and social changes; there are increasingly more caretakers than caregivers. One solution is to enable people, especially the elderly, to help themselves and to maintain autonomy by supportive technology. To this end, the AAL domain should provide both indoor and outdoor applications like those for emergency prediction, detection and prevention, assistance for cooking, drinking, shopping and medication, orientation services etc [2].

From the technical point of view, the AAL environment is made of heterogeneous devices that provide heterogeneous services. The services are dynamic in the sense that they can appear and disappear at any point during runtime. These can apply both to services that were envisioned at development time and new ones that become available when a new device is introduced in the environment, e.g., when bringing a wireless pager into the vicinity of a wireless network. Despite this dynamic, the user should interact with the system in an intuitive and as natural as possible manner [10]. Considering that most inhabitants in these environments are elderly or impaired, effortless interaction becomes a vital criterion to the user acceptance of AAL applications. Speech user interfaces as a natural and effortless way of interaction can play a significant part here [5].

Speech user interfaces are not novel to the field of human-computer interaction. They have gained momentum during the last years as the demand for well designed, usable and accepted speech dialog systems has pushed the development. In the industrial practice there are a few good examples on how usable speech user interface might look like [8]. In the beginning of this discipline, the systems merely understood single clearly spoken keywords, often only after having learned the specific acoustic pattern of the speaker. Today, the situation is much better with systems that do not need to be trained, that can process whole sentences, that can be interrupted, and that can handle background noise or unclear pronunciation [16]. Appling SUI in the AAL environment for elderly is therefore possible. This paper proposes a suitable interaction concept as a starting point.

The remaining part of this paper is structured as follows:. Some selected works related to AAL are briefly introduced in the next section. This is followed by a section presenting a unified digital interaction interface as mediator between the user and the AAL environment. Thereafter, canonical tasks in the AAL are defined. These are followed by the proposed interaction concept and a proposal for principles for designing voice messages. An interaction dialog as an implementation example then follows. The last section of this paper provides a conclusion and an introduction of future work.

RELATED WORK

Since the emergence of Ambient Intelligence (AmI), there have been many projects dealing with Ambient Assisted Living (AAL). The focus of most works has been to find the services that can be provided to the elderly user. For instance, Nehmer et al. provide examples of services like Emergency Treatment, Autonomy Enhancement or Comfort in [2].

Another issue has been to address context-awareness and multimodality in AmI environments. One of the known European projects towards this step is the Amigo ("Ambient Intelligence for the networked home environment") project [3]. In this project, an integrating middleware that supports mobile devices, consumer electronic devices, household appliances, as well as the PC domain was developed. By using the infrastructure provided, the system can monitor user activities, provide context-specific services as well as facilitate multimodal interaction. As the name implies, the main objective however was to facilitate home networking.

Some investigations have also been done on the application of speech for elderly users. For example, the experiment that was conducted in [3] on speech output revealed that male voices were preferred to female voices and natural voices were preferred to synthetic ones. This experiment also resulted in some recommendations such as keeping audio messages as short as possible, offering at most three selections when designing menus etc. In a work published by Hanson et al. [4] on speech inputs provides a complex picture on elderly users. For example, it was found that they speak slower than young ones and their speech input are hard to recognize. In [5] it was mentioned that correct recognition of speech input is vital to the usability of speech-based systems. "*Even a well-designed speech-based application will be undesirable if the user's speech is consistently misrecognized or the messages back to the user are not apprehended*." so the authors. These challenges must therefore be considered by a speech based application.

In addition, some speech interface design patterns for elderly users have been defined. Examples include designing confirmatory message (provide confirmation after user action), default input message (provide some default input values) and explanation message (tell the user what is going on) [6]. Though these were not made for the AAL context, they provide useful knowledge that can be relevant also for applications in AAL.

Assessing the works provided above, it can be concluded that a concept for speech dialogs for elderly users, especially for intelligent user interfaces in AAL is missing. The contribution of this paper is an interaction concept that describes how users perform their inputs and receive feedback from an AAL system that has volatile services integrated.

THE UNIFIED DIGITAL INTERACTION INTERFACE – A AAL MEDIATOR

Following the vision of AAL, the environment consists of many devices offering heterogeneous services and interaction concepts. Analogously, we talk of a user being immersed in digital environment. To smoothly and uniformly interact with such an environment, a unified digital interaction interface (UDI^2) mediating between the user and the user interfaces of several location specific interaction devices is required. This concept is different from that of unified user interfaces [17] in that after installation the UDI^2 is personalized to one user of the AAL environment since this remains the same over a long time. The UDI^2 becomes the sole location independent interface between the user and the AAL environment. In this way, it hides diversity of the system from the user and presents him an easily usable interface for all services. Figure 1 illustrates this concept, where the UDI^2 is symbolized by a thick ellipsis around the user, shielding away the interaction complexity. Note that the UDI^2 as a concept is not directly visible to the user but it can be implemented on a visible device like a PDA (see Figure 2).

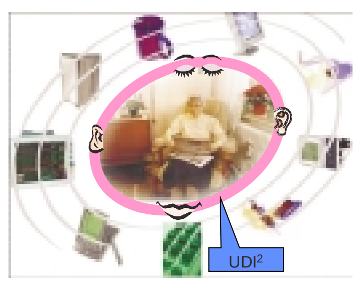

Figure 1. User, AAL Environment and UDI^2

A sample application scenario in AAL could be as follows. Mrs. White[1] bears her age well but has her Personal UDI^2

---

[1] This is a fiction name used on this paper. Even the scenario presented here has not been taken from a concrete interaction scene.

(PUDI²) implemented on a PDA. While being out for a stroll with her friend, PUDI² is connected to her insulin meter and her mobile phone with integrated GPS and routing. Using the location service of the mobile phone, PUDI² can tell the name of the street and where to go next to reach the café she was looking for. During the whole time Mrs. White's insulin level is monitored in the background by the insulin meter. Through PUDI², Mrs. White may ask for the current state of her medical condition at any time. PUDI² will remind her of her regular insulin intake by mentioning the upcoming medication time in advance. In case of decreased insulin level, e.g. if she has forgotten to inject her insulin, PUDI² will begin to blink and snore to catch the attention of Mrs. White. This is intended to give useful information in order to avoid a critical medical condition. If the condition further deteriorates PUDI² will begin to warn by making a peeping noise while trying to directly address Mrs. White, e.g. by speech. During the whole time but especially in the alert state, PUDI² will be ready to build up a telephone link with Mrs. White's physician to get help for her. If she falls into a diabetes coma, the insulin meter will start the alarm which PUDI² routes through the mobile phone to the emergency with all necessary data, e.g., the location and condition of Mrs. White. This is the worst-case scenario. Normally Mrs. White will go home after having had coffee with her friend. When she arrives at home, PUDI² recognizes the new environment, updates local data and gets the status of each service around. PUDI² reminds Mrs. White things that she is supposed to attend, for example, it will tell her that the washing machine that she started before leaving house is ready. In this new environment with the household being manageable through PUDI², Mrs. White has the opportunity to power on and off each electronic device, check the fire and smoke alarms, ask for external help in case of problems or alarm the fire-brigade or the police through a simple speech command.

To enable the operations necessary to make such a scenario come true, is not focus of this work but will be presented shortly. The high-level architecture of the AAL System with a PDA used as a UDI² is illustrated in Figure 2. The services can be accessed by the UDI² through the AAL Platform.

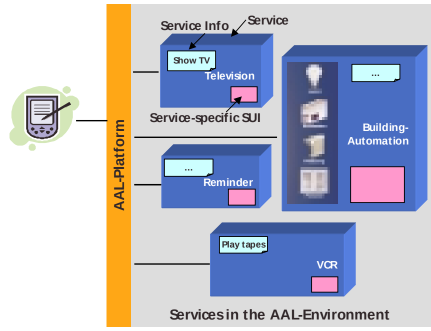

Figure 2. High-Level System Architecture

A NATURAL INTERACTION CONCEPT

In the previous section we have introduced UDI² as a mediator between the user and the complex AAL and provided examples of its application. In this section, we define canonical user tasks and an interaction concept for speech interaction with UDI² in the AAL environment. The tasks are derived from the scenario and examples provided above whereas the interaction concept is based on the metaphor of human to human interaction by speech.

Canonical User Task in AAL

As mentioned in the introductory part, one characteristic of the AAL environment is the presence of dynamic services. Services can arbitrarily come and go at runtime. The system is therefore never complete. Moreover available services could vary according to user preferences. For example while one user would like to have a TV service installed, another could prefer the internet service instead. Defining user interaction for SUI must therefore be service independent.

A solution we suggest here is to abstract from service details and come up with a set of canonical tasks. Canonical tasks are abstract tasks or task types that are defined to specify what kind or operations are supported by the system in the AAL environment to compose or augment functionality provided by the services. To develop these tasks, we can orientate ourselves at the sample scenario given in the previous section. We can also extract essential needs of interaction with a smart environment for the elderly and abstract them to more general types of interaction that represent a common basis for most AAL environments. In smart homes for elderly, the essential needs of the users that must be supported by the system are (according to [9]):

- Activate a central first-aid alarm
- Activate and de-activate of the monitoring of vital signs/functions
- Activate and de-activate the alarm and smoke detectors
- Give overview of electric devices
- Activate and de-activate every electric device
- Establish a direct contact to the doctor

- Remind of scheduled tasks

As it can be seen the tasks can be initiated by the user (e.g., "Request the current state of the medical condition") or the UDI^2 (e.g., "Remind of scheduled task"). Note that the UDI^2 is only a mediator but the fulfillment of the services is done by the respective components. For example, if the user requests information about some service, the UDI^2 transmits this request to the component providing this service (see Figure 2). The canonical tasks and their descriptions are shown in Table 1.

Name	Description	Example
User		
Explore services/ options	Go through existing services/options.	"Which programs are available today?"
Request information	Ask for information on some services.	"What does SMS mean?"
Create Program	Prepare a program for performing a certain task.	"Prepare VCR for recording"
Edit settings	Edit settings for UDI^2 or Program.	"Change voice type"
Suspend task	Leave a task incomplete.	"Suspend 'set up'"
Resume task	Continue with a suspended task.	"Resume 'set up'"
Initiate operation	Start an operation.	"Start recording"
Cancel operation	Stop a running task/operation.	"Stop/Cancel recording"
UDI^2		
Notify service appearance	Inform the user when a new service appears.	A signal for a new service.
Notify service disappearance	Inform the user when a existing service disappears.	A signal for a disappearing service.
Accept input	Inform the user that his input has been accepted.	A signal for accepting input.
Reject input	Inform the user that his input has been rejected.	A signal for rejecting input.
Send warning	Warn the user on nearing danger.	A signal for dropping pulse.
Send alarm	Alert the user on critical situations.	A signal for reduced water level.
Report status	Inform the user on the status of running operations.	A signal for the completion of dish washer
Provide information	Provide more details on services.	Provide requested details of a service

Table 1 Description of canonical tasks

Having presented the canonical tasks, the next step is to define the interaction between the user and the system in a way that is as intuitive and natural as possible.

The Shopping Metaphor – a natural interaction concept

One of the challenges of IUI is to provide intelligent support while leaving the user in control of the system [11]. In other words, an IUI should support the user without causing him to feel like being insulted or under estimated by it. Our solution approach is based on the following shopping example.

Observing people while shopping, one can realize that some customers get scared when the sales assistant speaks to them especially when they ask "*May I help you*?" once they enter the shop. Some will say "*Yes, thanks*", some will respond "*No, thanks*", others will say "*Well, let me first try my luck*" and still there are those who will remain silent, that is, they ignore the sales assistant. Of course, the same customer can give different answers at different times and to different assistants. Experienced sales assistants would try to first study the customer before they try to offer their services. They know the proper time and a proper way to approach different customers. In another case, some customers will directly ask for assistance before they spend time searching on their own, but this case is not of interest for us now. When an experienced sales assistant has successfully managed to get the customer accept to take his help/service, he will not stick on assisting the customer all the time until he leaves the shop. Instead, he will let the customer continue with his own decision making and at his own pace, without putting (much) pressure on him. In this way, the customer does not feel like being forced to buy something he would otherwise not have bought. In this case, other issues such as culture and language should be considered too, but these are beyond the scope of this paper.

Returning to our challenge, and substituting the experience sales assistant with the IUI, we get the picture of how this should behave. Since we have used a shopping scenario, we will call this the Shopping Metaphor. There are many scenarios similar to this, but we think that this is easy to remember and use. An interaction concept based on this metaphor can be summarized as follows:

1. IUI observes the user.
2. Depending on the user characteristics, the IUI kindly asks whether the user would like support.

3. Once the user has been supported, the IUI leaves him to continue on his own.

4. Should it realize that the user still needs support, step 2 above will be repeated.

Step 1 assumes that the IUI has the ability to observe the user and that user agrees to be observed.

Step 2 emphasizes that the user is treated gently when providing help. In this scenario the meaning of "kindly ask" is context[2] specific. For example, this can be an animated symbol blinking on the screen, a voice asking "*Can I help you*?", an acoustic signal etc. The important point here is to ask the user prior to providing the support. In this way, decision and control are left to the user.

In Step 3, the question of "how much support is enough before leaving the user" will have to be answered. When we go back to our shopping scenario, will it be enough to tell a customer who is looking for male garments that they can be found in the third floor? Or should the customer be accompanied and perhaps asked which concrete garments is he looking for? Of course, one could argue that it depends on the customer and the support required. Considering that it is the assistant who initiated the support dialog, it could be assumed that the customer's expectation will be higher than if the customer initiates a dialog. The support should therefore fulfill the customer's expectation (which will still need to be determined). Another related question is the quality of the provided support. Is it sufficient to simply tell the customer in the previous example only the number of the floor, or should the customer also be told how to get there? The least intrusive way is to try to anticipate the user's expectation, e.g. using heuristics or the history of the ongoing session.

In Step 4, it has been assumed that it can be realized the user still needs support. However, there is a need to clarify how this should be done. Simply asking the user "D*o you still need support?*" may not be adequately sufficient because, at some point, the user may have the feeling that he is not capable of supporting himself. This case can be treated like in step 1 with the extension that the two participants already know each other and there is a history of the interaction happened so far. For instance, this knowledge, especially the user's former decisions and reactions, can be used here to create a successful communication.

As the descriptions of the four steps imply, there are many questions that still need to be answered. We hope to find answers in the evaluation experiment which will be conducted soon.

[2] We use the world „context“ in its general meaning. It can be situation, culture, modality etc [7].

Designing voice messages for AAL

Although speech interaction is considered to be the most natural way of interaction, speech recognition and output is still far to be called "natural". In addition, the existing technological barriers and knowledge of properly designing voice message, especially for elderly, is still a challenge. There already exist some guidelines such as to keep the output speed slow or limit the number of menu options to at most three. Another guideline is to keep the audio messages as short as possible [3]. Unfortunately, this will result into long dialog that can be time consuming and boring to the user. It does also not clarify what the length of a "short" audio message should be. In this paper, we therefore propose two principles based on the human to human interaction:

- Reduce user input but maximize system output, and
- Limit the output message depending on the information density.

The first principle aims at reducing the user effort for information input. Keyword based applications are known examples of reduced user inputs. One disadvantage is that the user is forced to use a set of predefined keywords, which is unnatural. We believe that with better recognition and processing tools this limitation will no longer be required. The new and important point of this principle is to maximize the output. The user therefore becomes more information than his input. Let us demonstrate this in the following example. Assume that the user wants to record a TV program. The dialog might look like this:

User: "*Record a TV program*"

UDI^2: "*Which program shall be recorded?*"

User: "*The TV Doctor*"

UDI^2: "*Dou you mean the TV Doctor in Chanel One, today, at 19:30?*"

User: "*Yes*"

It can be seen that UDI^2 provides (almost) all necessary information related to the TV program to be recorded. There is no need of boring the user with questions like "*when*?", "*which Chanel*?" and "*at what time*". Of course it is assumed that the information is available in the system. After all we are talking about an "intelligent system". Note that there are still missing information. For example, whether the tape speed should be SP or LP. This brings us to the second principle of limiting the output message.

In designing graphical user interfaces, it is recommended to keep the density of the displayed information in a container for better perception between 25% and 60% [12, 13, 14]. Analogously, the density of audio information could be restricted to some density. The question is how this density can be determined. One possibility could be to define it as the total information content per unit time, i.e., information output divided by the time spent. The information content can in turn be determined as the total number of words contained in the message. For the message "*Dou you mean the TV Doctor in Chanel One, today, at 19:30?*" spoken by UDI² above, the information content could have been 12. Another possibility could be to define the information content by counting only terminological words (*Rhema*). In this case, the information content of the above example could have been 3, i.e., counting the words "*Chanel One*", "*today*" and "*19:30*". As an alternative to the information density, the number of new information in a message can be restricted. We can make use of the 7 (plus or minus 2) heuristic [15]. This could result into limiting the number of new words to the maximum of 9. Whether this number is acceptable for elderly still needs more investigation.

As it can be seen there are many questions to answer. We hope that we will be able to clear most of them in our evaluation experiment which is being developed.

APPLYING THE INTERACTION CONCEPT FOR UDI²

The interaction concept defined below will now be applied to a modern SUI for AAL. We will mainly focus on how UDI² should enter into a dialog with the user. According to Step 2 above, UDI² should attract the user's attention before giving out an acoustic message. This can be done by give a "proper/relevant" acoustic sign. This implies that there are different signals used for information coding depending on the context. For example a signal for indicating that a new service is available should differ from the one for indicating the removal of a service. Likewise, a signal for reminding the user on a pending task should differ from that for alerting him in some situation. When the user recognizes this signal and is ready for the message, he could then ask question like, "*What is new?*" This is an indicator for UDI² that the user is now ready for the message. It will then deliver a corresponding speech message. For example "*A new service for sending voice messages is now available. You are now able to send one to Mr. White. Shall I activate this service?*" Again the control is left to the user to decide what should be done. Note that the content of the message above provides all necessary information that might be required by the user. The user knows which service is available and what he can do with it. By providing more information, the interaction dialog is made short but yet without overload the user. It is the goal of UDI² to make the dialog as effective as possible by using the principles mentioned above.

CONCLUSION AND FUTURE WORK

In this paper, we have introduced and defined canonical tasks for the service independent description of user interaction in AAL. In this way, the it can be defined how the interaction should look like without knowing the concrete service. Based on the canonical tasks, an interaction concept using the shopping metaphor has been proposed that resembles human natural interaction and therefore simple to understand. The concept makes use of IUI but yet leaves the user in control of the action. In additional, this paper has proposed principles for designing voice messages for SUI in AAL. Such messages should provide maximum output for reduced user input, an important support for elderly users.

As future work, the ideas proposed in this paper will soon be evaluated in an experiment.

ACKNOWLEDGMENTS

The work presented in this paper was carried out in the BelAmI (Bilateral German-Hungarian Research Collaboration on Ambient Intelligence Systems) project, funded by the German Federal Ministry of Education and Research (BMBF), the Fraunhofer-Gesellschaft, and the Ministry for Science, Education, Research and Culture (MWWFK) of Rheinland-Pfalz. Special thanks to Tamás Kristóf Hámor who is now working on the design and implementation demo for evaluation purposes.

REFERENCES

1. Kurt Bittner, Ian Spence (2002). *Use Case Modeling*. Addison Wesley Professional, Amsterdam
2. Nehmer, J., Becker, M., Karshmer, A., and Lamm, R. 2006. Living assistance systems: an ambient intelligence approach. In *Proceeding of the 28th international Conference on Software Engineering* (Shanghai, China, May 20 - 28, 2006). ICSE '06. ACM Press, New York, NY, 43-50.
3. Amigo Project, Ambient intelligence for the networked home environment, http://www.amigo-project.org.
4. Hanson, V. L., J.T. Richards, and C.C. Lee. "Web Access for Older Adults: Voice Browsing?" Proceedings of HCI International - Universal Access in HCI. Beijing, China: Springer, July 22 - 27, 2007
5. Sara Basson, Peter G. Fairweather, Vicki L. Hanson: Speech recognition and alternative interfaces for older users. In interactions of ACM, Volume 14, Number 4 (2007), Pages 26-29
6. Zajicek, M. "Patterns for encapsulating speech interface design solutions for older adults." In *Proceedings of the 2003 Conference on Universal Usability* (Vancouver, British Columbia, Canada, November 10 - 11, 2003). CUU `03. New York: ACM Press, 2003, 54-60..
7. Dey, A.K. Abowd, G.D. Towards a Better Understanding of Context and Context-Awareness. CHI 2000 Workshop on the What, Who, Where, When, and How of Context-Awareness (2000)
8. Karger, R. "*Gewinner gekürt: VOICE Award 2006*" Pressemitteilung der Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Saarbrücken http://www.dfki.de/web/presse/pressemitteilungen_intern/2006/gewinner-gekurt-voice-award-2006
9. Ringbauer, B. & Heidmann, F. (2007). „Besonderheiten bei der User Interface Gestaltung für Senioren". In *Produktblatt Smart Home für Senioren*, Fraunhofer IAO, Stuttgart.

http://www.hci.iao.fraunhofer.de/de/projekte/smart_home_fur_senioren/index.html

10. Aarts, E., Encarnacao, J.: Into Ambient Intelligence. In: True Visions - The Emergence of Ambient Intelligence. pp. 1--16, Springer (2006)
11. European Commission: Software, Services and Complexity Research in the ISP Programme, FP VI 2006
12. Danchak, M.M.: CRT displays for power plants. In: Instruction Technology, 23 (10), S. 29-36, 1976
13. Kruschinski, V.: Layoutgestaltung grafischer Benutzungsoberflächen: Generierung aus OOA-Modellen. Heidelberg Berlin: Spektrum, Akademischer Verlag, 1999
14. Tullis, T.S.: Screen Design. In: Helander, M. (Ed.): Handbook of Human Computer Interaction. Amsterdam: Elsevier Science Verlag, 1998
15. Miller, G. A.: The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956. v. 63, S. 81-97
16. Steimel, B. (2006). "VOICE-Award 2006: Managed Services auf dem Vormarsch". In 'Marketing Börse' vom 01.09.2006 http://www.marketing-boerse.de/News/details/VOICE-Award-2006-Managed-Services-auf-dem-Vormarsch/3105
17. Stephanidis, C. The Concept of Unified User Interface. In *User Interfaces for All: Concepts, Methods, and Tools*. Laurence Erlbaum Associates, 2001.