Gestures Used by Intelligent Wheelchair Users

Dimitra Anastasiou[1] and Christoph Stahl[2]

[1] Computer Science/Languages Science, University of Bremen, Germany
anastasiou@uni-bremen.de
[2] German Research Center for Artificial Intelligence (DFKI), Bremen, Germany
christoph.stahl@dfki.de

Abstract. This paper is concerned with the modality of gestures in communication between an intelligent wheelchair and a human user. Gestures can enable and facilitate human-robot interaction (HRI) and go beyond familiar pointing gestures considering also context-related, subtle, implicit gestural and vocal instructions that can enable a service. Some findings of a user study related to gestures are presented in this paper; the study took place at the Bremen Ambient Assisted Living Lab, a $60m^2$ apartment suitable for the elderly and people with physical or cognitive impairments.

Keywords: assisted living, gestures, intelligent wheelchair, smart home.

1 Introduction

According to the International Society for Gesture Studies, gesture studies is a rich and old interdisciplinary field, broadly concerned with examining the use of the hands and other parts of the body for communicative purposes. Gesture researchers work in diverse academic and creative disciplines including anthropology, linguistics, psychology, history, neuroscience, communication, art history, performance studies, computer science, music, theater, and dance. Gesture is a necessary modality for people with speech disorders, but also in general, in situations where hand interaction is not feasible.

This paper is laid out as follows: in section 2 we present taxonomies of gestures based on a linguistic and computational point of view, and stress the importance of gestures in smart homes. Our research study and its findings are presented in section 3 and a conclusion and future prospects follow in section 4.

2 Gestures in Theory and Practice

Gesture is an already established scientific research field in verbal and particularly non-verbal communication, both from theoretical and practical points of view. Before the 1980s gesture was part of non-verbal communication research; only after the 1980s, gesture was closely tied with speech in creating meaning.

McNeill (1992 [1]), based on Kendon (1982) [2], laid the philological foundations about gestures and mind, and classified gestures into gesticulation, pantomime,

K. Miesenberger et al. (Eds.): ICCHP 2012, Part II, LNCS 7383, pp. 392–398, 2012.

emblem, and sign language. Various dimensions were assigned to these types of gestures, namely i) degree of speech accompaniment (reducing from gesticulation to sign language), ii) degree of linguistic properties (ascending from gesticulation to sign language), iii) conventionality (also ascending), and iv) semiotic differences (gesticulation is global/synthetic, pantomime segmented/synthetic, emblem global/ analytic, and sign language segmented/analytic).

Within the context of Human-Computer Interaction (HCI), there is a taxonomy of gestures developed by Quek (1994) [3]; based on this taxonomy, meaningful gestures are differentiated from unintentional movements. Meaningful gestures are classified into communicative and manipulative gestures. The former are used to act on objects in an environment and the latter have an inherent communicational purpose. Manipulative gestures can occur both on the desktop in a 2-D interaction using a direct manipulation device (mouse, stylus), as a 3-D interaction involving empty handed movements to mimic manipulations of physical objects (virtual reality interfaces), or by manipulating actual physical objects that map onto a virtual object in tangible interfaces. Wexelblatt (1998) [4] provided an overview of the primary classifications referred to in some computing literature too.

2.1 Gestures in Smart Homes

Ambient Assisted Living (AAL) has been the research topic of many scholars of the last decade. AAL is a research domain supported by national programmes (such as the German BMBF and the European Ambient Assisted Living Joint Programme) which promotes intelligent assistant systems for a better, healthier, and safer life in the preferred living environments through the use of Information and Communication Technologies (ICT). AAL technologies and applications are used in domotics, as the motivation for AAL research is to improve the lifestyle of the seniors in their domestic environments. The homes where such AAL technologies are applied to are called smart homes.

In smart homes multimodal applications are necessary to compensate specific limitations of physically challenged people. For example, people with motor disabilities would prefer or need speech interaction, while people with speech impairments prefer gestural interaction; more information on why gestures are needed in AAL can be found in Anastasiou (2011) [5]. Nazemi et al. (2011) [6] stated that gestural interaction is more natural and simpler for seniors, as they often have problems with precise movements to open applications through clicking on programme symbols at interfaces. Besides, it is difficult for them to use the common TV remote control, because the buttons as well as the text written on them are too small. Moreover, in situations where the hands are employed, such as cooking in the kitchen or talking to the phone, where verbal communication is impossible or constricted, gestural interaction is helpful.

As for intelligent wheelchairs/personal assistants in AAL, traditional electric-powered wheelchairs are normally controlled by users via joysticks, which cannot satisfy the needs of elderly and disabled users who have restricted limb movements caused by some diseases, such as Parkinson and quadriplegics (Jia et al., 2007 [7]). Generally speaking, unlike existing techniques, gesture-based interaction as a mode of explicit interaction is more natural and appealing to people while accessing various services (Chen et al., 2010 [8]).

As far as gesture recognition techniques applied in AAL is concerned, Jaimes & Sebe (2007) [9] stated that recognizing gestures and integrating them to access ambient services has been under-researched due to the lack of accuracy, limited set of gestures, extensive learning efforts, overall robustness of the particular gesture recognition techniques, and their special setup requirements and operating constraints. Many initiatives have been undertaken in AAL by applying three dimensional acceleration sensor information of the *WiiMote*, e.g. Nazemi et al. (2011) [6] and Neßelrath et al. (2011) [10].

3 User Study

A pilot user study[1] in BAALL took place in November-December 2011 and included a real-life everyday scenario of a human user using a wheelchair to navigate in their environment by means of speech and gesture. Both HRI and home device control are available in BAALL (see Krieg-Brückner et al. 2010 [11]) at the German Research Center for Artificial Intelligence. BAALL is an apartment suitable for the elderly and people with disabilities. It has an area of 60m^2 and contains all standard living areas, i.e. kitchen, bathroom, bedroom, and living room. In BAALL the autonomous wheelchair/robot *Rolland*, offers mobility assistance being equipped with two laser range-sensors, wheel encoders, and an onboard computer; *Rolland* has a spoken dialogue interface that allows the user to choose destinations and control devices in BAALL; control with smart phone or tablet PC is also possible. The intelligent wheelchair *Rolland* and a participant pointing to a landmark are depicted in Figure 1.

Fig. 1. Rolland and a participant pointing to the kitchen in BAALL

The goal of the study was to observe whether people would gesture and how, and what they would say, if they used a wheelchair in their smart home to carry out daily activities. One of the hypotheses was whether the participants would gesture more in case their addressees can see them. Indeed speakers gesture more in this situation than

[1] Acknowledgments to Daniel Vale, Bernd Gersdorf, Thora Tenbrink, and Carsten Gendorf.

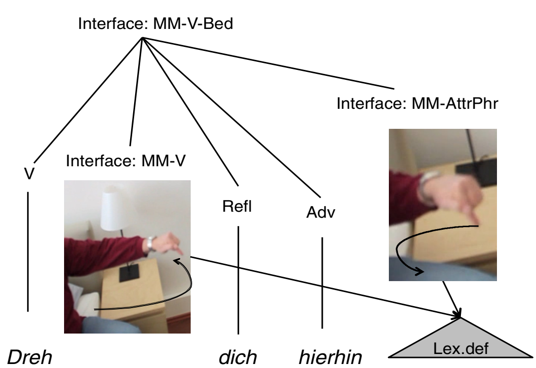
when their addresses cannot see them (Cohen, 1977 [12]; Alibali & Don, 2001 [13]). Moreover, Rimé & Schiaratura (1991) [14] said that gestures' type and density change in relation to the referent, the recipient, and the communication mean. 20 German student participants took part in the study (mean age 26) and were asked to act as if they were dependent on the wheelchair, so that the scenario is close to reality. The participants had to perform the following activities in this order:

- Rinse your mouth with water.
- Take something to eat or drink.
- Wash your hands.
- Take a book.
- Read the book on the sofa while Rolland recharges.
- Open the door when someone knocks.

The user study was Wizard-of-Oz (WoZ) controlled, i.e. an experimenter A was in an office observing by live audio and video streaming what was going on in BAALL in order to remotely set *Rolland*'s navigational goal. An experimenter B was inside BAALL giving instructions to the participants and following them during the task execution. There were three cameras in total which recorded the subjects' activities: two cameras placed in BAALL and one placed on *Rolland*'s back, so that *Rolland* could see and 'supposedly' recognize their gestures. The verbal instructions included introduction of the lab and *Rolland,* and instruction of the tasks that the participants should perform. The written instructions included the 6 tasks presented above and were also put as a note on *Rolland*'s armrest; they were not given as handouts, so that participants have their hands free to gesture. In the verbal and written instructions we did not refer to the landmarks, i.e. 'living room', 'bathroom', etc., but to the activities instead, e.g. "Read a book on the sofa" instead of "Go to the living room". This decision was made because one of the goals in this study, apart from observing gestural frequency and gesture types, was to collect empirical data of natural dialogue in HRI. There were tasks where the user is sitting on *Rolland*, but also when *Rolland* drives autonomously without the user (part of task 5), as differences in gesture may change based on the recipient. In the end of each session, a retrospective protocol approach was followed (Dorst & Dijkhuis, 1995) [15]; the participants were asked to go through the tasks that they just performed and say what they were thinking. They were also asked to recommend future improvements of the HRI. An example answer from the retrospective protocol was that one female participant expected *Rolland* to have a female voice (see Crowell et al. 2009 [16]). Most of the participants said that *Rolland* "parked" too far and a person with disabilities would not have been able to reach it.

As for the findings of the study in relation to gestures, in 7 sessions out of total (35%) participants employed at least one gesture during a session (6 sessions with deictic gestures, 1 with iconic). In 2 of the 7 sessions participants gestured more than once, while in the remaining 5 sessions, people gestured once. The deictic gestures were pointing at a place where *Rolland* should go to, as *Rolland* was too far from the landmark that the participants wanted to go. The iconic gesture by one participant was 'hand rubbing' under the tap to represent a state of washing her hands. In general, the study has shown that participants gestured mostly when something happened out of order, e.g. *Rolland* drove to a wrong destination point or stopped too far from the participant or was close to hit a wall or door. When everything went well, i.e. *Rolland* drove the participant where he/she wanted to, participants did not gesture. An exception was one female participant who gestured during all the tasks. This situation can be explained by personal influences, e.g. the user's personality (see Rehm et al., 2008) [17]. From the study also the attitude and expectations of the participants against the robot have been evaluated. The style, the volume of the utterance, the waiting time for *Rolland* to react as well as the content of the lexical content of the utterance itself have shown that humans' perception of robots vary and can significantly change during a study. For example, a participant waited firstly for 9 seconds for *Rolland* to react, then for 3 seconds, until finally she uttered the context-sensitive spatial instruction "come here": "Rolland, <break 9 sec> Rolland, <break 3 sec> komm her (*Rolland, Rolland, come here*)". It is worth noting that one female participant characteristically expected *Rolland* to have female voice. A video recording showing some of the scenes where participants employed gestures is available[2].

In order to make an analysis of gestures in the multimodal grammar, we follow the model by Hahn & Rieser (2010) [18]. Figure 2 represents an example of the speech-gesture alignment of the phrase *Dreh dich hierhin (Turn over here)* in the dialogue part. The arrows outside the pictures pointing towards the lexicon definition indicate that gesture content operates on lexical content.



***V*: Verb**
***Adv*: Adverb**
***AttrPhr*: Attribute phrase**
***Lex.def*: Lexicon definition**
***MM*: multi-modal**
***Refl:* reflexive pronoun**

Fig. 2. Interface of speech-gesture in HRI (adapted from Hahn & Rieser, 2010)

4 Conclusion and Future Prospects

In this paper we introduced a philological and computational taxonomy of gestures. We are interested in contact-free, touch less gestures, i.e. *gesticulation* according to

2 The video is available at `http://ai.cs.uni-sb.de/~stahl/d-anastasiou/DiaSpace/Resources/`

the philological taxonomy and *meaningful gestures* according to the computational taxonomy. Then a user study with the goal to collect empirical speech-gesture data was presented. The findings have shown that only in situations where *Rolland* parked too far and generally when something went out of order, participants gestured. In the future we plan to have participants coming from different countries to see whether cultural differences in pointing gestures in AAL domain exist (see Kita [19]). In addition, more user studies with specific constraints, e.g. participants are prohibited to speak, and/or ambiguous situations ("Bring me to the sofa", having many sofas in BAALL) etc. are planned for the near future in order to focus more on gesture generation. Gesture recognition will follow after we have established a theoretical framework about a set of natural and intuitive gestures used in a smart home.

Acknowledgments. We gratefully acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center SFB/TR 8 Spatial Cognition.

References

1. McNeill, D.: Hand and Mind: What Gestures reveal about Thought. University of Chicago Press (1992)
2. Kendon, A.: The study of gesture: some observations on its history. Recherches Semiotique/Semiotic Inquiry 2(1), 25–62 (1982)
3. Quek, F.: Toward a vision-based hand gesture interface. In: Proceedings of the Virtual Reality System Technology Conference, pp. 17–29 (1994)
4. Wexelblat, A.: Research Challenges in Gesture: Open Issues and Unsolved Problems. In: Wachsmuth, I., Fröhlich, M. (eds.) GW 1997. LNCS (LNAI), vol. 1371, pp. 1–11. Springer, Heidelberg (1998)
5. Anastasiou, D.: Gestures in assisted living environments. In: Proceedings of the 9th International Gesture Workshop (2011)
6. Nazemi, K., Burkhardt, D., Stab, C., Breyer, M., Wichert, R., Fellner, D.W.: Natural gesture interaction with accelerometer-based devices in ambient assisted environments. In: 4. AAL-Kongress, pp. 75–84 (2011)
7. Jia, P., Hu, H., Lu, T., Yuan, K.: Head gesture recognition for hands-free control of an intelligent wheelchair. Industrial Robot: An International Journal 34(1), 60–68 (2007)
8. Chen, K.Y., Chien, C.C., Chang, W.L., Teng, J.T.: An integrated color and hand gesture recognition approach for an autonomous mobile robot. In: Proceedings of the 3rd International Congress on Image and Signal Processing (2010)
9. Jaimes, A., Sebe, N.: Multimodal human–computer interaction: A survey. Computer Vision and Image Understanding 108(1-2), 116–134 (2007)
10. Neßelrath, R., Lu, C., Schulz, C.H., Frey, J., Alexandersson, J.: A Gesture Based System for Context – Sensitive Interaction with Smart Homes. In: Wichert, R., Eberhardt, B. (eds.) Ambient Assisted Living, vol. 63, pp. 209–219. Springer, Heidelberg (2011)
11. Krieg-Brückner, B., Röfer, T., Shi, H., Gersdorf, B.: Mobility Assistance in the Bremen Ambient Assisted Living Lab. GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry 23(2), 121–130 (2010)
12. Cohen, A.: The communicative functions of hand illustrators. Journal of Communication 27, 54–63 (1977)

13. Alibali, M.W., Don, L.S.: Children's gestures are meant to be seen. Gesture 1, 113–127 (2001)
14. Rimé, B., Schiaratura, L.: Gesture and speech. Fundamentals of nonverbal behavior. In: Studies in Emotion & Social Interaction, pp. 239–281 (1991)
15. Dorst, K., Dijkhuis, J.: Comparing paradigms for describing design activity. Design Studies 16, 261–274 (1995)
16. Crowell, C.R., Scheutz, M., Schermerhorn, P., Villano, M.: Gendered voice and robot entities: perceptions and reactions of male and female subjects. In: Proceedings of the 2009 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, IROS (2009)
17. Rehm, M., Bee, N., André, E.: Wave like an Egyptian: accelerometer based gesture recognition for culture specific interactions. In: Proceedings of the 2nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction, vol. 1, pp. 13–22 (2008)
18. Hahn, F., Rieser, H.: Explaining Speech Gesture Alignment in MM Dialogue Using Gesture Typology. In: Proceedings of the 11th Annual SIGdial Meeting on Discourse and Dialogue (2010)
19. Kita, S.: Cross-cultural variation of speech-accompanying gesture: A review. Language and Cognitive Processes 24(2), 145–167 (2009)