HAI-Workshop proposal

Title: The body of embodiment

Understanding Embodiment in human-agent interaction – an interdisciplinary approach

Overview of the workshop

Robots and virtual agents are progressively becoming more ubiquitous in our society. Although both types of agents may have an embodiment either physical or simulated (Pfeifer and Scheier,1999), a crucial difference exists, in terms of the possibility to share the same space of the users. There is currently an active discussion on the potential impact of physical presence on interaction and agent perception (e.g., Li 2015, Moosaei,et al. 2017, Ferrari & Eyssel 2016), both for the investigation of social cognition (e.g. Sciutti et al. 2015) and in applications of agents in the social domain (e.g. Fisher, Lohan, Foth 2012). For instance, is a physical agent better than a virtual agent in the context of teaching children or acting as receptionist?

The workshop "**The body of embodiment**" will focus on the role of an agent's embodiment and physical presence and their effects on social human-agent interactions.

The following core research questions will be explored with a multidisciplinary audience:

- Which is the role of **embodiment** as a constituent of social agents?
- Which is the role of **physical presence** as a constituent of social agents?
- Which impact embodiment and physical presence have in the **social perception** of robots and virtual agents?

These three key questions are at the core of ongoing discussions in the field of HRI, HCI and HAI and this workshop aims to shed light on the development of a common ground and mutual understanding among experts in the field of embodiment research, coming from robotics, neuroscience, psychology, computer science and AI. We believe that this event will strengthen the communication among the different research areas in the field of interaction between humans and novel technologies. Moreover, this workshop seeks to introduce junior researchers who have just entered the field to this highly relevant and complex concept.

Time		Speaker	Title
08:30-10:25			
	08:30 -09:00	Alessandra, Friederike, Katrin	Opening presentation
	09:00-10:00	Stefan Kopp	From embodied communication to social presence in dynamic human-agent interaction
	10:00-10:25	Paul Hemeren	Embodiment in social hand gestures from kinematic information in biological motion
Coffee Break			
11:00-12:25			
	11:00-12:00	Tom Ziemke	Brains, Bodies & Buddies
	12:00-12:25	Tomoko Koda	Shyness Level and Sensitivity to Gaze from Agents and Robots
Lunch			
13:45-15:10			
	13:45-14:45	Massimiliano Cappuccio	Embodied Social Robotics and its Theoretical Foundation
	14:45-15:10	Ingo Keller	The Effect of Speech Location on the Likability of a Robotic System
Coffee Break			
15:30-18:00			
	15:30-16:00	Linda Hirsch	Establishing the role of a social robot as memory trainer
	16:00-16:30		Panel discussion
	16:30-17:00	Alessandra, Friederike, Katrin	Closing Session

Invited Speakers

Prof. Stefan Kopp
Bielefeld University

Titel: From embodied communication to social presence in dynamic human-agent interaction

Abstract
The effects and role of the embodiment of virtual and physical agents have been discussed in many studies. However, accumulating evidence suggests that the embodiment of an artificial agent is less important than the presence it is able to create. In this talk, I will discuss what it takes from an artificial agent to create and keep up a social presence with its human user(s). I will argue that social presence, while initially being induced through the embodiment and appearance of the agent, is a volatile perception formed continuously by the user during dynamic human-agent interaction. Importantly, it does not only depend on the appropriateness or lifelikeness of single behavioral responses, but is eventually contingent upon the degree to which an artificial agent can engage in meaningful embodied and cognitive interpersonal communication and coordination. I will present work to model such abilities in virtual and physical agents.

Prof. Tom Ziemke
University of Skövde

Title: Brains, Bodies & Buddies

Abstract
Rule 4 of the "Principles of Robotics" formulated by a number of researchers in the UK states that robots - but this presumably applies to virtual agents as well - are "manufactured artefacts" whose "machine nature should be transparent", and that "the illusion of emotions and intent should not be used to exploit vulnerable users". While at a first glance Rule 4 makes very much sense (at least to this speaker), upon closer inspection it seems to depend on at least two assumptions that are far from uncontroversial in the cognitive sciences today. Firstly, the assumption that humans - and presumably other animals as well - are not machines. Secondly, the assumption that humans - and possibly other animals as well - have real emotions and intent, whereas machines are at best capable of creating the illusion thereof. The concept of embodiment is of course crucial to evaluating these assumptions, but unfortunately it is far from well-defined. The talk will try to clarify how different notions of embodiment and its role in cognition and emotion, as well as work in social neuroscience, relate to current research on human social interaction with different types of autonomous technologies, such as robots, virtual agents, and automated vehicles.

Prof. Massimiliano Cappuccio
United Arab Emirates University (UAEU)

Title: Embodied Social Robotics and its Theoretical Foundation

Abstract
We often hear that robots are embodied computational systems. It is assumed that any peripheral or actuator attached to a central processing unit constitutes ‘the body’ controlled by the robot. Such functionalist approach to robotics cannot do justice to the notion of embodiment. This approach risks confounding the real meaning of cognition by underestimating the role that the body plays in it. To correctly understand the legitimately cognitive role played by the body, I will introduce and briefly discuss the ecological, enactive, and phenomenological principles of embodied

cognition theory. This theory, often contrasted to the instructionist, representationist and cognitivistic approach to cognition, has provided one of the richest scientific characterizations of the body in both biological and artificial cognitive systems. According to embodied cognitive science, the body is the medium that discovers the worldly surroundings of a cognitive agent in terms of opportunities of perception and action. The body is the background of experience, affect, and purposefulness against which these surroundings can be mapped in terms of meaningful adaptive behaviors. The role played by the body is not only instrumental to cognition, because it is the body itself that assigns an adaptive significance to any act of perception and action, scaffolding at the same time higher and more abstract forms of cognitive interaction, decision, and problem-solving. This means the body is not only the physical implementer of computational functions, but one of the very preconditions of cognition itself. For a long time, these ideas have been explored by the situated, embodied/enactive, and "soft" approaches to robotics, which over the years have provided some of the key theoretical contributions to the development of the embodied cognitive science paradigm. Today, these ideas are tested in the arena of applied social robotics: social cognition is inherently embodied because the relationship of empathic attunement between social agents is necessarily mediated and scaffolded by the details of their bodily constitution and organization. If the cognition social agents were not embodied in such a rich sense, only inferential off-line forms of social attunement would be possible among them. A cognitive agent can recognize another cognitive agent and correctly interpret its behavior only becoming sensitive to the goal-oriented opportunities of action enabled by its body, which in turn can be detected only in analogy with the body of the cognitive agent itself, and its first-hand interactive and manipulative capabilities.

Massimiliano L. Cappuccio is associate professor of Cognitive Science at UAE University, where he directs the Interdisciplinary Cognitive Science Lab. His work addresses theoretical issues in embodied cognition and social cognition combining analytic, phenomenological, and empirical perspectives. He is the principal investigator of two UAEU/NRF-sponsored research projects that focus on performance under pressure and human-robot interaction, respectively. With Mohamad Eid and Friederike Eyssel he organizes and chairs the yearly Joint UAE Symposium on Social Robotics (JSSR). He is currently editing the MIT Press Handbook of Embodied Cognition and Sport Psychology.

Embodiment in social hand gestures from kinematic information in biological motion Human judgment data and machine learning

Paul Hemeren & Serge Thill

University of Skövde School of Informatics Skövde, Sweden paul.hemeren@his.se

Gauss Lee[1] & Peter Veto[2]

[1]ETraveli AB, Gothenburg, Sweden
[2] Department of Physics, Philipps-Universität Marburg, Germany Gauss.Lee@etraveli.com

Abstract—**Intention recognition by humans and three machine learning techniques (k-Nearest Neighbor, Locality Sensitivity Hashing-Forest and Dirichlet Process Gaussian Mixture Model) is assessed by using human classification data as a reference for evaluating the classification performance of machine learning techniques for thirty hand/arm gestures. The gestures are classified according to the degree of grasping and judgments of their social quality. The results indicate that the machine learning techniques provide a similar classification of the actions according to grasping kinematics and social quality. This supports previous findings that demonstrate a kinematic basis for perceiving intention in humans.**

Keywords—gesture recognition; biological motion; social actions; grasping; machine learning; kinematics

Shyness Level and Sensitivity to Gaze from Agents and Robots

Tomoko Koda
Osaka Institute of Technology
1-79-1 Kitayama, Hirakata, Osaka, Japan
tomoko.koda@oit.ac.jp

Takuto Ishioh
Osaka Institute of Technology
1-79-1 Kitayama, Hirakata, Osaka, Japan
m1m16a03@st.oit.ac.jp

Taku Hirano
Osaka Institute of Technology
1-79-1 Kitayama, Hirakata,
Osaka, Japan

ABSTRACT

This paper introduces our experiments on individual's shyness level, gaze sensitivity, and agent's embodiments. We report the results of our first experiment on how shy individuals perceived different amount of gaze from a virtual agent and how their perception of the gaze affected their perceived friendliness of the agent. The results indicate shy individuals are sensitive to even a very low amounts of gaze from the agent. However, contrary to our expectations, as the amounts of gaze from the agent increases, shy individuals had more favorable impression toward the agent, and they did not perceive the adequate amount of gaze as most comfortable. We then describe an outline of our second experiment with a humanoid robot that exhibits the same gaze patterns as the virtual agent in order to compare participants' perception and impressions of gazes from different embodiments.

CCS CONCEPT

• **Human-centered interaction** → Human computer interaction (HCI) → Empirical studies in HCI

KEYWORDS

HAI, HRI, IVA, gaze, non-verbal behavior, embodiment, personality, shyness, evaluation

1 INTRODUCTION

Gaze plays an important role in our social interactions such as controlling the flow of a conversation, indicating interest and intentions, and improving listener's attention and comprehension [1, 2]. As in humans, virtual agent's gaze behavior is also important to provide natural interaction. Previous research on modeling gaze behavior of virtual agents were conducted to make appropriate turn management [3], to express social dominance by gaze [4], all of which report modeling realistic human gaze behavior to an agent resulted in more natural and smooth interaction.

However, being gazed at can lead to discomfort from feeling observed, especially for shy individuals. Shyness is defined as "discomfort and inhibition in the presences of others, where these reactions derive directly from the social nature of the situation" [5]. Shy individuals tend to avert gaze and engage in more self-manipulations [6]. Thus, shy people might not prefer to interact with a virtual agent that exhibit a social, realistic human gaze behavior that facilitates smooth interaction.

Our first experiment [7] aimed to investigate adequate gaze behavior of a virtual agent for shy individuals to interact comfortably, and seek for answers for the following hypotheses: 1) Shy individuals are more sensitive to gaze from a virtual agent than those are not shy. 2) Shy individuals prefer lower amounts of gaze from the virtual agent, thus they perceive more friendliness from an agent that does not gaze at them.

On the other hand, HAI researchers have reported people's different behaviors and attitudes to a collocated robots, remote robots, and virtual agents that perform the same task [8, 9]. Humanoid robots used in HAI have fixed pupils in general, and gaze controls are mainly done by tilting or rotating their head. Thus, we need to verify the results from the first experiment using a humanoid robot with controllable pupils. This paper reports the results of our first experiment with a virtual agent and describes our second experiment procedure with a humanoid robot.

2 GAZE FROM A VIRTUAL AGENT

2.1 Experimental Procedure

We designed a conversational virtual agent with gaze behaviors based on [10] that proposed a Japanese gaze model controlled by a probabilistic state transition (shown in Fig. 1 right). We designed four gaze conditions. Firstly, the agent gazes toward a participant all the time during a conversation (Full gaze condition); secondly, the agent gazes toward the participant 67% of the time, which is defined as an adequate gaze amount to facilitate smooth interaction in [10]; thirdly, the agent gazes toward the user 33% of the time (low gaze condition); and lastly, the agent gaze away from the user all the time (no gaze condition). The agent's gaze-at and gaze-aversion states are shown in Fig. 1.

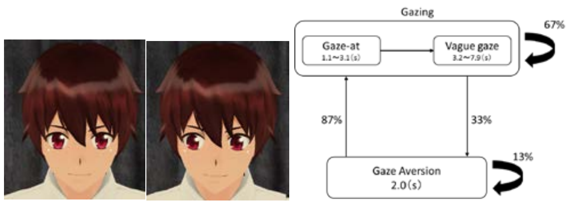

Figure 1: Agent's Gaze-at state (left), Gaze-aversion state (middle), and Japanese gaze state transition model (right)

19 university students participated in the Woz experiment and had pseudo conversations with the all four gaze conditions. They answered the Shyness scale questionnaire [11] before the experiment. We divided the participants into two groups based on their shyness level scores. 12 participants were categorized as high shyness group (HS), 7 participants as low shyness group (LS). The post-experiment questionnaire was conducted on

perceived gaze amount from the agent, and perceived friendliness of the agent after each conversational session.

2.2 Results and Discussion

The results of 2-way ANOVA repeated measures showed a significant main effect of gaze condition in "perceived amount of gaze from the agent" (F=32.95, p<0.01). The more gaze the agent gives toward the participants, the higher they felt the agent was looking at them. HS was more sensitive to change of gaze amount between no gaze and low gaze condition (score=1.83, 3.42; F=11.46, p<0.01), while LS was more sensitive to the one between low gaze and adequate gaze condition (score=2.00, 4.14; F=10.81, p<0.01). Thus, hypothesis 1 is supported.

In terms of perceived friendliness toward the agent shown in Fig. 2, HS liked the agent less than LS in general. LS rated the adequate gaze condition as most friendly (score=4.73, p<0.05), while HS did not perceive the difference in friendliness between adequate and full gaze, although they are aware of the differences in the amount of gaze between the two conditions. The results did not support the hypothesis 2.

These results indicate the answer to the questions in the following; 1) HS are sensitive to gaze even in the low gaze condition, where the agent gaze at the participant only 33% of the interaction duration. 2) However, HS perceived lowest friendliness toward the agent that does not gaze at them at all. On the contrary, they perceived the highest friendliness toward the agent that give them full gaze, while the low shy-ness group rated the friendliness of the agent highest when its amount of gaze was adequate. This suggests that HS were sensitive to little amounts of gaze but not sensitive to / aware of "adequate level of gaze" (66 % of the interaction duration), which is recognized and attributed to agent's friendliness by the low shyness group.

One of the reasons the hypothesis 2 was not supported is that the interaction partner was not a collocated physical robot, nor human, but a virtual agent. Thus, we need to conduct a sequel experiment with a humanoid robot with controllable pupils.

3 GAZE FROM A HUMANOID ROBOT AND FUTURE DIRECTIONS

The humanoid robot used in the second experiment is CommU, developed by Vstone, which has 8 DOFs in its eyes, head, and 6 DOFs in its body. CommU's eye movements and timings are fully controlled using the same gaze transition models and gaze conditions as in the virtual agent in the first experiment. The only difference between the first and second experiment is the embodiment of the agent. Our question is whether shy individual's perception of gaze and friendliness from the robot are similar to the ones from the virtual agent. Our hypotheses are: 1) Shy individuals' sensitivity to gaze is affected by the embodiment of the agent, 2) Shy individuals perceive more friendliness toward the virtual agent than the robot and humans, 3) Not-shy individuals perceive the equal friendliness from the virtual agent, the robot, and humans. The experiment is currently conducted.

Further study should compare perception of gaze with wider variety of appearances of agents, with different realism. Also we should consider other personality measurements, i.e., introvert/extrovert in addition to shyness. Moreover, quantitative analysis of eye tracking data of participants' gaze, especially whether they look at the agent's face or eyes is needed. We believe this research would lead to investigate comfortable gaze behaviors of agents for shy/introvert individuals, and such agents could be applicable to train adequate gaze behavior for them.

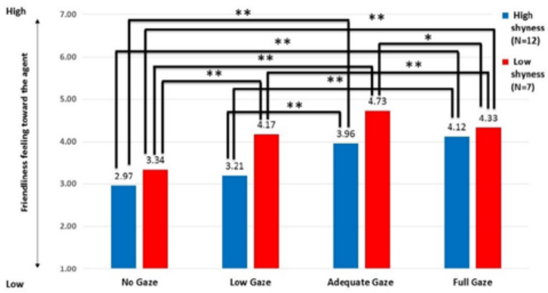

Figure 2: Perceived Friendliness from the Agent

Figure 3: Robot's Gaze state (left) and Gaze-away state (right)

ACKNOWLEDGMENTS

This work was partially supported by JSPS KAKENHI JP26330236, and JP17K00287.

REFERENCES

[1] Argyle, M., Cook, M. 1976. Gaze and mutual gaze. Cambridge University Press Cambridge.

[2] Bayliss, A.. et al. 2006. Gaze cuing and affective judgments of objects: I like what you look at. Psychonomic bulletin & review 13, 6, l06l-1066.

[3] Pelachaud, C., and Bilvi, M. 2003. Modelling gaze behavior for conversational agents. ln Proc. of IVA2003, Springer, 93-100.

[4] Bee, N., et al. 2010. Bossy or Wimpy: Expressing Social Dominance by Combining Gaze and Linguistic Behaviors. In Proc. of IVA2010, 265-271.

[5] Jones, Warren H., et al. 1982. The social reticence scale: An objective instrument to measure shyness. Journal of personality assessment 46, 629-631.

[6] Cheek, J. M., & Buss, A. H. 1981. Shyness and sociability. Journal of Personality and Social Psychology, 41, 330-339.

[7] Koda, T. et al. 2016. Shyness Level and Sensitivity to Gaze from Agents. in Proc. IVA 2016, 359-363.

[8] Powers, A., Kiesler, S., Fussell, S., and Torrey, C. 2007. Comparing a computer agent with a humanoid robot. HRI '07, 145-152.

[9] Hoffmann, L. and KräMer, N.C. 2013. Investigating the effects of physical and virtual embodiment in task-oriented and conversational contexts. Int. J. Hum.-Comput. Stud. 71, 7-8, 763-774.

[10] Ishii, R, et al. 2008. Avatar's Gaze Control to Facilitate Conversation in Virtual-Space Multi-User Voice Chat System, in J. of HI, 10 (1), 87-94.

[11] Aikawa, A. 1991. A study on the reliability and validity of a scale to measure shyness as a trait. The Japanese Journal of Psychology, 62(3), 149-155.

The Effect of Speech Location on the Likability of a Robotic System

Ingo Keller[1], Murat Gunana[1], and Katrin Lohan[1]

***Abstract*— In this paper, we will present a Human Robot Interaction (HRI) pilot study, that investigates the question, which impact the location of speech production has on a simple gesture interpretation game with the iCub robot. Participants where asked to communicate about different objects on a table by using a pointing gesture. As a feedback the robot would produce sentences confirming the object selected by the human. Using the Godspeed questionnaire we found that the robot has a higher likability, the animation is preferred and there are differences in the behaviour of the participants between the two speech conditions.**

I. Introduction

Robot embodiments and especially the humanoid type are an important factor not only to create robots that are able to deal with an environment build for humans but also to create robots that are more natural to interact with [4]. One of these ways is communicating through voice. For example, Hastie et. al. [7] are discussing the interaction between voice and appearance in the embodiment of a Robot Tutor in a pedagogical setting. They could show that the type of voice used has an impact on the perceived role of the robot in such a setting. It is widely recognized that audition dominates time perception, while vision dominates space perception [5]. They are suggesting that the visual dominance for space and the auditory dominance for time could reflect a cross-sensory comparison of vision in the spatial visuo-audio task and a cross-sensory comparison of audition in the temporal visuo-audio task. In our paper, we are investigating the aspect of speech produced by an iCub robot. We chose a common task that involves feedback from the robot and for this we are manipulating the location of the speech production to investigate the impact on the acceptance of the robot as an embodied agent.

Since humanoid robots are more commonly treated as social interaction partners/companions, we included an analysis of proxemics in our pilot study. The principle of proxemics was introduced by Hall [6] and it describes a number of Zones or personal spaces that humans implicitly use to attribute a level of intimacy for their interpersonal relation which are personal, social, and public Zone in order of increasing distance to the person. These spaces are attributed to certain members of the public, like family, friends, or strangers. Their individual distant range is not only based on culture but also varies due to age and social status of the person [1]. Provided there is ample space to move, by allowing people to enter and stay within certain distances we are communicating our relationship to that person

[1]School of Mathematical and Computer Science, Heriot-Watt University, Edinburgh, UK `k.lohan@hw.ac.uk`

In our pilot study we are investigating the impact of the location of speech production on the likeability of the robot by varying the distance to the robot. Manipulating the sound source location within the social space, we are also analysing the variation's impact for the recognition capabilities of the robot.

II. Experiment

A. Participants

13 participants were taking part in this experiment and were randomly assigned to either of our two conditions (6 assigned to the speech from behind and 7 to the speech from robot condition). They are all students or working at the university in the area of computer science or robotic. Thus had previous experience with robots.

B. Setup

In our experiment the participants were asked to interact with the iCub robot. Our iCub robot was equipped with an algorithm able to detect participants pointing behaviour towards 3 objects on a table. The robot would confirm through speech which of the objects was pointed at. This feedback was on the one hand to make participants as well as the experimenter aware of the limitations of the system and on the other hand we used a manipulation on the location of the speaker to understand the participants acceptance of the robot as an agent. The participant was asked to stand in 3 different locations (Zone1-Zone3) from the robot and repeat the interactions with the iCub. More details of the setup can be found in figure 1.

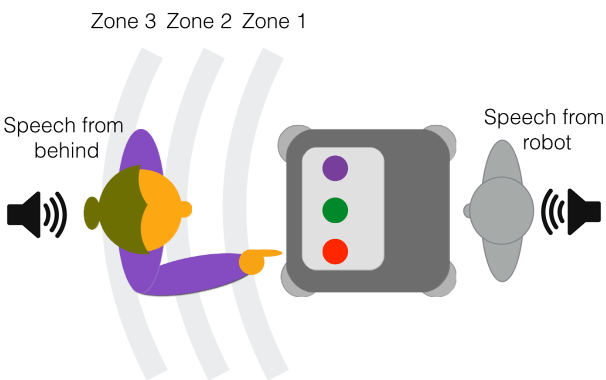

Fig. 1: Experiment setup showing the different speech source conditions and the distance Zones.

C. Procedure

After being asked for their consent and reading an instruction on how to interact with the robot, the participants had the chance to observe the experimenter showing the interaction with the robot and try the system themselves once before starting the experiment. During the experiment they were asked to perform the pointing towards the 3 objects on the table with their right hand standing in the 3 different Zones. Zone 1, 2 and 3 (1.25m, 1.5m and 1.75m) are within the social interaction Zone according to [6], [9]. All Zones are within a well recognized distance from the 3D sensor[13]. The participants asked to randomly point at an object in each Zone for 1.5 minutes and where told to move to the next Zone from the experimenter after this time. During this time the experimenter would note down the performance of the detection of the system as well as the order of objects pointed to. On conclusion of all 3 Zones participants were asked to answer the prepared Godspeed questionnaire [2].

D. System

The system was developed as a finite state machine, using a 3D sensor (ASUS xtion) to detect the human skeleton [10] as well as using a Hough Circle Transform [3], [11] and color filter to detect our objects. For the verbal feedback we used the lip-synchronization tool developed by Keller at el. [8]. This would use the iCub talking head [12], facial actuators to form the visemes in synchronization with the phonemes of the produced speech feedback.

E. Condition

The experiment carried out had a 2X2 design with 3 repetitions. In one condition the speech would be produced from a speaker behind the participant and in the other from the iCub speaker. The repetition was carried out in the 3 Zones. Speech produced by the system was kept the same between both conditions.

III. Result

We evaluated the experiment using 2 methods, the godspeed questionnaire and behaviour measures collected during the experiment. Results of these two measures will be presented separately below.

A. Questionnaire results

A one-way between subjects ANOVA was conducted to compare the effect of Anthropomorphism, Animation, Likeability, Perceived Intelligences and Perceived Safety on the presented system in the speech from behind and speech from the robot conditions (see Fig. 2). There was not a significant effect of Anthropomorphism for the presented system at the $p < .05$ level for the two conditions $[F(1, 11) = .855, p = .375]$. There was a significant effect of Animation for the presented system at the $p < .05$ level for the two conditions $[F(1, 11) = 5.608, p = .037]$. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the speech from behind condition (M = 2.25, SD = .546, N = 6) was rated significantly lower than the speech from robot condition (M = 3.07, SD = .679, N = 7). There was not a significant effect of Likeability for the presented system at the $p < .05$ level for the two conditions $[F(1, 11) = 3.553, p = .086]$. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the speech from behind condition (M = 2.967, SD = .871, N = 6) was rated lower than the speech from robot condition (M = 3.857, SD = .830, N = 7). There was not a significant effect on Perceived Intelligences for the presented system at the $p < .05$ level for the two conditions $[F(1, 11) = 1.036, p = .331]$. There was not a significant effect on Perceived Safety for the presented system at the $p < .05$ level for the two conditions $[F(1, 11) = .327, p = .579]$.

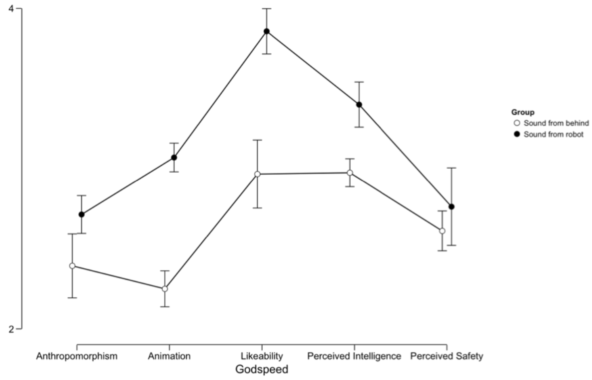

Fig. 2: Average Likert scale results of the Godspeed questionnaire items for each condition.

B. Behaviour results

A repeated-measure between subjects ANOVA was conducted to compare the effect of distance to the robot on the presented system in the speech from behind and speech from the robot conditions. There was a significant effect of the purple object for the presented system at the $p < .05$ level for the two conditions $[F(2, 22) = 5.731, p = .010]$. Post hoc comparisons using the Bonferroni test indicated that the mean score for Zone 3 was significantly lower than the results for Zone 1 (MD = -.455, SE = .121, t(12) = -3.753, p = .008) (see Fig. 3 on the left).

Further we investigated the speed the task was performed in and we found no significant difference using a repeated measure ANOVA (see Fig. 3 on the right).

IV. Discussion and Conclusion

Overall, we can see that there are 2 different effects present in our system. One is shown in the behaviour results on the different recognition capabilities of the system. The purple object was placed on the left side of the table which resulted in lower recognition by the right arm being in front of the body of the participant, which was worst if the participant was in Zone 3 (the furthest from the Robot).

The other effect is present in the GodSpeed results for the Animation as well as in a trend of the Likeability

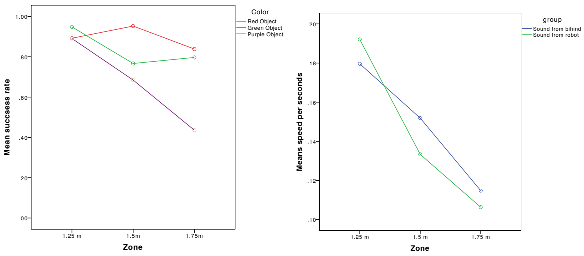

Fig. 3: Left figure shows the Mean success rate of the pointing feedback loop for each object in each distance Zone. Right figure presents the mean speed for task performance per seconds in each distance Zone.

items. Here we find that participants prefer to interact with the robot if the speech is produced from direction of the robot. It is interesting that there is no effect on either Anthropomorphism, Perceived Intelligences and Perceived Safety, which indicates that participants do not take speech location into account for these aspects.

We believe to answer the question if the speech location has an impact on the acceptability of a system like this and the implications on the relation between gesture communication and speech as feedback with further experiments need to be carried out. Including but not limited to using more locations for the speech as well as manipulation on the relations between gestures and speech produced.

We are especially interested in the distance between the mouth as visual output cue and the relative distance of the speech production as the audible cue in order to understand how much distance can be between them without irritating the user. We expect that this information can be useful for the development of new robots in order to accommodate for technical constraints, such as space limitations in humanoid heads, while maintaining as less irritation as possible.

References

[1] John R Aiello and Tyra De Carlo Aiello. The development of personal space: Proxemic behavior of children 6 through 16. *Human ecology*, 2(3):177–189, 1974.

[2] Christoph Bartneck, Elizabeth Croft, and Dana Kulic. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, 2009.

[3] Xing Chen, Ling Lu, and Yang Gao. A new concentric circle detection method based on hough transform. In *Computer Science & Education (ICCSE), 2012 7th International Conference on*, pages 753–758. IEEE, 2012.

[4] Kerstin Dautenhahn, Bernard Ogden, and Tom Quick. From embodied to socially embedded agents–implications for interaction-aware robots. *Cognitive Systems Research*, 3(3), 2002.

[5] Monica Gori, Giulio Sandini, and David Burr. Development of visuo-auditory integration in space and time. *Frontiers in integrative neuroscience*, 6, 2012.

[6] Edward T Hall. *The hidden dimension*. Anchor Books New York, 1969.

[7] Lohan K. Deshmukh A. & Aylett R. Hastie, H. The interaction between voice and appearance in the embodiment of a robot tutor. In *Proceedings of the 9th International Conference on Social Robotic (ICSR).*, 2017.

[8] Ingo Keller and Katrin Solveig Lohan. Lip synchronisation for the icub talking head. In *RoMan 2016*, 5 2016.

[9] Katrin Solveig Lohan, Hagen Lehmann, Frank Broz, Christian Dondrup, and Hatice Kose. *Enriching the human-robot interaction loop with natural, semantic and symbolic gestures*. Springer, 7 2017.

[10] Cyrille Migniot and Fakhreddine Ababsa. 3d human tracking in a top view using depth information recorded by the xtion pro-live camera. In *International Symposium on Visual Computing*, pages 603–612. Springer, 2013.

[11] opencv dev team. Hough circle transform.

[12] Alberto Parmiggiani, Marco Randazzo, Marco Maggiali, Giorgio Metta, Frederic Elisei, and Gerard Bailly. Design and validation of a talking face for the icub. *International Journal of Humanoid Robotics*, 12(03):1550026, 2015.

[13] Frank Weichert, Daniel Bachmann, Bartholomäus Rudak, and Denis Fisseler. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, 13(5):6380–6393, 2013.

Program Committee members

PC Member	Affiliation
Adeline Chanseau	University of Hertfordshire
Amol Deshmukh	Glasgow University
Christian Dondrup	Heriot-Watt University
Ingo Keller	Heriot-Watt University
Hagen Lehmann	Istituto Italiano di Tecnologia
Francesco Rea	Istituto Italiano di Tecnologia
Giulio Sandini	Istituto Italiano di Tecnologia
Eduardo Benitez Sandoval	Bielefeld University
Christiana Tsiourti	University de Geneve
Eli Sheppard	Heriot-Watt University

Organizers

Alessandra Sciutti

Istituto Italiano di Tecnologia

Alessandra Sciutti is the head of the Cognitive Robotics and Interaction Laboratory of the Robotics, Brain and Cognitive Sciences Department of the Italian Institute of Technology (IIT) in Genoa, (Italy), where she investigates the sensory-motor bases of human-human and human-robot interaction. She received her Ph.D. in Humanoid Technologies from the University of Genova in 2010 and after a research period at the Robotics Lab of the Rehabilitation Institute of Chicago (2011) and at the Emergent Robotics Laboratory of Osaka University (2014) she became Researcher at IIT. She has authored 20 papers in international journals, 2 book chapters and has presented her work at a number of international conferences in the field of human-robot interaction, cognitive and perceptual sciences. More information at:
https://www.iit.it/people/alessandra-sciutti

Katrin S. Lohan
Heriot-Watt University

Katrin Lohan joined the school of Mathematical and Computer Sciences at Heriot-Watt University as an assistant Professor in 2013. She is deputy director of the Robotics Lab. She became SICSA team leader in the Cyber Physical Systems research theme in 2016. She was General Chair for the European Robotics Forum 2017. She is hired under the Global Platform Recruitment for Research Leaders and part of the Edinburgh Centre for robotics. Previously, she was working at the Istituto Italiano di Tecnologia (IIT) as a Post Doc in the RobotDoc project funded by the Marie Curie Fellowship. She obtained her Ph.D. in Engineering from Bielefeld University, Germany in 2012, where she was associated with the ITALK Project. Her main research interests are in understanding the learning mechanisms between parents and infants, between adults and adults, and between humans and robots in order to create a natural interaction with a robot. Furthermore, she is interested in deep learning of semantic objects, both through vision and speech.

Friederike Eyssel
Bielefeld University

Friederike Eyssel is Professor of Psychology and head of the Group „Applied Social Psychology and Gender Research“ at the Center of Excellence Cognitive Interaction Technology (CITEC) at Bielefeld University, Germany. Friederike Eyssel has earned her Masters Degree in psychology from University of Heidelberg (Germany) in 2004. She received her PhD in social psychology from University of Bielefeld in 2007. Friederike Eyssel has held visiting professorship of social psychology at the University of Münster, the Technical University of Dortmund, the University of Cologne, and New York University in Abu Dhabi. Moreover, she co-founded a new conference series, the 'Joint UAE Symposium on Social Robotics' (JSSR 2015, 2016, 2017. Her main research interests in the domain of social robotics focus on psychological mechanisms of successful human-machine interaction, anthropomorphism, education robotics, and social implications of assistive technology developments.

References

- Pfeifer, R., Scheier, C., 1999. Understanding Intelligence. MIT Press, Cambridge, MA.
- Li, J., 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum Comput Stud 77: 23–27.
- Moosaei, M., Das, S. K., Popa, D. O., & Riek, L. D., 2017. Using Facially Expressive Robots to Calibrate Clinical Pain Perception. In Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. 32-41.ACM.

- Ferrari, F., Eyssel, F. A. 2016. Toward a hybrid society: The transformation of robots, from objects to social agents. In A. Agah, J. - J. Cabibihan, A. M. Howard, M. A. Salichs, & H. He (Eds.), Lecture Notes in Artificial Intelligence (LNAI): Vol. 9979. Social Robotics. Proceedings of the 8th International Conference on Social Robotics (ICSR 2016)/ (pp. 909-918). Heidelberg/ Berlin: Springer.
- Sciutti, A., Ansuini, C., Becchio, C., & Sandini, G., 2015. Investigating the ability to read others' intentions using humanoid robots. Front Psychol 6, 2015.
- Fischer, K., Lohan, K., & Foth, K. (2012, March). Levels of embodiment: Linguistic analyses of factors influencing hri. In *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on* (pp. 463-470). IEEE.

Acknowledgments

This workshop is supported by the European CODEFROR project (PIRSES-2013-612555) and by SICSA (http://www.sicsa.ac.uk/).