Gaze-Based Hints During Child-Robot Gameplay

Eunice Mwangi[1(✉)], Emilia I. Barakova[1], Marta Diaz[2], Andreu Català Mallofre[2], and Matthias Rauterberg[1]

[1] Designed Intelligence Group, Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands
emwangi@tue.nl

[2] Technical Research Centre for Dependency Care and Autonomous Living (CETpD), Universitat Politècnica de Catalunya, Vilanova i la Geltrú, Barcelona, Spain

Abstract. This paper presents a study that examines whether gaze hints provided by a robot tutor influences the behavior of children in a card matching game. In this regard, we conducted a within-subjects experiment, in which children played a card game "Memory" in the presence of a robot tutor in two sessions. In one session (Help), the robot tutor gives gaze hints to help children find matching cards and, in the other session (No_Help), the robot tutor does not provide help to the children. Gaze hints involved looking toward the correct matching cart. We analyzed the child-robot interaction regarding execution performance, gaze behavior, and level of engagement. Children performance was measured using the number of trials and overall time used to complete the game. We found that children used significantly fewer trials in the Help condition than in the No_Help condition. In addition, there were more instances of mutual gaze in the Help condition than in the No_Help condition. These findings suggest that designing a robot with dynamic gaze increases the effectiveness of the robot tutor as a helping agent.

Keywords: Gaze in HRI · Gaze perception · Games with social robots · Robots for (special) education, Joint attention design of human-robot interactions

1 Introduction

The adoption and use of social robots in education and intervention programs for children with autism has grown considerably in the last decade [3, 4, 6]. The potential benefits of social robots in this domain have been discussed in numerous studies [6, 9, 12]. For example, robots can provide personalized interactions by adapting to the cognitive and affective needs of children, robots can provide repetitive feedback, which is challenging for individual teachers in classrooms, and they have been shown to elicit interest from children [11]. Though there are many prospective benefits of social robots as mentioned, their successful implementation requires that robots provide appropriate and intuitive interactions with human users. Thus, the design of the social interaction between humans and robots is becoming critical.

A. Kheddar et al. (Eds.): ICSR 2017, LNAI 10652, pp. 413–422, 2017.
https://doi.org/10.1007/978-3-319-70022-9_41

Eye gaze behavior has been reported as a significant nonverbal cue and is considered essential for the establishment of mutual acknowledgment in human communication [10, 14]. An important communicative property of gaze, in particular, is its ability to direct attention to objects of interest [16]. Moreover, psychological studies indicate that, at a very young age, infants (from the age of three months) can follow the gaze of their parents or their caregivers [5]. Such gaze following behavior facilitates the formation of joint attention, which is essential for children's development and learning [10, 16]. Subsequently, joint attention is considered to be the foundation of the theory of mind and perspective taking, which are important in early childhood social cognition. Therefore, being able to understand and provide gaze cues is an important aspect in developmental robotics. Most remarkably, both psychologists, Astington and Jenkins [2], and roboticists, Dautenhahn [9]; Krämer et al. [15], draw on the theories of mutual understanding to explain how humans and robots may acquire the ability to represent each other's mind.

Prior studies have addressed the effects of gaze in human-robot interaction [1, 8]. However, to date, it is still not yet established the extent to which children can read, understand, and attribute meaning or intentions from a robots' gaze. To address the gaps mentioned above, we developed an interaction task to examine the effects of gaze hints exhibited by a social robot NAO, with children in a task-oriented activity. Our objective was to determine whether gaze hints from the robot tutor can direct the attention of children and, thus, influence their selections in a card game task. Based on the nonverbal theories of gaze [13, 14], we hypothesize that gaze hints will direct the attention of the children, either implicitly or explicitly to the matching card and, thus, children will perform better with the help of the robot tutor than without help. We further projected that occurrences of mutual gaze and joint attention would facilitate the effectiveness of the robot tutor as a helping agent.

2 Related Work

A lot of gaze research in human-robot interaction has taken a human-centered approach to examine the ability of humans to read, perceive, and interpret social cues exhibited by a robot gaze. For example, Mutlu and Colleagues [18], the authors consider the extent to which people can understand and attribute meaning to leakage signals using Geminoid and Robovie platforms. The authors claim that, in general, the gaze cue led to better performance on a task, and even better with Robovie than with Geminoid. Mwangi and colleagues [19] examined how people perceive gaze cues and head angles directed towards different target positions on a table when a human and NAO robot are facing each other as in board game scenarios. This study showed that when the head pitch angle is higher (24 ± 2) and the depth is less, approximately 20 cm from the robot, participants detect the positions with good accuracy for this particular robot.

Gaze has also been used to create more efficient interactions with robots [1, 7, 17, 24]. For example, Yoshikawa et al. [24] show that responsive robot gazes induce stronger feelings of being looked at compared to non-responsive gazes. A lot of work has also addressed the role of gaze in physical human-robot interaction to support collaborative work and as a mechanism to control turn-taking. For instance, in a dictation task, Palinko et al. [21] found that mutual gaze is an efficient means of controlling turn-taking with human partners with different needs. Yamazaki et al. [23] also demonstrated the importance of gaze timing in turn-taking interactions. Further, in a handover setting where the robot handed bottles to human subjects, [17] empirical evidence reveals that gaze cues can improve hand over timing and the subjective evaluation of the robot.

Several studies have explored how human behaviors differ in interactions with robots compared to humans. For instance, using a word-learning task [25], a micro-level analysis shows that participants spend more time gazing at the face of the robotic agent than at the human's face when naming objects. Boucher et al. [7] studied gaze effects in human-human and human-robot interaction on task completion. Their findings demonstrate that human participants can use the gaze cues of a human or of a robot partner in a physical interaction. The differences in how human behavior changes in interaction with humans compared to the robots has been attributed to many factors: for example, the novelty effect of the robot [20, 25], the robot's gaze being very overt with large head motions compared to more natural gaze movements that are much more subtle [20] as humans mostly use their eyes without much head movement.

3 Method

3.1 Participants

The participants were eighteen typically developing children, aged between 4 and 11 years. Ten of the participants were boys and the other eight were girls. A picture of the experimental setting can be seen in Fig. 1.

Fig. 1. Experimental setting for the experiment. Child interacting with the NAO robot in a card matching task. An experimenter is present at the session to support the child if necessary, this setting is typical for child-robot interaction sessions.

3.2 Experimental Setup

The setup included a humanoid robot NAO developed by Aldebaran Robotics [22], a personal computer, a webcam, and a game. The participant and a NAO robot sat on the two sides of a table facing each other. The table is approximately 80 cm in width, and the height of the table is 72 cm. A board grid with the card positions is fixed on the table. The layout has 18 squares (8 * 8 cm) organized in six (6) columns and three (3) rows. The 18 squares correspond to 18 card positions for the game. The squares are 10 cm apart in depth (y-axis) and 6 cm apart in width (x-axis). The layout is 600 mm in width and 900 mm long. To develop the game, we used Java Programming Language. Each card was labeled with a unique card code, and placed in a fixed position on the board layout marked with a head pitch and yaw angle on the computer layout. The algorithm was applied such that, after scanning the code of the selected card, the robot head angles shifted to the card position of the chosen card, then to the face of the participant (assumed at NAO's initial position), and then to the location of the matching card.

3.3 Procedure and Conditions

The children were asked to find matching pairs of cards on the table. In the beginning, the cards were laid face down on the board, and the child was required to find matching pairs of cards by turning the first one and then guess the matching one. If the two cards turned face up were identical (a pair), the child continued to a new try. Otherwise, the child turned the cards face down and made a new try/move. The game ended when the child found all the matching pairs.

Two experimental conditions, Help and No_Help, were executed by the robot tutor. In the help condition, the tutor provided gaze hints to help the participant find the matching cards. While introducing the game, the robot tutor informed the children that it would help them. However, the tutor did not explicitly reveal the modality it would use to help. In the Help condition the tutor remained silent the entire session; when the child turned the card upwards, the tutor looked at the flipped card followed by the glance to the child's face and then to the matching card. In the No_Help condition, the robot tutor remained silent and only looked at the child.

When each child completed a session, he/she left the room to allow the experimenter to re-arrange the game and was later welcomed back (approximately after four minutes) for the next session. At the end of each session, the robot verbally thanked the child. During the experiment, both the experimenter and the researcher were present in the room in each session. The researcher sat in a corner in front of a big screen, remained silent for the entire duration, and tracked and recorded the actions of the child during the game from the webcam display. To control the robot, the researcher scanned on the keyboard the numbers corresponding to the cards picked by children so the robot could then execute the corresponding to this particular card position head movements. After both sessions, the experimenter asked each child a few post-experiment questions on whether they noticed the help cues from the tutor and, if they did, whether that influenced their choice of cards.

3.4 Measurements and Hypothesis

The main goal of this experiment is to examine the effect of help on child-robot interaction. In this regard, we formulated two hypotheses: how help will impact children's execution performance of the task and, also how help will affect gaze behavior of the children.

Hypothesis One (H1): Children will perform better (in terms of time and number of trials) in the Help condition than in the No_Help condition.
Hypothesis Two (H2): Children will look more into the robot's face in the Help condition than in the No_Help condition.
H2.1: There will be more events of mutual gaze in the Help condition than in the No_Help condition.

3.5 Child-NAO Behavioral System

We developed a coding scheme focused on three aspects: gaze behavior, engagement, and emotions experienced during the interaction.

Gaze behavior. We created three categories of gaze behaviors as follows:

Child: The observation scheme captures children' gaze at the Robot face, Card 1, Card 2, Board, Experimenter, and Elsewhere. "Card 1 and Card 2" refer to the two cards the participant manipulates while making every trial. Looking at the board refers to generally gazing at cards on the board, probably when figuring which one to flip next. Looking elsewhere refers to objects outside the game and its associated components (robot, board, experimenter), for example, looking at computers, windows, chairs, doors, or at the roof. Looking elsewhere can indicate moments of disengagement during the game. In the further analysis, we included the matching card category to help examine patterns of successful joint attention.
Robot: The robot first looks at the card the child flips (Card 1), then at the child's face, and finally looks to the matching card.
Simultaneous gaze: The purpose of the simultaneous gaze analysis is to examine the interaction between the gaze behavior of the child and the robot. We focus on two social concepts of gaze, including mutual gaze and joint attention.

Gaze behaviors were defined as states and behaviors by the same subject were mutually exclusive and exhaustive. During the experiment, we recorded a total of 35 videos using a high-resolution webcam to capture the view of both the robot and the child, a total of 30 videos of 15 children both in the Help and No_Help condition were coded and included for analysis. The behaviors for the child and the robot were coded separately; that is, the researcher first coded the videos for child behaviors and then at the second round for robot behaviors. To code, we used Noldus Observer XT software, developed and commonly used for behavioral analysis. Reliability checks were performed throughout the coding process through randomly selecting and checking videos.

Engagement: The second part of the scheme focused on the engagement of the child with the game, the robot, and the experimenter. We defined two categories as states: (1) disengaged. During coding we marked the attributes of behavior associated with disengagement such as closing eyes, looking down or head down, the child covering their face, the child becoming impatient, and looking elsewhere as defined in the gaze section; (2) For the engaged state, the assumption was that if the child was not disengaged, i.e. does not exibit either of the above mentioned cues, then she/he was engaged.

4 Results

For this analysis, we considered a total of 15 children (age 6–11; Mean Age = 7.6), for a total of 15 trials in the Help condition and 15 trials in the No_Help conditions. Three (3) children were excluded from this analysis owing to the following reasons: one of the children declined to participate in the second session and the other two were very young (below age 5) and needed a lot of help from the experimenter to play the game. From the post-experiment interview, eight (8; approximately 53%) out of the fifteen (15) children said they noticed the help hints from the tutor while the others (7: 47%) stated that they did not see the help gaze cues.

4.1 Performance Analysis

To evaluate the effects of gaze hints on children's performance, we identified two measures: (1) *Duration:* the time it takes participants to find all pairs of matching cards on the table; and (2) *Trials:* The number of trials required to find all matching cards. A trial consists of choosing two cards. For this analysis, we conducted a repeated measures ANOVA, with Help_Type (Help vs. No_Help) as the within subject factor.

Duration: There was no significant difference in duration of the game between the Help and No_Help conditions (Help = 167.87 s; No_Help = 169.6 s; F (1, 14) = 0.015, $p = 0.905$.). Furthermore, there was no significant correlation between noticing gaze and the duration the children took to play the game ($p = 0.\ 530$, two-tailed).

Number of trials: There was a significant difference in the number of trials of the game between the Help and No_Help condition (Help = 14.07 trials; No_Help = 17 trials; F (1, 14) = 5.331, $p = 0.03$.). Further analysis shows a significant correlation between noticing gaze hints and the number of trials ($p = 0.001$, two-tailed).

4.2 Behavioral Analysis

The visualization below shows a time event plot of the gazing behaviors of a child and the robot in the Help condition.

The upper sequence in Fig. 2 shows the behaviors of the child and the bottom sequence shows the behavior of the robot. Each behavior appears on the visualization if the behavior is coded at least once in the system. The length of the boxes indicates the duration and the incidences of similar colored bar rectangles show the frequency of the occurrences of each behavior. For the gaze behavior of the child we analyzed both the duration of looking and frequency of gaze as follows: We conducted a repeated measure ANOVA, with Help Type (Help vs. No_Help) as the within-subject factor. The paragraphs below report on the results.

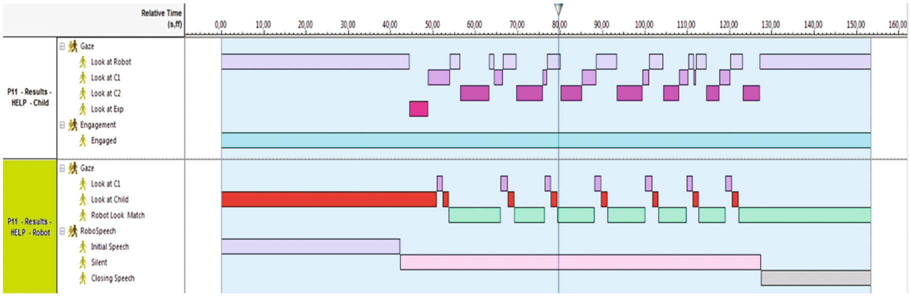

Fig. 2. Visualization of the behaviors of the robot and the child in Help condition. In the plot it is seen that the gaze behavior of the children co-occurs with the ones of the robot

Eye-gaze analysis: Based on quantified coded data, the children's eye gaze frequency and durations were calculated for each condition (Help and No_Help conditions). *Frequency:* There was a significant difference in the number of occurrences of looking at the robot during the game between the Help and No_Help condition: (Help = 14.07 tries; No_Help = 17 trials; F (1, 14) = 5.331, p = *0.03.*). However, we found no significant differences for the rest of the categories of gaze behavior: looking at card 1, looking at card 2, looking at the board, looking at the experimenter, and looking elsewhere.

Duration: The duration was defined as the entire session between the child starting to play the game and completing it. The percentage shows the ratio of eye gaze duration to the total time the child took to play the game, as duration varied with every game session. There was a significant difference in eye gaze durations percentages for looking at the robot between the Help and No_Help conditions (Help = 28.41%; No_Help = 7.21% duration; F (1, 14) = 14.06, p = *0.002.*). Children looked at the robot significantly longer when the tutor was helping than when there was no help. Similarly, there was a significant difference in eye gaze durations for looking at Card 1 between the Help and No_Help conditions (Help = 21.91% (observation duration); No_Help = 32.88% (observation duration); F (1, 14) = 18.25, p = *0.001.*). Children looked significantly longer to Card 1 when the tutor was not helping than when the tutor was helping. However, there was no difference in looking duration at Card 2 in between the two conditions. A plausible explanation for this is given in the discussions section. In percentage duration for looking at the board (Help = 1.38%; No_Help = 4.46% duration; F (1, 14) = 3.75, p = *0.073.*), children looked longer at the board when there was no help than when there was help. However, the difference was not significant. There was no significant difference in percentage durations of looking at the experimenter (p = *0.56.*) and looking elsewhere *(p = 0.65.)* between both conditions.

Mutual Gaze: Mutual gaze was examined as a combination of RLC (Robot: Look at Child) – CLR (Child: Look at Robot). We found a significant difference in the number of occurrences of the mutual gaze between the Help and No_Help conditions (Help =

5.53; No_Help = 1.67; $p = 0.001.$). There were significantly more occurrences of mutual gaze when the tutor was helping compared to the when there was no help.

4.3 Engagement

We examined the percentage levels of engagement during Help and No_Help conditions. From our findings, there was no significant difference in engagement during the game between the Help and No_Help conditions (Help = 99.38% (Observational duration); No_Help = 97.89%; p = 0.247.). Similarly, there was also no significant difference in disengagement during the game between the Help and No_Help conditions (Help = 0.62% (Observational duration); No_Help = 2.01% (Observational duration); p = 0.281).

5 Discussion and Conclusion

In this study, we examined if gaze hints provided by a robot tutor influenced the behavior of children in Memory game (card matching). We analyzed the child-robot interaction regarding execution performance, gaze behavior and level of engagement. The performance was measured through the number of trials, and the overall time to complete the game. The findings partially support our first hypothesis in that children used significantly fewer trials in the Help condition than in the No_Help condition. However, there was no significant difference in duration taken by children between the Help and No_Help conditions.

There was no significant correlation between noticing the gaze and the game duration. There are several possible explanations for this. Firstly, as soon as the children noticed that the robot tutor was helping them with gaze hints, they waited until the robot showed them the matching card, even when they had an idea of where the matching card was. Secondly, the novelty effect of the robot, which is supported by the larger duration of looking at the robot even when the number of trials was less, indicates that the children who noticed gaze could have spent more time looking at the robot. Another probable reason is the duration of head movements of the robot during attention shifts from the flipped card to the face of the participant and then to the matching card. Lastly, a few of the children spent some time asking the experimenter questions during the game either because of confusion or when they saw the robot movement and could not interpret what it was doing.

The findings support hypothesis (H2.1) in that children looked significantly longer at the robot's face in the Help condition than in the No_Help condition. We also found that children gazed at Card 1 was significantly longer in the No_Help condition as compared to the Help condition. This could be because of the fact when the children flipped the first card (Card 1) they did not look at it as they knew the robot would help. Again, interestingly, we found no significant difference of eye gaze duration at Card 2. This could be explained by the fact that in the Help condition, the children were looking at Card 2 longer to confirm it is the match i.e. the same, since all cards were a shadue images of dogs. We found negligible percentages of time durations of looking elsewhere. This result confirms our results on engagement, as we found no significant differences during the Help and No_Help conditions - there was high of engagement during both sessions of the game. Results also reveal more events of mutual gaze in the Help condition than in the No_Help condition. This supports our (H2.2), which further projected that occurrences of mutual gaze would facilitate the effectiveness of the robot tutor as a helping agent. In i follow up study, we will analyse the dynamic and sequential analysis of gaze, which will determine the patterns of timing related to joint attention. From the subjective feedback we got from the post-experiment interview, children who noticed the help from the robot regarded the robot as friendly, and they defined the robot as being “cool”.

The results from this study differ a bit from our previous study with adult participants using similar setting. In the study with adults, we found that the participants performed better when the tutor was helping than without help [19]. Likewise, in both settings for the adult and children, most indicated that they expected verbal help from the tutor. We assume the design of the robot behavior may have led participants to expect verbal support from the tutor since, in the beginning, the robot verbally gave instructions to the participant but remained silent during the entire interaction, only giving gaze clues. In our ongoing analysis, we examine the sequential dynamics of joint attention gaze including the temporal patterns for this kind of interaction.

Acknowledgement. This work was supported in part by the Erasmus Mundus Joint Doctorate in Interactive and Cognitive Environments, which is funded by the EACEA Agency of the European Commission under EMJD ICE FPA no 2010-0012. We thank Kyra Fredricks, for her valuable support during the experiment, and all the children who participated in the study, the teachers and all parents/guardians who made this study possible.

References

1. Admoni, H., Scassellati, B.: Social Eye gaze in human-robot interaction: a review. J. Hum.-Robot Interact. **6**(1), 25–63 (2017)
2. Astington, J.W., Jenkins, J.M.: Theory of mind development and social understanding. Cogn. Emot. **9**(2–3), 151–165 (1995)
3. Balch, T., Summet, J., Blank, D., Kumar, D., Guzdial, M., O’hara, K., Walker, D., Sweat, M., Gupta, G., Tansley, S., Jackson, J.: Designing personal robots for education: hardware, software, and curriculum. IEEE Pervasive Comput. **7**(2), 5–9 (2008)
4. Barakova, E.I., Bajracharya, P., Willemsen, M., Lourens, T., Huskens, B.: Long-term LEGO therapy with humanoid robot for children with ASD. Expert Syst. **32**(6), 698–709 (2015)
5. Baron-Cohen, S.: Mindblindness: An Essay on Autism and Theory of Mind. MIT Press, Cambridge (1997)
6. Belpaeme, T., Baxter, P.E., Read, R., Wood, R., Cuayáhuitl, H., Kiefer, B., Racioppa, S., et al.: Multimodal child-robot interaction: building social bonds. J. Hum.-Robot Interact. **1** (2), 33–53 (2012)
7. Boucher, J.D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., Ventre-Dominey, J.: I reach faster when I see you look: gaze effects in human–human and human–robot face-to-face cooperation. Frontiers in Neurorobot. **6**(3), 1–11 (2012)
8. Broz, F., Lehmann, H., Mutlu, B., Nakano, Y.: Gaze in Human-Robot Communication. John Benjamins Publishing Company (2015)

9. Dautenhahn, K.: Roles and functions of robots in human society: implications from research in autism therapy. Robotica **21**(4), 443–452 (2003)
10. Frischen, A., Bayliss, A.P., Tipper, S.P.: Gaze cueing of attention: visual attention, social cognition, and individual differences. Psychol. Bull. **133**(4), 694 (2007)
11. Gordon, G., Breazeal, C., Engel, S.: Can children catch curiosity from a social robot? In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI 2015), pp. 91–98. ACM, New York (2015)
12. Huskens, B., Verschuur, R., Gillesen, J., Didden, R., Barakova, E.: Promoting question-asking in school-aged children with autism spectrum disorders: Effectiveness of a robot intervention compared to a human-trainer intervention. Dev. Neurorehabil. **16**(5), 345–356 (2013)
13. Kendon, A.: Some functions of gaze-direction in social interaction. Acta Physiol. **26**, 22–63 (1967)
14. Kleinke, C.L.: Gaze and eye contact: a research review. Psychol. Bull. **100**(1), 78 (1986)
15. Krämer, N.C., Eimler, S., von der Pütten, A., Payr, S.: Theory of companions: what can theoretical models contribute to applications and understanding of human-robot interaction? Appl. Artif. Intell. **25**(6), 474–502 (2011)
16. Langton, S.R.H., Watt, R.J., Bruce, V.: Do the eyes have it? Cues to the direction of social attention. Trends Cogn. Sci. **4**(2), 50–59 (2000)
17. Moon, A., Troniak, D.M., Gleeson, B., Pan, M.K.X.J., Zheng, M., Blumer, B.A., Croft, E. A., et al.: Meet me where I'm gazing. In: Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction, HRI 2014, pp. 334–341 (2014)
18. Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H., Hagita, N.: Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior. In: Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, pp. 69–76, March 2009
19. Mwangi, E., Barakova, E., Zhang, R., Diaz, M., Catala, A., Rauterberg, M.: See where I am looking at: perceiving gaze cues with a NAO robot. In: Proceedings of the Fourth International Conference on Human Agent Interaction, pp. 329–332. ACM, October 2016
20. Mwangi, E.N., Barakova, E.I., Díaz, M., Mallofré, A.C., Rauterberg, M.: Who is a better tutor?: gaze hints with a human or humanoid tutor in game play. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp. 219–220. ACM, March 2017
21. Palinko, O., Sciutti, A., Schillingmann, L., Rea, F., Nagai, Y., Sandini, G.: Gaze contingency in turn-taking for human robot interaction: advantages and drawbacks. In: 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 369–374, August 2015
22. Robotics, A.: Nao robot reference manual. Version **1**(10), 10 (2010)
23. Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M., Kuzuoka, H.: Precision timing in human-robot interaction: coordination of head movement and utterance. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 131–140, April 2008
24. Yoshikawa, Y., Shinozawa, K., Ishiguro, H., Hagita, N., Miyamoto, T.: Responsive robot gaze to interaction partner. In: Robotics: Science and Systems, August 2006
25. Yu, C., Schermerhorn, P., Scheutz, M.: Adaptive eye gaze patterns in interactions with human and artificial agents. ACM Trans. Interact. Intell. Syst. (TiiS) **1**(2), 13 (2012)