Development and Perception Evaluation of Culture-specific Gaze Behaviors of Virtual Agents

Tomoko Koda[1,2] ✉, Taku Hirano[1], and Takuto Ishioh[2]

[1] Faculty of Information Science and Technology, Osaka Institute of Technology, Osaka, Japan
[2] Graduate School of Information Science and Technology, Osaka Institute of Technology, Osaka, Japan

tomoko.koda@oit.ac.jp, m1m16a03@st.oit.ac.jp

Abstract. Gaze plays an important role in human-human communication. Adequate gaze control of a virtual agent is also essential for successful and believable human-agent interaction. Researchers on IVA have developed gaze control models by taking account of gaze duration, frequency, and timing of gaze aversion. However, none of this work has considered cultural differences in gaze behaviors. We aimed to investigate cultural differences in gaze behaviors and their perception by developing virtual agents with Japanese gaze behaviors, American gaze behaviors, hybrid gaze behaviors, and full gaze behaviors. We then compared their effects on the impressions of the agents and interactions. Our experimental results with Japanese participants suggest that the impression of the agent is affected by participants' shyness and familiarity of the gaze patterns performed by the agent.

Keywords: gaze, shyness, intelligent virtual agents, non-verbal behavior, cross-culture, perception, evaluation

1 Introduction

Intelligent virtual agents (IVAs) that interact face-to-face with humans are beginning to spread to general users across cultures, and IVA research is being actively pursued. IVAs require both verbal and non-verbal communication abilities to achieve natural interaction with humans. Among those non-verbal behaviors, gaze plays an important role in our social interactions, including controlling the flow of a conversation, indicating interest and intentions, and improving the listener's attention and comprehension [1, 2]. As in humans, a virtual agent's gaze behavior is important for facilitating natural interaction. Previous research on modelling gaze behavior of virtual agents has investigated appropriate turn management [3], where to look [4], making idle gaze movements [5], expressing social dominance by gaze [6], and the appropriate amount of gaze to facilitate interaction [7, 8]. All of these studies have modelled realistic human gaze behavior to an agent, resulting in more natural and smooth interaction.

Gaze perception and preferences are also affected by personality. For example, being gazed at can lead to discomfort from feeling observed, especially for shy people [9]. Shyness is defined as "discomfort and inhibition in the presence of others, where

[Tapez ici]

J. Beskow et al. (Eds.): IVA 2017, LNAI 10498, pp. 213-222, 2017.
DOI 10.1007/978-3-319-67401-8_25

these reactions derive directly from the social nature of the situation" [9]. Shy people tend to avert gaze and engage in more self-manipulation [10, 11]. Thus, shy people might not prefer to interact with a virtual agent that exhibits the realistic social human gaze behavior that is believed to facilitate smooth interaction. Our previous research investigated how shy people perceive different amounts of gaze from a virtual agent and how their perception of gaze affects the comfortableness of the interaction [12]. The results indicated that shy people are sensitive to even a very low amount of gaze from the agent. However, contrary to our expectations, as the amount of gaze from the agent increased, shy people had a more favorable impression of the agent, and they did not perceive the typically adequate amount of gaze (66%) as most comfortable. On the other hand, non-shy people perceived the gaze condition and recognized the adequate amount of gaze as most friendly.

None of the above IVA research, however, has addressed cultural differences in gaze behaviors, despite researchers in psychology having reported cultural difference in gaze behaviors and their perception. We believe there is a strong need to develop enculturated agents by making them exhibit culture-specific non-verbal behaviors such as gaze. Although this study focused on gaze behaviors, the importance of culturally adaptive IVAs for successful agent interactions has been suggested by Rehm and Nakano [13], who focused on gestures and postures, by Koda [14] in an investigation of facial expressions, and by Kuhne and Finkelstein [15, 16] in research on linguistic alignment. Culturally aware agent applications have been implemented as a culture training system [17] and to raise cultural awareness [18].

In terms of culture-specific gaze behaviors, there are findings from observation and video analysis of human-human and human-agent interactions that show cultural differences. Mayo [19] found that gaze patterns differ according to the culture of the conversant by analyzing gaze behaviors in video recordings of human-human conversations. Elzinga [1] reported that Japanese individuals had "more frequent and shorter lasting other-directed gazes" than did Australian participants, and that English-speaking participants looked at the other person to signal turns, while Japanese participants did not [20]. Argyle found that Swedes gaze at their conversation partner more than English participants do (50% vs. 38% of the time) [1].

In terms of perception of gaze behaviors, studies have indicated cultural preferences in the gaze amount that one receives. According to Cook [21], favorableness of impression is a linear function of the amount of gaze a person receives, and 50% of gaze amount gave the most favorable impression toward the human gazer in an experiment conducted in UK. Fukayama et al. [22] changed the amount of gaze from a virtual agent by 25%, 50%, 75%, and 100% and compared the agent's impressions by Japanese evaluators. The results showed that 50% gaze was perceived as most friendly, followed by 75%, then 25%. Impressions of friendliness plummeted in the full gaze (100%) condition.

If there are cultural differences in performing gaze behaviors, there should also be cultural differences in perceiving gaze behaviors of other cultures. We aimed to investigate cultural differences in gaze behaviors and their perception by developing virtual agents with Japanese gaze behaviors, American gaze behaviors, their hybrid gaze behaviors, and full gaze behaviors, and to compare their effects on the impressions of the

agents and interactions. In our previous work [12], non-shy people recognized the adequate amount of gaze (66%) and perceived the condition as most comfortable, while shy people did not. Accordingly, it is possible that the impression of the agent is affected by participants' shyness and the familiarity of the gaze patterns performed by the agent. Considering our findings that shy people's sensitivity to gaze amount and non-shy people's recognition of an adequate amount of gaze [12], we formed the following two hypotheses: H1) Shy people form worse impressions of agents whose gaze model does not originate from the same culture; H2) Non-shy people are more tolerant of gaze models from other cultures. This paper reports our experimental results with Japanese participants.

2 Gaze Models

We implemented American gaze behaviors (AG), Japanese gaze behaviors (JG), hybrid gaze behaviors (HG), and full gaze behaviors (FG) in our virtual agent in order to compare the impression of different cultural gaze behaviors.

AG was implemented in accordance with the gaze model proposed by Cassell et al. [23]. Their model shows American gaze patterns by analyzing video recordings of human dyad conversations. The model shows the probability of "looking away" at the beginning (44%) and end (84%) of an utterance. Fig. 1a shows the state transition diagram of AG at the beginning of an utterance, and Fig. 1b shows the AG at the end of an utterance as implemented in our AG model. Our AG also includes a gaze pattern at the end of a question, where the agent "gazes at" the user (human participant). The agent "looks away" from the user for 0.5 seconds 44% of the time at the beginning of an utterance, then shifts toward a "gaze-at" state. The agent keeps its gaze toward the user during the utterance. The agent "looks away" from the user for 2 seconds 84% of the time at the end of the utterance. The "look away" timing at the end of the utterance is calculated by estimating the duration of the synthesized speech. The agent keeps its "gaze-at" state toward the user while listening.

JG was implemented in accordance with the gaze model proposed by Ishii et al. [7, 8]. Their model shows Japanese gaze patterns by analyzing video recordings of three-way human conversations. We implemented JG by modifying their gaze model for dyad conversations by eliminating the state transitions to the third person. Fig. 2a shows the state transition diagram of JG. The agent "gazes at" the user at the beginning of an utterance, maintains the gaze for 1.1 to 3.1 seconds, and then shifts its gaze to "vague gaze" (described in section 3) for 3.2 to 7.9 seconds. The agent shifts its gaze pattern to "gaze-at" state 67% of the time or to "averted gaze" for 2.0 seconds 33% of the time after the "vague gaze." "Gaze aversion" is continued 13% of the time or shifted to the "gaze-at" state 87% of the time at the end of gaze aversion. The agent follows the gaze transitions during its utterance and while it is listening.

HG was implemented by combining JG and AG. As gaze behaviors and patterns are dependent on culture [19], we implemented HG as a culture-independent model that was neither American nor Japanese. Fig. 2b shows the state transition diagram at the beginning of and during an utterance. The agent follows the transition of AG at the

beginning of an utterance, and then follows JG during the utterance. The agent follows the state transition diagram of AG while listening. In addition to AG, JG, and HG, we implemented FG, a full gaze model, as a control gaze condition.

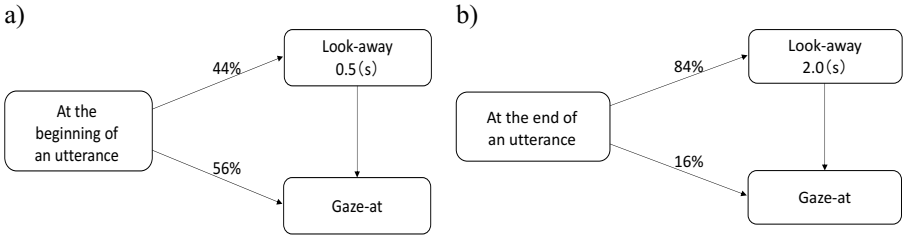

Fig. 1. a) State Transition Diagram of American Gaze Behavior at the Beginning of an Utterance, b) at the End of an Utterance

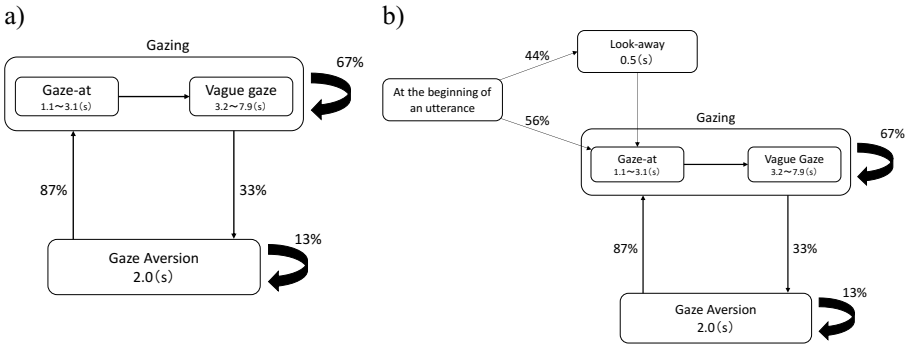

Fig. 2. a) State Transition Diagram of Japanese Gaze Behavior, b) Hybrid Gaze Model

3 Virtual Agent and Gaze Animations

The agent's appearance and gaze animations were developed by Unity 5.2.1fl (https://unity3d.com/) and Taichi Character Pack asset (https://www.assetstore.unity3d.com/jp/#!/content/15667). The agent's voice was synthesized with AITalk (http://www.ai-j.jp/english/). The gaze behaviors implemented in the agent were the four types described in section 2, namely, "gaze-at," "vague gaze," "look-away," and "gaze-aversion".

"Gaze at" is a state in which the agent keeps gazing at a user (shown in Fig. 3a: left, b: top). "Vague gaze" is described as follows: "in order to express less-face-threatening eye-gaze in virtual space avatars" [7, 8], which was implemented with the agent looking five degrees lower than the user's eye position (shown in Fig. 3a: right, b: bottom). "Look-away" was implemented as an animation in which the agent discontinues its gaze for 0.5 seconds and looks up, as in Gambi's agent [24], which was implemented in accordance with Cassell's American game model [23]. The agent looks up (in "look-away" state for 0.5 seconds) before an utterance (shown in Fig. 4a). "Gaze aversion" was implemented in two directions, to the right and left, and each aversion lasts for 2

seconds, as described in [8] (shown in Fig. 4b). A validation check for each gaze animation was conducted by 8 university students. The agent performs blinks, head rotation, lip sync, and the gaze patterns. The gaze behaviors are fully automated according to a dialogue the agent speaks. The amount of gaze in each condition is not fixed, as it is affected by the duration of utterance from the agent and a participant, as well as randomized transition processes of each model. However, we can estimate that the amount of gaze becomes greater, in order, for HG, JG, AG, and FG.

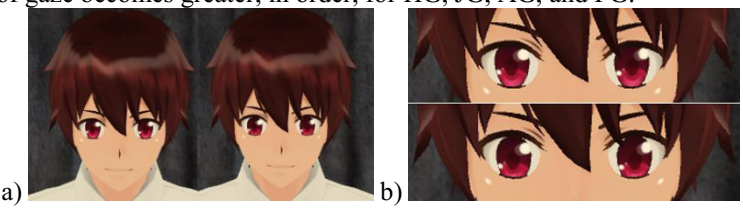

Fig. 3. Agent's Gaze-at State (a: left, b: top) and Vague Gaze State: 5 degrees lower than gaze-at (a: right, b: bottom)

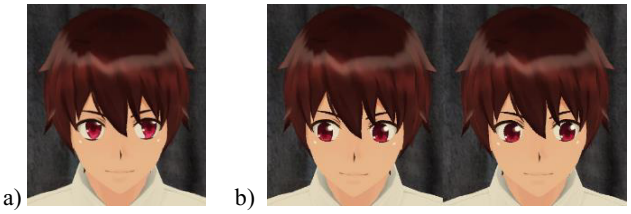

Fig. 4. a) Agent's Look-away state, b) Agent's Gaze Aversion States (to left and right, in either direction)

4 Experiment

4.1 Experimental Procedure

The experiment was conducted as a Wizard of Oz experiment. Participants were asked to have four formal conversational sessions with a conversational virtual agent to assess its functions. The true purpose of the experiment was not explained to the participants during the experiment. The agent's gaze models and conversational topics were randomly assigned in each of four conversation sessions in order to minimize the effect of conversational content. The topics included the US Election, Senior Driving, Pokémon GO, and POP Icons. Each conversation lasted about 2 minutes. The agent brought up the issue at asked the participants for their opinions. The agent's reply was controlled by a Wizard. The distance between the agent and participant was set to 1.8 meters, which is defined as an appropriate social distance by Hall [25]. The agent's upper torso was displayed on a 42-inch display, and the display's height was adjusted to the eye level of each participant. The participants were asked to answer a questionnaire after each session, indicating the agent's perceived shyness, perceived friendliness, their

friendly feelings toward the agent, and comfortableness and naturalness of the conversation.

In this experiment, we recruited 18 Japanese university students (15 men and 3 women, 19 to 23 years old). Their shyness level was measured by the Shyness Scale [26] beforehand. We divided the 18 participants into three groups according to their shyness scores. The average shyness score of Japanese university students in engineering was reported in a previous study to be 47.56 [26]. Seven participants with scores above 51 were categorized as the high shyness group (HS), 4 with scores between 45 and 48 were categorized as the mid shyness group (MS), and 7 with scores below 41 were categorized as the low shyness group (LS). We only analyzed the results of HS and LS to investigate the effects of shyness. The experimental conditions were gaze model (four models: FG, AG, JG and HG) and participant's shyness (HS/LS).

5 Results

5.1 Analyses of Perceived Shyness of the Agent and Naturalness of the Interaction

For perceived shyness of the agent, a 2-way repeated measures ANOVA showed a significant main effect of gaze condition ($F = 11.105$, $p < 0.01$). Fig. 5 shows the results of multiple comparisons. AG was perceived as significantly less shy than were other gaze conditions by both the HS ($F = 5.45$, $p < 0.01$) and LS individuals ($F = 5.69$, $p < 0.01$). There were no significant interactions. This indicates that regardless of the shyness of the participants, AG was perceived as less shy than were the other gaze models. It is interesting that FG was not perceived as least shy, when the amount of gaze was 100%. As for perceived naturalness of the interaction, a 2-way repeated measures ANOVA showed no significant main effects or interactions. This indicates that all interactions are perceived as equally natural regardless of the gaze conditions and participants' shyness (ratings of 4–5 on a 1–7 scale).

5.2 Analyses of Friendliness from / toward the Agent and Comfortableness of the Conversation

Regarding perceived friendliness from the agent, a 2-way repeated measures ANOVA showed a significant main effect of gaze condition ($F = 4.462$, $p < 0.01$). Fig. 6 shows the results of multiple comparisons. The HS group perceived JG and FG as significantly more friendly as compared to AG and HG (JG–AG: $F = 5.770$, $p < 0.05$; JG–HG: $F = 7.296$, $p < 0.01$; FG–AG: $F = 7.705$, $p < 0.01$; FG–HG: $F = 6.573$, $p < 0.05$). However, the LS group perceived AG to be equally as friendly as JG and FG, and significantly more friendly than HG ($F = 5.517$, $p < 0.05$). The perceived friendliness of AG showed a significant difference by participants' shyness. Specifically, the LS group evaluated AG as significantly more friendly than did the HS group ($F = 6.229$, $p < 0.05$).

These results indicate that HS individuals are sensitive to the change of gaze patterns and perceived less friendliness from gaze patterns with which they are not familiar (i.e., AG and HG). LS individuals are more tolerant of the change in gaze patterns and

perceived equal friendliness from gaze patterns with they are not familiar, except for HG. These results are in accordance with the result of our previous study, which suggested that HS individuals are sensitive to even a very small amount of gaze from the agent as well as changes in gaze amount [12].

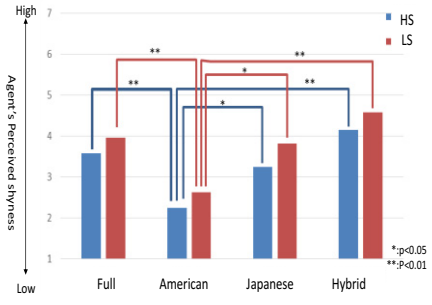

Fig. 5. Agent's Perceived Shyness

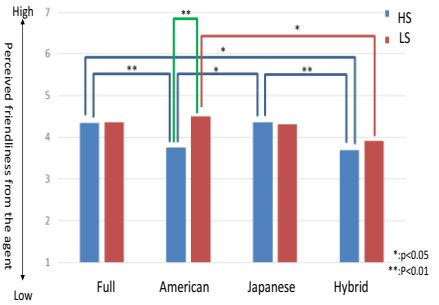

Fig. 6. Perceived friendliness from an agent

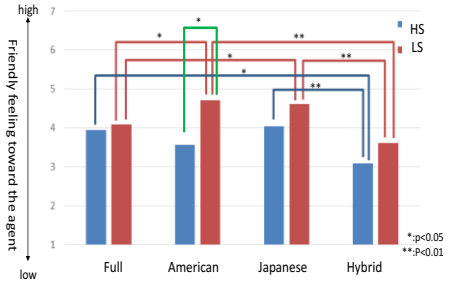

Fig. 7. Friendly feeling toward the agent

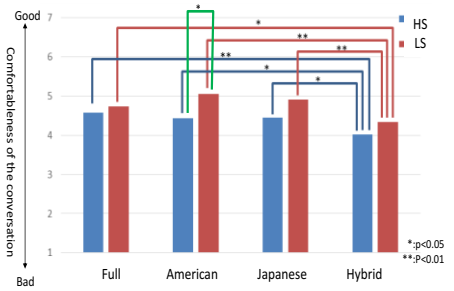

Fig. 8. Comfortableness of the Conversation

As for friendly feelings toward the agent, a 2-way repeated measures ANOVA showed a significant main effect of gaze ($F = 7.728$, $p < 0.01$) and shyness ($F = 4.151$, $p < 0.05$). The result of multiple comparisons is shown in Fig. 7. HS individuals felt significantly more friendly toward JG ($F = 8.565$, $p < 0.01$) and FG ($F = 6.31$, $p < 0.05$) as compared to HG. While LS individuals had more friendly feelings toward JG and AG as compared to FG and HG (JG–HG: $F = 9.443$, $p < 0.01$; JG–FG: $F = 4.345$, $p < 0.05$, AG–HG: $F = 12.747$, $p < 0.01$, AG–FG: $F = 4.356$, $p < 0.05$). Friendly feelings toward AG showed a significant difference between shyness groups. LS individuals had significantly more friendly feelings toward AG than did HS individuals ($F = 8.649$, $p < 0.01$).

HS individuals tended to have less friendly feelings toward AG (although not significant) and HG as compared to FG and JG, while LS individuals felt friendly feelings toward AG and JG. HS individuals' overall friendly feelings toward the agent were lower than were those of LS individuals. These results show similar trends to the results for agent's perceived friendliness.

As for comfortableness of the conversation, a 2-way repeated measures ANOVA showed a significant main effect of gaze condition ($F = 7.979$, $p < 0.01$). Fig. 8 shows

the results of multiple comparisons. HS individuals felt less comfort than did LS individuals in all gaze models. In particular, AG was evaluated as less comfortable by HS than by LS individuals ($F = 5.355$, $p < 0.05$), a similar trend to that for friendly feelings. HG was evaluated as less comfortable than were other gaze models by both HS and LS individuals ($p < 0.05$). The agent frequently changed its gaze states in HG, which might have led to discomfort in the conversation.

6 Discussion

Our previous research showed that perceived shyness of an agent was inversely proportional to the amount of gaze from the agent [12]. In this experiment, the amount of gaze was increased from HG to JG, AG, and FG, and thus, we expected perceived shyness to increase in the inverse order. However, contrary to our expectation, AG was perceived as significantly less shy than were other gaze models. One possible reason is that amount of gaze is not the only factor that affects an agent's perceived shyness, and that gaze patterns are also critical. This result is in line with a report that expression of dominance can be controlled though an agent's gaze patterns [6].

The perceived friendliness, friendly feeling, and comfortableness of HG were significantly lower than they were in the other gaze conditions. HG was implemented by combining JG and AG as a culture-independent model that was neither American nor Japanese. HG had the least amount of gaze compared to the other gaze models, and the agent shifted its gaze patterns more often than it did in the other gaze models. Although interactions in all gaze conditions were regarded as equally natural, the Japanese participants did not feel friendly feeling toward the agent that shifted its gaze frequently with unfamiliar patterns.

FG was perceived as equally friendly and comfortable by both shyness groups (HS and LS). FG was perceived as equally friendly and comfortable as was JG by HS individuals. One possible reason for this result is that the Japanese participants were familiar with agents or computer characters that do not move their eyes and keep staring at them, as many virtual agents are not programmed to shift their gaze.

JG was perceived as equally friendly and comfortable by both shyness group, although there was an overall tendency for LS individuals to form more positive impressions of the agents and interactions as compared to HS individuals. This result is reasonable, as the participants were familiar with JG patterns.

In order to verify whether our hypotheses were supported, we compared perceptions of AG evaluated by both shyness groups. The hypotheses were as follows: H1) Shy people form worse impressions of agents whose gaze model does not originate from the same culture; and H2) Non-shy people are more tolerant of gaze models from other cultures.

HS individuals formed worse impressions of AG than they did of FG and JG, as low as their impressions of HG. This suggests that HS individuals' perceptions and impressions of AG are as unfriendly and uncomfortable as their impressions of HG are. One possible explanation is that this poor impression is caused by unfamiliarity of gaze patterns in AG, in addition to HS individuals' reduced tolerance or adaptation skills for

unfamiliar gaze patterns. HS individuals had similarly positive impression of FG and JG, demonstrating the same tendency as a previous study [12]. HS individuals had more favorable impressions of the agent as the amount of gaze from the agent increased. Other characteristics of HS individuals include their sensitivity to changes in gaze patterns, as their perception of friendliness changed significantly across the four gaze patterns. These results are in accordance with previous findings [12]. HS individuals' evaluations of friendliness and comfort were lower than were those of LS individuals in general. In particular, their impressions of AG were significantly lower than the ones made by LS individuals. These results support H1.

LS individuals showed the opposite impressions of AG. They maintained a stable positive impression of AG, which was as positive as their impression of JG. Their friendly feeling toward the agent in AG was higher than it was for FG. LS individuals' impressions of HG were consistently lowest. These results support H2. The findings suggest that LS individuals are more tolerant to unfamiliar gaze patterns from different cultures because of their social skills.

7 Conclusion

We aimed to investigate cultural differences in gaze behaviors and their perception by developing virtual agents with Japanese gaze behaviors, American gaze behaviors, hybrid gaze behaviors, and full gaze behaviors. We then compared their effects on impressions of and interactions with the agents. Our experimental results with Japanese participants suggest that impressions of the agent are affected by participants' shyness as well as the familiarity of the gaze pattern performed by the agent.

We will continue the experiment with participants from the US and other cultures. The limitations of this study include the agent's appearance, and that most participants were male, with a limited sample size. Our follow-up study should include a female agent and more female participants, with a larger sample size, people with mid-level shyness. Finally, the agent should exhibit non-verbal behaviors other than gaze.

We believe this study will draw more attention to awareness of cultural differences of gaze behaviors, which we usually control unconsciously. One potential application of this outcome is cultural training for the typical gaze behaviors of different cultures in order to facilitate mutual understanding and decrease the likelihood of misunderstandings through misinterpretation of other cultures' gaze behaviors.

Acknowledgements
This research is supported by JSPS KAKENHI JP26330236 and JP17K00287.

References

1. Argyle, M., Cook, M. Gaze and mutual gaze. Cambridge University Press Cambridge, 1976.
2. Bayliss, A.. Paul. M., Cannon, P, and Tipper, S. Gaze cuing and affective judgments of objects: I like what you look at. Psychonomic bulletin & review 13, 6, pp. l06l-1066, 2006.

3. Pelachaud, C., and Bilvi, M. Modelling gaze behavior for conversational agents. ln IVA2003, Springer, pp. 93-100, 2003.
4. Lee, J., Marsella, S., Traum, D., Gratch, J., and Lance, B. The rickel gaze model: A window on the mind of a virtual human. In IVA2007, Springer, pp. 296-303, 2007.
5. Cafaro, A., Gaito, R., and Vilhjrilmsson, H. Animating idle gaze in public places. In IVA2009, Springer, pp. 250-256, 2009.
6. Bee, N., Pollock, C., André, E., Walker, M. Bossy or Wimpy: Expressing Social Dominance by Combining Gaze and Linguistic Behaviors. In IVA2010, Springer, pp 265-271, 2010.
7. Ishii, R, et al. Avatar's gaze control to facilitate conversational turn-taking in virtual-space multi-user voice chat system. In IVA2006, Springer, pp.458-458, 2006.
8. Ishii, R, et al. Avatar's Gaze Control to Facilitate Conversation in Virtual-Space Multi-User Voice Chat System, in J. of Human Interface, 10 (1), pp. 87-94, 2008. (in Japanese)
9. Jones, Warren H., and Dan Russell. The social reticence scale: An objective instrument to measure shyness. Journal of personality assessment 46.6, pp. 629-631, 1982.
10. Cheek, J. M., & Buss, A. H. 1981 Shyness and sociability. Journal of Personality and Social Psychology, 41, pp.330-339, 1981.
11. Daly, S. Behavioral correlates of social anxiety. British Journal of Social and Clinical Psychology, 17, pp. 117-120, 1978.
12. Koda, T., Ogura, M, and Matsui, Y. Shyness Level and Sensitivity to Gaze from Agents. - Are Shy People Sensitive to Agent's Gaze? In IVA 2016, Springer, pp. pp 359-363, 2016.
13. Rehm, M. et al. From Observation to Simulation - Generating Culture Specific Behavior for Interactive Systems. AI & Society 24, Springer, pp. 267–280, 2009.
14. Koda, T. et al. Avatar Culture: Cross-Cultural Evaluations of Avatar Facial Expressions. AI & Society, Springer, pp. 237–250, 2009.
15. Kühne, V. et al. Using linguistic alignment to enhance learning experience with pedagogical agents: the special case of dialect. In IVA2013, Springer, pp. 149-158, 2013.
16. Finkelstein, S. et al. The effects of culturally congruent educational technologies on student achievement. AIED2013, Springer, pp. 493-502, 2013.
17. Johnson, L., et al. Tactical language and culture training systems: Using artificial intelligence to teach foreign languages and cultures. In: Proc of IAAI, pp,73-83, 2008.
18. Aylett, R. et al. But that was in another country: agents and intercultural empathy. In: Proceedings AAMAS, pp. 329–336, 2009
19. Mayo, C., and La France, M. Gaze Direction in Interracial Dyadic Communication. In Harper, R.G. et al.(ed.) Meeting of Eastan Psychological Association. 1978.
20. Elzinga, R.H. Temporal Organization of Conversation, Sociolinguistics Newsletter, 9, 2 (summer), pp.29-31, 1978.
21. Cook, M. and Smith, M.C. The Role of Gaze in Impression Formation, Br. J. Clinical Psychology, Vol.14, pp.19-25, 1975.
22. Fukayama, A. et al.. Messages embedded in gaze of interface agents - impression management with agent's gaze. Proc. of the SIGCHI (CHI '02). ACM, New York, pp.41-48, 2002.
23. Cassell, J., Obed E. T., and Prevost, S. Turn taking versus discourse structure. Machine conversations. Springer US, pp.143-153, 1999.
24. Gambi, C., Staudte, M., and Torsten, J. The role of prosody and gaze in turn-end anticipation. Procs. of the Annual Conference of the Cognitive Science Society. 2015.
25. Hall, E.T. The hidden dimension, Doubleday and Company, 1966.
26. Aikawa, A. A study on the reliability and validity of a scale to measure shyness as a trait. The Japanese Journal of Psychology, 62(3), pp.149-155, 1991. (in Japanese)