



Interaction Initiation with a Museum Guide Robot-From the Lab into the Field

Laura-Dora Daczo, Lucie Kalova, Kresta Bonita, Marc Domenech Lopez,
Matthias Rehm

► To cite this version:

Laura-Dora Daczo, Lucie Kalova, Kresta Bonita, Marc Domenech Lopez, Matthias Rehm. Interaction Initiation with a Museum Guide Robot-From the Lab into the Field. 18th IFIP Conference on Human-Computer Interaction (INTERACT), Aug 2021, Bari, Italy. pp.438-447, 10.1007/978-3-030-85613-7_30 . hal-04292376

HAL Id: hal-04292376

<https://inria.hal.science/hal-04292376>

Submitted on 17 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Interaction Initiation with a Museum Guide Robot — From the Lab into the Field

Laura-Dora Daczo, Lucie Kalova, Kresta Louise F. Bonita, Marc Domenech Lopez and Matthias Rehm

Department of Architecture, Design, and Media Technology
Aalborg University, 9000 Aalborg, Denmark

Abstract. The use of social robots is making its way into public spaces such as schools and museums. In this paper we investigate initiation of interaction between a museum guide robot and visitors. First, we conducted a lab experiment, based on a previous Japanese study, which confirmed that participants prefer the robot with appropriate greeting behavior to the robot without. Thus, we concluded that the results obtained in the original study can be applied in the Danish cultural context. Based on these finding, we conducted a field study in a Danish museum to evaluate whether the same principles apply in the wild and to investigate spatial behavior between robot and visitors in real world settings.

Keywords: Human robot interaction · field study · cross cultural comparison

1 Introduction

Social robots are envisioned to provide services in public spaces such as schools and museums, which means that people will be interacting with robots more often. This has brought researchers to study the different aspects of the robot's behaviour and to find out how humans react to it. For instance, Shiomi et al. [14] studied the interaction between a humanoid robot and visitors in a science museum, and Yousuf et al. [20] studied the best way to start interaction through greeting with a humanoid robot.

In all these studies, mobile robots provide services like education, information, entertainment and assistance to people [11]. However, most studies were not conducted in the wild, even though studying social robots in the real world plays an important role in improving their acceptance. We decided to replicate the study by Yousuf et al. [20], to find out if their findings also apply to the Danish culture. We then verified this in the context of a museum through a field study.

2 Related Work

Human-Robot Interaction Within the field of HRI, multiple studies focus on how humans approach robots [1, 3], and how robots should behave after the initiation

phases of the interaction [15, 19]. However, we found only two studies focusing on how a robot should initiate interaction through greetings. Heenan et al. [5] implemented some of Kendon's [7] greeting behaviours from human human interaction into a 58 cm tall humanoid robot. Based on informal observations of how people reacted to their prototype they found the robot's greeting capability effective, however there's no indication about how they measured effectiveness. Yousuf et al. [20] used a 108 cm tall, humanoid robot with hands and wheels, to study how robots that perform greeting impact people's perception. They compared the robot that greeted people before providing information, with one that did not. The sighting phase consisted of participants looking at the robot. During the distance salutation, the robot positioned itself to form a transactional segment, while in the approach phase the robot decreased the original distance from the participants to a distance of 1 to 1.3 m. Finally, the close salutation consisted of a verbal greeting, "Welcome", without additional gestures. The experimental area was 3.5 m², so the starting distance was within the social distance. The robots were assessed through a questionnaire that used Likert-scale on greeting capability and general impression, participants preferring the robot that performed the interaction initiating behaviors.

Interaction with museum guide robots Other studies focus on exploring how humanoid robots can serve as guides [20, 15, 19, 21, 8]. Shiomi et al. [15] studied a guide robot that served as an advertiser in a shopping mall, while the other studies investigated the interaction in the context of museums. Apart from Yousuf et al. focusing on greetings, the other studies investigated the coordination of verbal and non-verbal interactions as well as spatial formations. Two of these studies modelled the interaction based on observations and video analysis of human guides [19, 21]. From their ethnographic studies, they found that humans automatically form F-formations, and if a visitor's facial or body orientation was not facing the guide or the exhibit, the human guide stopped and restarted sentences in order to re-capture the visitor's attention. Further, they also found that guides frequently turned their heads to look at both the visitors and the exhibit while explaining, and pointed towards the exhibit that was being explained. These non-verbal cues were most often performed at sentence endings. If these patterns were adopted by robots, participants found them more favorable compared to cases where they were missing.

In the wild Except for Yamazaki et al. [19] and Shiomi et al. [15], all experiments mentioned here were conducted in labs. Considering the lack of studies in which museum guide robots were tested in the wild, we explored challenges that can appear in real world settings when interacting with a robot. In in-the-wild studies, people use systems without being explicitly instructed about tasks, so the ecological validity is higher than in case of lab studies. This on the other hand leads to the researchers not being able to control the outcome [9, 16]. According to Marshall et al. [9], these studies are most beneficial early in the development process to gain insights into design challenges. However, Siek et al. [16] mentions that the used prototypes should be at a level where they are robust enough to be

used *in situ*. Not being highly familiar with the environment and the participants can often lead to unexpected situations that are very specific to the study and that the researchers cannot prepare for based on other studies and literature. This can lead to situations in which actions are taken based on the environment and not on the needs of the researchers [4]. This can eventually lead to further findings about aspects that need to be improved to achieve a prototype that is robust enough to function in the wild [17, 2]. While doing experiments in the wild Williamson and Williamson and Mueller et al. [18, 10] found that it can take a long time to find people who are willing to interact with their artifacts placed at public places, as, unlike lab experiments, these sessions are not scheduled. Therefore, it is important to plan ahead when conducting these studies, keeping in mind when there would be people willing to interact with a given artifact [16]. Furthermore, permission to conduct or, on occasion, document these studies should also be obtained [9]. While not everything can be planned, the prototype should be usable and the researchers should be prepared to encounter unexpected situations.

3 Lab Experiment

This experiment was a replication of the study by Yousuf et al. [20], conducted a) to verify their results; b) to explore if there are any culture-related effects; and c.) to explore whether using a different robot (we use a Pepper robot from SoftBank Robotics) influences the results. The experiment was conducted in a lab, where we marked an area of 3.5 m^2 , within which the participants had to stay, just as in the original experiment. We placed a drawing at eye level in one corner of the area, while the robots starting position was in the opposite corner.

16 Danish two person groups participated (8 female, 24 male), both graduate and undergraduate students recruited at Aalborg University Create campus. One of the participants had previous experience with humanoid robots. The experiment was counterbalanced and consisted of two separate conditions, condition A (experimental condition) and condition B (control condition), with within-subject and post-test only design. Participants filled out a set of two questionnaires. The original questionnaire consisted of three questions on a 7-point Likert scale: (1) Do you think the robot was able to greet you properly?; (2) Do you think that the robot attended to you adequately?; (3) What is your overall evaluation of the robot? Additionally, participants filled out a *Godspeed Questionnaire Series* (GQS) for additional data, after each condition. The experiment took approximately 8-10 min to complete including filling in the questionnaires.

In the experimental condition (A), the guide robot approached the person looking at it and said: Welcome, may I explain the exhibit to you?. If the participant replied Yes, the robot moved towards the exhibit and started the explanation. In the case that neither of the participants was looking at the robot in the beginning of the experiment and they were both looking at the exhibit instead, or anywhere else in the room, the guide robot approached closer, eventually moving into the participants field of view to get their attention. With at

least one of the participants gazing at the robot, the experiment continued in the same way as in the case mentioned before. In the control condition (B) the robot moved directly to the painting and started the explanation immediately without any greeting.

3.1 Results

Original questionnaire In a Wilcoxon signed ranked test, we discovered that the participants found the robot’s greeting capability significantly higher in condition A (Mdn=5) compared to condition B (Mdn=2.5) $T = 16.5, p < 0.01, r = -0.55$. In question 2, the Wilcoxon signed ranked test revealed that the participants found robots attendance in condition A (Mdn=5) significantly more adequate compared to the condition B (Mdn=4), $T = 18, p < 0.01, r = -0.45$. Finally, the participants evaluated the robot significantly higher in condition A (Mdn=5) compared to condition B (Mdn=4), $T = 24, p = 0.01, r = -0.45$.

Godspeed questionnaire On average, the participants rated the robot in condition A higher in all of the categories. Each of these categories were then individually compared using a Wilcoxon signed rank test. The results revealed that the anthropomorphism of the robot in condition A (Mdn=3) was significantly higher, than the anthropomorphism of the robot in condition B (Mdn=3) $T = 268, p < 0.01, r = -0.69$. The animacy lead to similar results, condition A (Mdn=3) being significantly more animous than than B (Mdn=3) $T = 1953, p < 0.01, r = -0.81$. We also found that the likeability (Mdn=4) in condition A was significantly higher than of condition B (Mdn=4), $T = 739, p < 0.01, r = -0.52$. The perceived intelligence (Mdn=4) was also significantly higher than in condition B (Mdn=3), $T=2654, p < 0.01, r = -0.45$. Finally, there was no significant difference in perceived safety (Mdn=3, Mdn=3), $p = 0.23, r = -0.15$.

3.2 Discussion

In the original study, participants rated the robot significantly higher in all three questions in condition A. We obtained the same results, which also seem to be supported by the answers to the GQS, where we got significantly higher results in four out of the five categories. Looking at the categories of the GQS individually, we can see that the Likeability results are aligned with the findings from the follow-up comments. Further, people attributed higher intelligence to the robot that was greeting and asking for permission to explain the exhibit, despite the fact that in both conditions the robot was exhibiting a pre-programmed behaviour and was unable to interact freely with the participants.

Based on these findings we can conclude that starting the interaction by greeting is appreciated in Denmark. Having similar results to original study it is difficult to conclude how the differences in the robots influenced greetings. Overall, achieving significant preference for the experimental condition in both our and the original study suggests that these results are strong enough to conduct an experiment in the wild with a robot that starts interacting with people by greeting them.

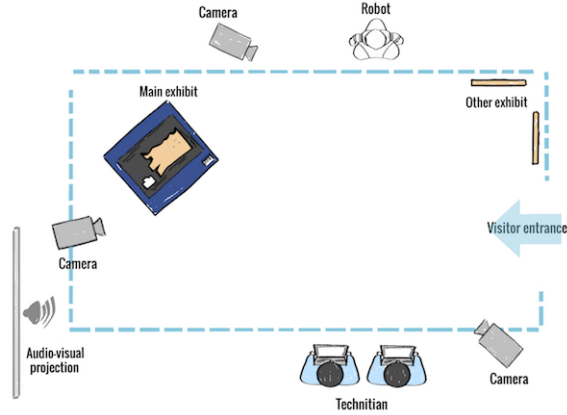


Fig. 1. The setup for the field study inside the museum.

4 Field Study

Following the success of the lab study, we decided to test the robot in a real museum. The purpose of the study was to evaluate the performance of our robot as a guide and to see how visitors in the context of a real museum interact with it. For data collection, we used the qualitative method of video analysis to observe the interaction, and interview to gather user responses from participants.

The second experiment was conducted at a local museum, the Utzon Center in Aalborg, Denmark. In the museum, we occupied an area where a wooden model of the Romsdal museum was used as the main exhibit for this field study experiment (see Fig. 1). We set up three video cameras at different angles to document the whole experiment.

Seeing that Williamson and Williamson [18] and Mueller et al.[10] warned about how difficult it can be to find participants in public places, we wanted to make sure we could recruit enough people willing to participate in the experiment. Thus, we chose the weekend day after a new exhibit was opened, when the Utzon Center usually has the most visitors. There was a total of 26 participants, 15 female and 11 male. Three people participated alone, while there were eight groups of two, one group of three and one group of four participants in the experiment. Two of the participants had previous experience with robots in the form of robot vacuum cleaners. All participants were Western European, two of them native English speakers from the UK and 24 Danish with various levels of knowledge of the English language.

Three researchers were present during each test i.e. one facilitator and two technicians. The facilitator lead and explained each test while the two technicians operated the robot manually to avoid any unpredictable problems that might arise when using the fully automatic behaviour. Each test started with the facilitator verbally explaining the procedure and showing the participants



Fig. 2. Left: L-shaped F-formation; Center: Horseshoe arrangement; Right: Horseshoe arrangement with Pepper.

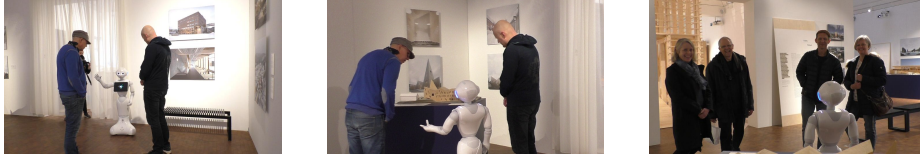


Fig. 3. Left: Circular F-formation with Pepper; Center: Participants leaning in to hear the robot better; Right: Participants smiling when interacting with the robot.

to the designated experiment area. Then, the robot would approach the group, greet them and ask if it could explain the Romsdal Folk model. Shortly after, the robot would ask participants to follow it towards the model to explain the specific exhibit. The entire test took 2 to 5 minutes to complete for each group, followed by a 3 minutes follow-up group interview.

4.1 Findings

Observation The videos were analyzed using Interaction Analysis, following a subset of the categories defined by [6]. These were turn-taking, spatial organization and trouble and repair. In the turn-taking theme, we observed 12 instances of groups of participants waiting with speaking or walking after robot has done so first. We also noticed that 5 participants were replying to the robot even when it was not needed for the interaction and were polite, saying things like "Yes, please" even though politeness does not matter when interacting with a robot.

With spatial organization we focused on the F-formation, looking into which one participants tend to use and when. We uncovered that every time a participant was interacting with Pepper alone they were in a vis-a-vis F-formation in the beginning, usually after they greeted each other and before Pepper lead them to the exhibit. After that, they regrouped into an L-shaped F-formation when the robot was explaining the exhibit (Fig. 2 left). In groups with several participants (2 to 4) there was more diversity in where people stood during the experiment. Most popular was the side-by-side F-formation, where 4 groups of 2 people chose to stand next to each other, opposite to Pepper, both in the greeting and the explanation phase. Furthermore, we observed 4 horseshoe arrangement cases, 2 of which consisted of people in this formation facing Pepper (Fig. 2 center) and 2 groups that included Pepper in this arrangement (Fig. 2

right). In two cases groups of 2 people formed a circular F-formation together with Pepper, once during the explanation phase and once during greeting (Fig. 3 left).

Analysing the videos through the trouble and repair theme brought to our attention that when the robot suddenly stopped moving during 3 trials, the participants became uncomfortable and confused and waited for the robot to start moving again. We also confirmed that the background noise, which was present during the testing, was intrusive and made it harder for the participants to hear and understand the robot, as two participants leaned in when it was talking to hear better (Fig. 3 center) and two people mentioned it was difficult to understand the robot. In general, we observed that participants in 5 groups were smiling and in 4 laughing when interacting with the robot, especially when they were in a group (Fig. 3 right).

Follow-up Interviews We analysed the follow-up interviews by six topics.

Main impressions of the experience Participants main impressions from the experiment were mostly positive, 6 people thinking it was fun, 2 that it was fine, 1 saying that it was a special experience and 1 calling it a positive experience. On the other hand, it was also apparent that none of the participants have ever interacted with a humanoid robot before, as 4 of them mentioned it was strange, 4 said it was a new experience and 2 participants felt it was surreal.

Participants feelings and preoccupations during the experiment During the experiment, 6 participants felt confused, because they were not sure what they should say or where the robot would lead them. Nevertheless, 4 participants felt curious and 1 was fascinated by the robot's movement and gestures.

Participants overall impressions of the robot The participants described the robot with positive attributes. 3 of them regarded it polite, 2 found it friendly and 2 others found it useful. There were 2 participants who described the robot as cute.

The robots voice However, most participants had problems with the robots voice, 10 of them mentioning that it was difficult to understand what it was saying, 3 participants explaining that the problem was the loud background voice coming from the neighbouring exhibit. Furthermore, 3 people thought the robot had a monotonous voice, while 3 others said it was mechanical and 2 regarded its voice high-pitched.

Further problems Other problems people had with robot were its speed, 4 participants saying that it moved too slowly, while 2 complained about the lack of open interaction and 1 about the lack of eye contact.

4.2 Discussion

We observed that people anthropomorphized the robot by treating it as they would be treating a human guide, being polite and answering its greetings. During the interviews they also referred to it with terms that are commonly used to describe people. On the other hand, they were fully aware that it was impossible to have the same kind of interaction with the robot as they have with each other e.g. impossible to have an open conversation. These findings are aligned with Reeves' media equation [12] and have also been confirmed by other studies [13].

Conducting the study in the wild resulted in the benefit of uncovering environmental influences on the efficiency of the robot. During the lab test people understood what the robot explained but in the museum the presence of another synthetic voice made it difficult to follow what the robot was saying. While people were able to understand each other, the similarity of the robot's voice to the sound coming from other exhibits made it difficult to understand. Besides human voices being easier to understand, people also mentioned preferring a human voice. However, even though this would be both preferred and easier to understand it would not only break the design by not conforming to expectations but also limit the verbal interaction to predefined speech. Additionally, changes in pitch, intonation and timbre could be made.

In terms of formations, there was no overall preference for a specific type of F-formation. As Kendon found in human interactions, the formations created with the robot were also dependent on the number of participants and the interaction phase. This suggests that during HRI both the number of participants and the specifics of the interaction should be considered before aiming to establish a formation. Due to the fact that participants sometimes excluded the exhibited model from their formations, focusing only on the robot, the robot should actively aim to draw attention to the exhibit. Yamazaki et al. [19], observed that gesturing towards the exhibit was often used during explanations. In our study we observed that this and the exhibit made the participants look at it, therefore these behaviours should be employed more frequently.

In general it can be noticed that the aspects that participants complained about were different from how humans interact with each other e.g. lack of possibility for free interaction or synthetic voices. These aspects should be improved to achieve an interaction that is more natural for the participants.

5 Conclusion

In this paper we have proposed a museum guide robot able to initiate interaction with museum visitors, using human human interaction as reference. We have conducted an experiment based on Yousuf et al.'s paper [20] and confirmed that their findings are correct and also apply to Denmark. Seeing that there was a lack of studies investigating HRI in the wild, our second step was to conduct a field study in a real museum setting. We can conclude that greetings are a good way of initiating interaction between a museum visitor and a museum guide robot in

Denmark. Additionally, our observations suggested that aspects of interaction that were different from human human interaction, such as the speed of the robot and its voice, were not appreciated. Therefore, these aspects should be further investigated. Another area that could be investigated in the future is whether these findings apply in other contexts, such as a library or a shopping mall.

References

1. Aaltonen, I., Arvola, A., Heikkilä, P., Lammi, H.: Hello pepper, may i tickle you? children's and adults' responses to an entertainment robot at a shopping mall. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. p. 5354. Association for Computing Machinery, New York, NY, USA (2017). <https://doi.org/10.1145/3029798.3038362>
2. Behrens, M., Valkanova, N., gen. Schieck, A.F., Brumby, D.P.: Smart citizen sentiment dashboard: A case study into media architectural interfaces. In: Proceedings of The International Symposium on Pervasive Displays. p. 1924. Association for Computing Machinery, New York, NY, USA (2014). <https://doi.org/10.1145/2611009.2611036>
3. Bergstrom, N., Kanda, T., Miyashita, T., Ishiguro, H., Hagita, N.: Modeling of natural human-robot encounters. In: 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 2623–2629 (2008). <https://doi.org/10.1109/IROS.2008.4650896>
4. Boyd, L.E., Rector, K., Profita, H., Stangl, A.J., Zolyomi, A., Kane, S.K., Hayes, G.R.: Understanding the role fluidity of stakeholders during assistive technology research "in the wild". In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. p. 61476158. Association for Computing Machinery, New York, NY, USA (2017). <https://doi.org/10.1145/3025453.3025493>
5. Heenan, B., Greenberg, S., Aghel-Manesh, S., Sharlin, E.: Designing social greetings in human robot interaction. In: Proceedings of the 2014 Conference on Designing Interactive Systems. p. 855864. Association for Computing Machinery, New York, NY, USA (2014). <https://doi.org/10.1145/2598510.2598513>
6. Jordan, B., Henderson, A.: Interaction analysis: Foundations and practice. *Journal of the Learning Sciences* 4(1), 39–103 (1995). https://doi.org/10.1207/s15327809jls0401_2
7. Kendon, A.: Conducting interaction: Patterns of behavior in focused encounters. Cambridge University Press, New York, NY, US (1990)
8. Kobayashi, Y., Shibata, T., Hoshi, Y., Kuno, Y., Okada, M., Yamazaki, K.: Choosing answerers by observing gaze responses for museum guide robots. In: 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 109–110 (2010). <https://doi.org/10.1109/HRI.2010.5453240>
9. Marshall, P., Rogers, Y., Pantidi, N.: Using f-formations to analyse spatial patterns of interaction in physical environments. In: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work. p. 445454. Association for Computing Machinery, New York, NY, USA (2011). <https://doi.org/10.1145/1958824.1958893>
10. Müller, J., Walter, R., Bailly, G., Nischt, M., Alt, F.: Looking glass: A field study on noticing interactivity of a shop window. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. p. 297306. Association for Computing Machinery, New York, NY, USA (2012). <https://doi.org/10.1145/2207676.2207718>

11. Pateraki, M., Trahanias, P.: Deployment of Robotic Guides in Museum Contexts, pp. 449–472. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-49607-8_18
12. Reeves, B., Nass, C.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. CSLI Publications (1996)
13. Rehm, M.: She is just stupid: Analyzing user-agent interactions in emotional game situations. *Interacting with Computers* **20**(3), 311–325 (2008). <https://doi.org/https://doi.org/10.1016/j.intcom.2008.02.005>, special Issue: On the Abuse and Misuse of Social Agents
14. Shiomi, M., Kanda, T., Ishiguro, H., Hagita, N.: Interactive humanoid robots for a science museum. In: *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. p. 305312. Association for Computing Machinery, New York, NY, USA (2006). <https://doi.org/10.1145/1121241.1121293>
15. Shiomi, M., Kanda, T., Ishiguro, H., Hagita, N.: A larger audience, please! encouraging people to listen to a guide robot. In: *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*. p. 3138. IEEE Press (2010)
16. Siek, K.A., Hayes, G.R., Newman, M.W., Tang, J.C.: *Field Deployments: Knowing from Using in Context*, pp. 119–142. Springer New York, New York, NY (2014). https://doi.org/10.1007/978-1-4939-0378-8_6
17. Vaizman, Y., Ellis, K., Lanckriet, G., Weibel, N.: Extrasensory app: Data collection in-the-wild with rich user interface to self-report behavior. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. p. 112. Association for Computing Machinery, New York, NY, USA (2018). <https://doi.org/10.1145/3173574.3174128>
18. Williamson, J.R., Williamson, J.: Understanding public evaluation: Quantifying experimenter intervention. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. p. 34143425. Association for Computing Machinery, New York, NY, USA (2017). <https://doi.org/10.1145/3025453.3025598>
19. Yamazaki, A., Yamazaki, K., Burdelski, M., Kuno, Y., Fukushima, M.: Coordination of verbal and non-verbal actions in humanrobot interaction at museums and exhibitions. *Journal of Pragmatics* **42**, 2398–2414 (2010). <https://doi.org/https://doi.org/10.1016/j.pragma.2009.12.023>
20. Yousuf, M.A., Kobayashi, Y., Kuno, Y., Yamazaki, A., Yamazaki, K.: How to move towards visitors: A model for museum guide robots to initiate conversation. In: *2013 IEEE RO-MAN*. pp. 587–592 (2013). <https://doi.org/10.1109/ROMAN.2013.6628543>
21. Yousuf, M.A., Kobayashi, Y., Kuno, Y., Yamazaki, A., Yamazaki, K.: Development of a mobile museum guide robot that can configure spatial formation with visitors. In: Huang, D.S., Jiang, C., Bevilacqua, V., Figueroa, J.C. (eds.) *Intelligent Computing Technology*. pp. 423–432. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)