You Can Leave Your Head On

Attention Management and Turn-Taking in Multi-party Interaction with a Virtual Human/Robot Duo

Jeroen Linssen, Meike Berkhoff, Max Bode, Eduard Rens, ✉Mariët Theune, and Daan Wiltenburg

Human Media Interaction
University of Twente
Drienerlolaan 5, 7522 NB Enschede
The Netherlands
J.M.Linssen@utwente.nl, M.Theune@utwente.nl

Abstract. In two small studies, we investigated how a virtual human/ robot duo can complement each other in joint interaction with one or more users. The robot takes care of turn management while the virtual human draws attention to the robot. Our results show that having the virtual human address the robot, highlights the latter's role in the interaction. Having the robot nonverbally indicate the intended addressee of a question asked by the virtual human proved successful in all cases when the robot was first addressed by the virtual human.

1 Introduction

In this paper, we investigate attention management and turn-taking in interactions with R3D3: the Rolling Receptionist Robot with Double Dutch Dialogue, which is intended to serve as an assistant to visitors of public places. R3D3 consists of a robot and a virtual human, which is carried on a tablet by the robot (see Fig. 1a). The virtual human can interact with people using Dutch spoken language. The robot does not speak, but takes on a supportive role by providing nonverbal cues with its eyes and head. However, our initial experiences with R3D3 showed that when talking to the virtual human, users tended to ignore the robot, putting the robot's added value for the conversation into question [3].

Our main research question for this paper is therefore how we can give the robot a clear role in the interaction. Specifically, we investigate how the virtual human can draw attention to the robot and how the robot can manage turn taking in multi-party conversation. Turn-taking, between humans [6] as well as between humans and both virtual characters [1] and social robots [2], is seen as an important factor in managing fluent interactions. Especially in crowded places, interactions between robots and multiple users can benefit from turn management through gaze, both for enabling a robot to express its intentions and controlling users' attention [2, 7].

Below we describe two small studies we carried out to investigate attention management and turn-taking with our virtual human/robot duo.

2 Study 1: Attention Management

The first study involved interactions between the virtual human (VH), the robot and a single user. The VH talked to the user about the research topics of our department, while the robot nonverbally supported the VH's utterances, nodding in confirmation and showing emotions in line with what the VH said. We used a Wizard-of-Oz setup, with a hidden experimenter controlling R3D3's behaviour.

The interactions took place in one of two conditions. In Condition 1 (C1), the robot nonverbally reacted to the participant's answers before the VH replied. In Condition 2 (C2), the VH explicitly addressed or referred to the robot before the robot showed its nonverbal behaviour. This is illustrated in Table 1. In each condition, five participants (students aged 18 to 25 years) interacted with R3D3.

Table 1. Excerpts from the interaction of Study 1 (translated from Dutch), with the difference between the conditions shown underneath the line.

VH: *If you had to choose, which of the following research topics would you like to hear more about: [...]?*

Participant mentions one of the topics.

C1	Robot nods enthusiastically. VH: *Good choice, I think that is very interesting too.*	C2	VH: *Good choice, Robot and I think that is very interesting too.* Robot nods enthusiastically.

Analysis of the participants' gaze behaviour showed that although they did not pay less attention to the robot in C2 compared to C1, their attention was better timed. In C2, the participants tended to gaze at the robot after it had been mentioned or addressed by the VH, and also at pauses in the speech of the VH. The latter suggests that they saw the robot as a side-participant in the conversation from whom they expected backchannelling behaviour at the appropriate places (see, e.g., [8]). In C1, the participants tended to look at the robot only after it had already started performing its nonverbal behaviour. They may have been alerted to this by the slight noise caused by the robot's head movements.

3 Study 2: Turn Allocation

In the second study we gave a bigger role to the robot by having it use gaze and head movement to allocate turns in interactions with two users. The intentional direction of gaze [5] has been shown to be a highly effective turn-taking mechanism in human-robot interaction [4].

As in Study 1, we used a between-subjects design with two conditions. In each condition, five pairs of participants (students aged 18-25 years) had a short conversation with R3D3, which was again controlled by a wizard. During each interaction, one of the participants was given a cap to wear; see Fig. 1b.
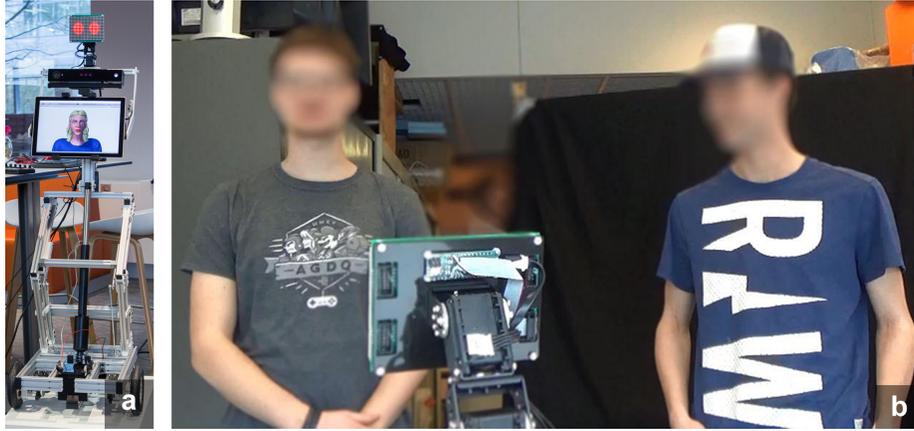
Fig. 1. (a) R3D3 at the time of the studies. (b) The setup of Study 2, with two participants standing in front of R3D3. The robot gazes at the person without the cap.

Table 2 shows the interaction scenarios for the two conditions. In both conditions, the final question was meant to be answered by Participant 2, the one without the cap. This addressee could not be derived from the VH's utterance. Instead, the robot's gaze was used to disambiguate the addressee, while the gaze direction of the VH remained neutral. In C2 but not in C1, the VH explicitly addressed the robot to draw the participants' attention to it before posing the final question.

Table 2. The interaction scenarios of Study 2 (translated from Dutch), excluding closing sequences. The difference between conditions is shown between the two lines.

Robot gazes at Participant 1 (with cap)

VH: *Hello, does the cap fit well?*

Participant 1 responds.

VH: *It looks great on you.*

C1	Robot gazes at Participant 2 (without cap). VH: *What do you think?*	C2	VH looks up at robot (called EyePi). VH: *What do you think, EyePi?* Robot nods and gazes at Participant 2. VH: *Would you agree?*

Either Participant 1 or Participant 2 (the intended addressee) responds.

In C1, in which the robot was not explicitly addressed by the VH, the intended addressee (Participant 2) responded only two out of five times. Participant 1 responded the other three times. The two participants who correctly took the turn

in C1 clearly responded when they saw the robot head turning in their direction. The others did not look at the robot when the question was asked (or slightly after). In C2, in which the robot was addressed before gazing at Participant 2, this participant responded in all five cases. This confirmed our expectation that using the robot's gaze for turn allocation would be more effective in combination with explicit addressing of the robot by the VH (C2) than without (C1).

4 Conclusion

We conducted two studies to investigate turn management with R3D3, a virtual human/robot duo. We found that by having the virtual human address or refer to the robot, the users' attention could be drawn to the robot as a participant in the conversation. We also found that this helped the robot to assign turns in interactions with multiple users. Although too small scale to draw strong conclusions, our studies suggest that the robot head can be used successfully to complement the behaviour of the virtual human, if proper attention is drawn to it. For the ongoing development of R3D3, this means it can leave its head on.

Acknowledgements

References

1. Bohus, D., Horvitz, E.: Models for multiparty engagement in open-world dialog. In: Proceedings of SIGDIAL '09. pp. 225–234 (2009)
2. Leite, I., Hajishirzi, H., Andrist, S., Lehman, J.: Managing chaos: Models of turn-taking in character-multichild interactions. Proceedings of ICMI '13 pp. 43–50 (2013)
3. Linssen, J., Theune, M.: R3D3: the Rolling Receptionist Robot with Double Dutch Dialogue. In: Proceedings of the Companion of HRI '17. pp. 189–190 (2017)
4. Mutlu, B., Kanda, T., Forlizzi, J., Hodgins, J., Ishiguro, H.: Conversational gaze mechanisms for humanlike robots. ACM Transactions on Interactive Intelligent Systems 1(2), 1–33 (2012)
5. Ruhland, K., Peters, C.E., Andrist, S., Badler, J.B., Badler, N.I., Gleicher, M., Mutlu, B., McDonnell, R.: A review of eye gaze in virtual agents, social robotics and HCI: Behaviour generation, user interaction and perception. Computer Graphics Forum 34(6), 299–326 (2015)
6. Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn taking for conversation. Language 50(4), 696–735 (1974)
7. Shiomi, M., Kanda, T., Koizumi, S., Ishiguro, H., Hagita, N.: Group attention control for communication robots with wizard of oz approach. In: Proceedings of HRI '07. p. 121 (2007)
8. Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M., Kuzuoka, H.: Precision timing in human-robot interaction. In: Proceedings of CHI '08. pp. 131–139 (2008)