It Gets Worse before it Gets Better:
Timing of Instructions in Close Human-Robot Collaboration

Lars C. Jensen, Kerstin Fischer

Department of Design and Communication
University of Southern Denmark

Dadhichi Shukla

Institute of Computer Science
University of Innsbruck

Franziska Kirstein

Blue Ocean Robotics
Odense, Denmark

Özgur Erkent, Justus Piater

Institute of Computer Science
University of Innsbruck

Summary

In an empirical study (n=36) we investigate mutual adaptation processes in repeated tasks during human-robot collaboration. In the joint action scenario under consideration, human and robot have to fulfill the same task four times in order to complete the whole task, and we study the ways in which the human users adjust to the robot. However, while interactions generally become more fluent in each repetition, most interactions exhibit a novel problem in the second task, which disappears in further iterations. We argue that this problem indicates that human users take interactional achievement for granted, which characterizes human, but not human-robot interaction.

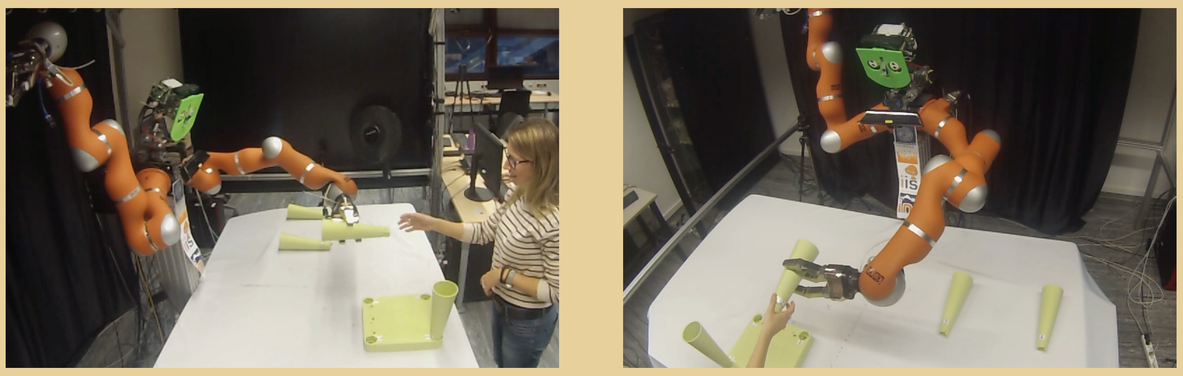


Figure 1: Participant interacting with the robot

Introduction

In human interaction:

- interaction partners adjust to each other over the course of interactions, and develop **shared representations** [1,2].
- Human interaction is characterized by considerable **interactional coordination** [3,4,5].
- In general, interaction partners respond to each other in a time frame of about **300-500 milliseconds**, which requires the successful **prediction** of next action [3,6].
- Thus, in **iterative tasks**, with the next action becoming more predictable, interactions between humans become increasingly **fluent and tightly coupled**.

In human-robot interaction:

- people have been found to **adjust** to robotic communication partners **over time** [7].

Based on these findings, it can be expected that human-robot interactions become **increasingly coordinated**; that an interaction that is running smoothly should become **less smooth in a repetition** is thus rather unexpected.

Method and Data

The Task:
The participants' task was to instruct the robot to fetch the legs of a stool, while the participants themselves had to perform the actual assembly of the stool. On average, the interactions lasted for about **5** minutes.

The Robot:
The robot comprises two **KuKa** arms, each equipped with a **Schunk 3-finger** gripper. However, for this study the robot made only use of its left arm, and a **KIT head**. The robot acted **semi-autonomously** during the experiments, needing only a confirmation for the planner to execute.

The Participants:
36 students and staff from the University of Innsbruck (age range 19-39, 11 female, 25 male)

The Method:
The method for the analysis is **ethnomethodological conversation analysis** [3], which proceeds sequentially by reconstructing each **participant's interpretation** of the respective partner's turn;the underlying assumption is that people in interaction need to **signal** to each other **constantly** how they **understand** each others' actions, which then provides a methodological resource for the analyst.

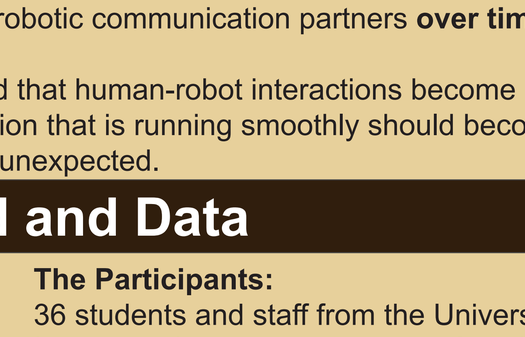
Analysis

First Handover
1. Robot: lifts arm with stool leg
2. **900 milliseconds pause**
3. Human: reaches for the stool leg

Second Handover
4. Robot: lifts arm with stool leg
5. **5.9 seconds pause**
6. Human: holds out hand
(waiting for robot)

Third Handover
7. Robot: lifts arm with stool leg
8. **1.7 seconds pause**
9. Human: holds out hand
(waiting for robot)

Fourth Handover
10. Robot: lifts arm with stool leg
11. **1.7 seconds pause**
12. Human: holds out hand
(waiting for robot)



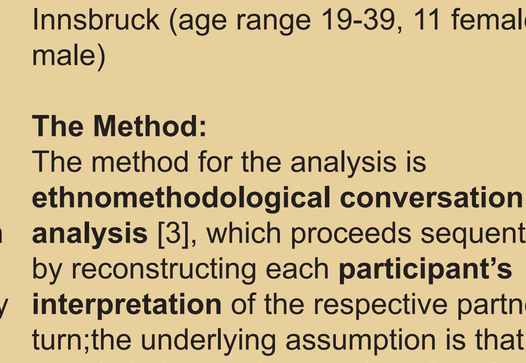


Figure 2: Example 1

Results

Participants indeed **adjusted to the robot over time**

- They became **increasingly savvy** about how to interact with the robot best.
- They are **less fluent in the second execution** of the task than they are in the first.
  - Initially, when the robot stops after it has lifted the first leg, participants initiate the next action after a **short** delay.
  - However, in round 2, they **hesitate even longer**, indicating that they **expect** the robot to carry out its task **autonomously.**

Thus, participants assume that the **robot understands** that the current task is a **repetition of the previous** one and that it has successfully **learned from the previous interaction** what the next step will be, namely to hand over the leg after it has picked it up, **without being explicitly signalled** to do so again.

Discussion & Implications

- Human-robot collaborations **do not simply become more fluent over time**, as previous work would suggest;
  - instead, people's **expectations** that the robot will **build on previous interactions** results in **longer response times** and hence **less fluent** interaction.

- While robot designers may try to implement **all human behavior** into robots, it may actually suffice to implement **those that people expect** the robot to have.
  - The current study has identied one such **expectation that should be considered** in robot design **for all repetitive collaborative tasks**

- The fact that almost all participants in our interactions make this error indicates that this is **an expectation** that may need to be accounted for in human-robot interactions (see also [8]).

- Participants **recover** from this erroneous assumption relatively quickly - within 5.9 seconds in the example.
  - This suggests that while people may carry **expectations** into human-robot interactions, they also **recover quickly** from the violation of these expectations.
  - This finding has also consequences for concepts like the 'uncanny valley', which might be overcome very quickly in interaction (cf. [9]).

References

1. Herbert H. Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. Cognition, 22:1-39, 1986.
2. Gregory Mills. Dialogue in joint activity: Complementarity, convergence and conventionalization. New Ideas in Psychology, 32:158-173, 2014.
3. Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn-taking for conversation. language, pages 696-735, 1974.
4. Herbert H. Clark. Using Language. Cambridge: Cambridge University Press, 1996.
5. Herbert H. Clark and Meredith A. Krych. Speaking while monitoring addressees for understanding. Journal of Memory and Language, 50(1):62-81, 2004.
6. Gail Jefferson. Glossary of transcript symbols with an introduction. Pragmatics and Beyond New Series,125:13-34, 2004.
7. Kerstin Fischer. Designing Speech for a Recipient. Amsterdam: John Benjamins,2016.
8. Raquel Fernandez, Stafan Larsson, Robin Cooper, Jonathan Ginzburg, David Schlangen, et al. Reciprocal learning via dialogue interaction: Challenges and prospects.2011.
9. Jennifer Robertson. Cyborg able-ism: Critical insights from the not so "uncanny valley" of Japan. Plenary Lecture, Robophilosophy'16.

Acknowledgements

This research was partially funded by the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 610878, 3rdHAND.