

Prototyping a User Interface for Multi-Robot Speech Control

Alexandra Bejarano*
abejarano@mines.edu
Colorado School of Mines
Golden, CO, USA

Terran Mott
Colorado School of Mines
Golden, CO, USA
tmott@mines.edu

Sebastian Negrete-Alamillo*
snegretealamillo@mines.edu
Colorado School of Mines
Golden, CO, USA

Tom Williams
Colorado School of Mines
Golden, CO, USA
twilliams@mines.edu

ABSTRACT

The Wizard-of-Oz (WoZ) method is a common and beneficial means of enabling researchers with control over robots in experimental settings. However, there is a lack of general robot control tools for WoZ that are publicly available and easily adaptable to various research domains and needs. In particular, existing control interfaces that may be used in WoZ experiments often do not support the control of multi-robot interactions. As such, in this work, we present the design of three prototypes for a multi-robot speech control interface that would enable the control of multi-robot dialogue interactions.

CCS CONCEPTS

• **Computer systems organization** → **Robotics**; • **Human-centered computing** → **User studies**; **Natural language interfaces**.

KEYWORDS

Wizard-of-Oz, robot control, interface design, multi-robot interaction

ACM Reference Format:

Alexandra Bejarano, Sebastian Negrete-Alamillo, Terran Mott, and Tom Williams. 2024. Prototyping a User Interface for Multi-Robot Speech Control. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion)*, March 11–14, 2024, Boulder, CO, USA. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3610978.3640703>

1 MOTIVATION

In Human-Robot Interaction (HRI), the Wizard-of-Oz (WoZ) method is often used by researchers to quickly and remotely control robot actions within experiments, typically while concealing the involvement of a human controlling the robot [1, 13]. One particular benefit of WoZ is that it provides HRI researchers with a means of robot control that enables the evaluation of robot designs and interactions without needing to fully implement functional systems and interaction methods. As such, WoZ provides an easy way of exploring

*Both authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution International 4.0 License.

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0323-2/24/03.
<https://doi.org/10.1145/3610978.3640703>

different dimensions of HRI such as testing robot behaviors and assessing human perceptions of and interactions with robots.

However, there is a lack of general robot control tools for WoZ that are publicly available and easily adaptable to various research domains and needs. In particular, current control interfaces are particularly underpowered in the area of multi-robot control. For instance, for robots like Misty¹ and Stretch², out-of-the-box control interfaces are limited to only controlling the particular capabilities of the robot they come with and only a single robot at a time can be controlled. As such, researchers often must either build custom multi-robot control interfaces, or control each robot independently, resulting in synchronization challenges.

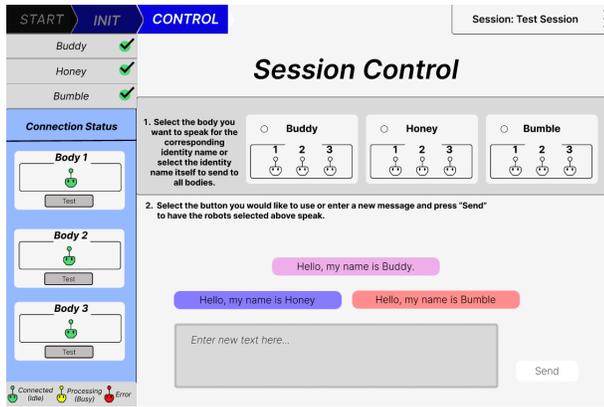
Due to the lack of accessible multi-robot control tools available to researchers, we present prototypes of a user interface for multi-robot speech control. Our user interface includes tools to organize and initialize multi-robot dialogue interactions that account for potentially different configurations of robot identity. Additionally, our prototype interface includes a variety of features that support researchers in preparing dialog ahead of time, as well as adapting new dialog on-the-fly during an interaction. As such, our prototype makes progress towards the development of accessible, generalizable research tools for WoZ research in multi-robot interactions.

2 RELATED WORK

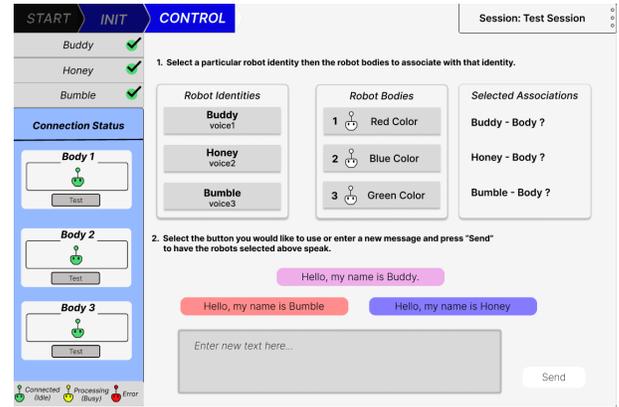
The WoZ method is commonly used in HRI to remotely control robot capabilities without necessitating the implementation of complete or autonomous systems [4, 13]. As such, this method has been used in a variety of HRI research such as to assess how humans may behave and perceive robots [8, 11]. To support such research in HRI, many robot control interfaces have been developed by the research community itself to address specific research needs (e.g. [6, 14, 17, 18]). Additionally, most robots used in HRI, such as the Misty and Stretch robot, come with a means of controlling those particular robots' capabilities through a user interface. However, most of these interfaces focus on and are designed to only enable single robot control, not accounting for HRI research consisting of interactions with multiple robots. There is especially a need for multi-robot control interfaces as there is an increasing focus on non-dyadic interactions in HRI [15]. As such, in this paper, we focus on the design of a multi-robot control interface that would enable to control of multiple robots in an interaction.

¹mistyrobotics.com

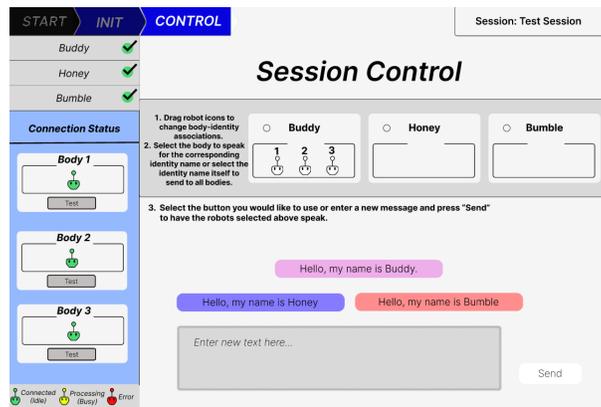
²hello-robot.com



(a) Prototype V1



(b) Prototype V2



(c) Prototype V3

Figure 1: Prototypes for a multi-robot speech control interface

Several multi-robot WoZ control interfaces have been developed (e.g. [5, 10, 16]), yet no general purpose and domain-adaptable multi-robot control tools have been made publicly available. Thus, our end goal is to create a multi-robot control interface that may be easily adapted for different HRI experiments. To simplify our initial interface design, we in particular focus on the control of robot speech as that is a common means of interaction used in HRI.

Furthermore, existing multi-robot control interfaces do not account for the added complexity of robot identity (performed persona) that may arise in multi-robot interactions. Robot identity can easily be manipulated through a variety of observable such as through robot speech to change how robots may be understood and interacted with [2, 7]. Groups of robot in particular do not have to maintain a static association between a robot’s physical body and a particular identity presentation, which can complicate who/what people may think they can interact with [3, 19] as well as how people control dialogue for robots. For instance, an identity performance strategy that may need to be considered in multi-robot interactions is *re-embodiment* in which a robot identity can switch between robot bodies, potentially to ease particular interactions [9, 12]. As such, in our interface design, we aimed to enable control over this flexibility in robot body-identity association.

Overall, in this work, we aim to facilitate multi-robot speech control while incorporating robot identity into the core of the interface. In the following section, we discuss our progress thus far in designing a multi-robot control interface that can uniquely enable WoZ for multi-robot HRI experiments.

3 DESIGN PROCESS

3.1 Brainstorming Requirements

We first had to brainstorm the requirements of the interface that may be needed to enable a user to control simultaneous multi-robot speech. This brainstorming led to seven key activities that such an interface should allow: (1) connecting to multiple robot bodies, (2) creating robot identities with distinct names and vocal parameterizations, (3) authoring of text ahead of time that robots might need to say during experiments, (4) changing which identities are associated with which robot bodies, (5) inputting text to be spoken by the connected robots, (6) triggering the speech of that text by selected robot bodies, and (7) assessing the status of connected robots.

Most of these requirements (1, 3, 5, 6, 7) were needed to achieve base functionality for the control of multiple robots and their speech. Additionally, requirements 2 and 4 were needed to account for the

complexities of robot identity that may be present among groups of robots. In particular, we determined that enabling a user to change robot body-identity association on-the-fly could provide users with a means of controlling robot identity performance such as enacting the re-embodiment strategy when needed.

3.2 Sketching and Prototyping

After identifying the interface requirements, we divided them into two key phases: *Robot Group Initialization* (requirements 1-3) and *Session Control* (requirements 4-7). For this interface, we decided to focus on the “Session Control” phase, in particular to explore the different ways users may change robot body-identity associations (requirement 4). We then sketched several possible layouts for a Session Control interface. Adhering to the idea of minimizing interactions necessary to change the robot body-identity association, we used “click” counts (i.e. number of user inputs to the interface) as a metric to guide the first iteration of sketches. The most promising sketches were then turned into design prototypes using Figma, an online interface design tool. Figure 1 shows the three design prototypes we created.

Thus far throughout our design process, we assumed that a user had already completed the “Robot Group Initialization” phase to facilitate the design and demonstration of these interfaces. As such, our prototypes demonstrate that a user had already (1) connected the interface to three robot bodies, (2) created three robot identities with unique names (Buddy, Bumble, and Honey) and voices, and (3) pre-authored three speech buttons (each intended to prompt speech saying “Hello, my name is ...”). To communicate these configurations set during the “Robot Group Initialization” phase, in each prototype, the leftmost side of the interface provides users with a list of the identities created (comprised of user-defined names and voices), and a list of connected robot bodies with color-coded icons to denote connection and processing status (idle, busy, or faulty). Each prototype also has two methods for users to input speech: (1) user-defined buttons of preplanned speech, and (2) a text box for on-the-fly input.

The three prototypes primarily differed in the workflow required to determine which robot bodies to use to utter inputted text, and which robot identities to use to parameterize that speech (otherwise referred to as setting the robot body-identity associations). Specifically, these interfaces differed in the specific visuals used to convey the robot body-identity association, and the types of input modalities intended for changing that association (e.g., radio buttons versus toggle buttons versus drag-and-drop).

3.3 Interviewing Potential Users

To help us iteratively refine the design of our multi-robot WoZ interface prototype, we conducted IRB-approved Zoom interviews with six HRI researchers. The recruited participants all had experience conducting human-subject WoZ studies and/or investigating interface design for robot control.

In these interviews, participants were shown the three multi-robot speech control interface prototypes in a semi-counterbalanced order. Specifically, three participants were shown prototypes in order (v1, v3, v2), and three were shown prototypes in order (v2, v1, v3). Prototype v3 was always shown directly after v1 because it was

designed as a modification to v1. When shown the first prototype, participants were told the general purpose of the interface and the configurations set during the “Robot Group Initialization” phase. For each prototype, participants were first asked for their initial impressions, and were then given a walkthrough of the expected use of the interface, through Figma *action prototyping*. Next, participants were asked for their thoughts on each prototype’s design and use. In particular, we asked (1) if the instructions presented in each prototype were sufficient to explain how to use the interface, (2) how each prototype compared to the others seen thus far (and which they preferred), and (3) what features/qualities might still be needed.

Overall, participant responses gave us insights on how to refine our prototyped design. In particular, these insights included information about (1) what users may need to learn how to use the interface (e.g. detailed instruction guide or simply enough time to engage with the interface), (2) what general design considerations to account for (e.g. standardization of design elements, and providing obvious means of input), (3) how to clearly communicate the connection between interface elements and the robots being controlled in the real world (e.g. including actual robot images rather than representative icons), and (4) what features and qualities more advanced users may want (e.g. the inclusion of robot movement in association with speech and potentially a means of auto-generating text).

3.4 Design Iteration and Future Work

Based on the insights from the potential user interviews, we intend to iterate over our prototyped designs by outlining any additional interface requirements brought up by participants and reconfiguring the layout of our multi-robot speech control interface (including the means by which requirement 4 is met). Moreover, to enable direct user testing of the interface and its eventual use within HRI experiments, our next goal is to begin the implementation of a working interface that could serve as a general tool in multi-robot speech control. Once an initial interface implementation is complete, we intend to conduct usability studies with potential users (HRI researchers) to test the functionality of the interface within a HRI experimental context as well as to receive further feedback on how to further improve our design.

4 CONCLUSION

There is a current lack of generally applicable, robot control interfaces that can be used to conduct WoZ experiments, especially those involving multiple robots. As such, in this work, we prototyped a user interface that would not only help with conducting multi-robot WoZ experiments in HRI, but also allow for the inclusion and exploration of robot identity. Through our design process, we developed three prototypes for such an interface and plan to iterate over these designs based on feedback received in interviews with HRI researchers. Overall, in this work we made strides towards the development of an accessible, generalizable research tool for WoZ research in multi-robot interactions.

ACKNOWLEDGMENTS

This work was supported by ARL DCIST W911NF-17-2-0181.

REFERENCES

- [1] Paul Baxter, James Kennedy, Emmanuel Senft, Severin Lemaignan, and Tony Belpaeme. 2016. From characterising three years of HRI to methodology and reporting recommendations. In *Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE.
- [2] Alexandra Bejarano, Samantha Reig, Priyanka Senapati, and Tom Williams. 2022. You Had Me at Hello: The Impact of Robot Group Presentation Strategies on Mental Model Formation. In *Proceedings of the 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.
- [3] Alexandra Bejarano and Tom Williams. 2023. No Name, No Voice, Less Trust: Robot Group Identity Performance, Entitativity, and Trust Distribution. In *32nd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*.
- [4] Abhinav Dahiya, Alexander M Aroyo, Kerstin Dautenhahn, and Stephen L Smith. 2023. A survey of multi-agent Human–Robot Interaction systems. *Robotics and Autonomous Systems* (2023).
- [5] Dylan F Glas, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2008. Simultaneous teleoperation of multiple social robots. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*.
- [6] Siying Hu, Hen Chen Yen, Ziwei Yu, Mingjian Zhao, Katie Seaborn, and Can Liu. 2023. Wizundry: A Cooperative Wizard of Oz Platform for Simulating Future Speech-based Interfaces with Multiple Wizards. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023).
- [7] Ryan Blake Jackson, Alexandra Bejarano, Katie Winkle, and Tom Williams. 2021. Design, performance, and perception of robot identity. In *Workshop on Robo-Identity: Artificial identity and multi-embodiment at HRI*.
- [8] Edith Law, Vicky Cai, Qi Feng Liu, Sajin Sasy, Joslin Goh, Alex Blidaru, and Dana Kulić. 2017. A wizard-of-oz study of curiosity in human-robot interaction. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE.
- [9] Michal Luria, Samantha Reig, Xiang Zhi Tan, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2019. Re-Embodiment and Co-Embodiment: Exploration of social presence for robots and conversational agents. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 633–644.
- [10] Matthew Marge, Stephen Nogar, Cory J Hayes, Stephanie M Lukin, Jesse Bloecker, Eric Holder, and Clare Voss. 2019. A Research Platform for Multi-Robot Dialogue with Humans. *NAACL HLT 2019* (2019).
- [11] Bastian Pflöging, Norman Seyffer, Zina Beek, and Sebastian Zug. 2023. A Wizard-of-Oz Setup to Assess Pedestrian Behavior and Teach Delivery Robots on Sidewalk Encounters. In *Workshopband Mensch und Computer 2023: MCI-WS17: Robots for Public and Social Spaces - Design for Intent Communication, Collaboration and Acceptance*. GI. <https://doi.org/10.18420/muc2023-mci-ws17-459>
- [12] Samantha Reig, Michal Luria, Elsa Forberger, Isabel Won, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2021. Social robots in service contexts: Exploring the rewards and risks of personalization and re-embodiment. In *Designing Interactive Systems Conference 2021*. 1390–1402.
- [13] Laurel D Riek. 2012. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction* 1, 1 (2012).
- [14] Finn Rietz, Alexander Sutherland, Suna Bensch, Stefan Wermter, and Thomas Hellström. 2021. WoZ4U: an open-source wizard-of-oz interface for easy, efficient and robust HRI experiments. *Frontiers in Robotics and AI* 8 (2021).
- [15] Eike Schneiders, EunJeong Cheon, Jesper Kjeldskov, Matthias Rehm, and Mikael B Skov. 2022. Non-dyadic interaction: A literature review of 15 years of human-robot interaction conference publications. *ACM Transactions on Human-Robot Interaction (THRI)* 11, 2 (2022).
- [16] Brett Stoddard, Abrar Fallatah, and Heather Knight. 2021. A web-based user interface for HRI studies on multi-robot furniture arrangement. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.
- [17] Hamish Tennent, Wen-Ying Lee, Yoyo Tsung-Yu Hou, Ilan Mandel, and Malte Jung. 2018. Paperino: Remote wizard-of-oz puppeteering for social robot behaviour design. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*.
- [18] Michael Villano, Charles R Crowell, Kristin Wier, Karen Tang, Brynn Thomas, Nicole Shea, Lauren M Schmitt, and Joshua J Diehl. 2011. DOMER: A Wizard of Oz interface for using interactive robots to scaffold social skills for children with Autism Spectrum Disorders. In *Proceedings of the 6th International Conference on Human-robot interaction (HRI)*.
- [19] Tom Williams, Daniel Ayers, Camille Kaufman, Jon Serrano, and Sayanti Roy. 2021. Deconstructed trustee theory: Disentangling trust in body and identity in multi-robot distributed systems. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 262–271.