Beyond Words: Standardizing Multimodal Robot Communication for HRI based on Design Patterns

Kathrin Pollmann
Fraunhofer Institute for Industrial Engineering IAO
Stuttgart, Germany
kathrin.pollmann@iao. fraunhofer.de

ABSTRACT

Robots increasingly act as interaction partners, making intuitive, inclusive communication essential for effective human–robot interaction (HRI). This paper argues that speech-based systems alone are insufficient to achieve this goal and introduces SHAPE_HRI, a structured approach for designing multimodal, non-verbal robot behaviors, integrating a human-centered design process, formal pattern notation, and a modular pattern language to support consistent and reusable interaction design standards. Empirical evaluations across two use cases provide insight into the empirical validation of the patterns generated by the SHAPE_HRI approach. By advancing standardized multimodal communication and outlining an adaptive AI-based framework, this work conceptually contributes to the future development of more intuitive, accessible, and context-aware HRI applications.

CCS CONCEPTS

• Human-centered computing → Interaction design → Interaction design process and methods → User centered design

KEYWORDS

Multimodal human-robot interaction; behavioral design patterns

ACM Reference format:

Kathrin Pollmann (2026). Beyond Words: Standardizing Multimodal Robot Communication for HRI based on Design Patterns. In *HRI 2026 Workshop on Empowering Human-Robot Communication Through Non-Linguistic Pathways (HuRoCo @ HRI 2026), 2026, Edinburgh, Scotland, UK.*

1 Introduction

As robots increasingly become human interaction partners and collaborators rather than autonomous machines, designing intuitive, safe, and acceptable interaction capabilities is essential. Researchers in human–robot interaction (HRI) are developing skills such as object handover, accompaniment, and context, speech, or face recognition to ensure smooth cooperation with humans. While these technical abilities form the basis for interaction, they are not sufficient to guarantee success: It is also important to make sure that a robot's action is understood intuitively by the human interaction partner and that the interaction is experienced as pleasant.

1.1. Multimodal Robot Behavior as a Success Factor for HRI

Although HRI is rapidly expanding, there is still no comprehensive understanding of how robot behavior should be designed to create positive user experiences. Today, robot communication is dominated by speech, relying on voice- or text-based formats. This narrow focus overlooks the fact that, as embodied technologies, robots naturally communicate through their movements, lights, sounds, and other non-verbal cues—whether intentionally designed or not. When verbal and non-verbal signals misalign, users may feel confused or uncomfortable, which reduces the user experience [13]. Moreover, speech is not always the best suited modality to communicate robot states, intentions and actions. In many situations, implicit cues that can be perceived in passing are preferable. Relying primarily on speech as the only communication channel can also exclude people with differing linguistic, cognitive, or cultural backgrounds. Spoken dialogue systems, while intuitive for some, are not universally accessible and often fail to generalize across contexts [16].

In contrast, non-verbal cues such as gesture, gaze, posture, or light-based signaling offer intuitive, implicit ways for users to understand a robot's state or intention in an instance, broaden accessibility and enrich the interaction experience. Designing robots with multimodal communication therefore provides a more inclusive, natural, and reliable basis for HRI. It enables users to read robots at a glance, reduces misunderstandings, and better matches the capabilities of embodied machines, ultimately leading to more pleasant and effective HRI.

1.2 Towards Design Standards and a Universal Robot Communication Language

Robots can make use of a broader range of communication modalities than humans, which comes with strengths and weaknesses: While it is yet difficult for most robots to mimic human gestures, facial micro expressions or speech modulation in a convincing and pleasant way, they offer communication channels that humans do not have, e.g. light and sound. Why not rely on these unique channels to enhance communication instead of trying to simulate human abilities?

I advocate for creating a unique robot communication language that leverages their communicative strengths and provides standards for HRI design. Current robot behaviors, are often inconsistent across platforms and use cases, making it difficult for users to interpret communication cues instantly and to establish a robust mental model for HRI. Establishing design standards reduces cognitive load, lowers entry barriers for a wider range of users, and fosters a sense of familiarity and trust, thereby enhancing the long-term acceptance and integration of robots in daily life. In addition, design standards in HRI reduces the effort and increases the efficiency of HRI design tasks: When defining the interaction concept for their robot, the development team can rely on reusable communication cues, instead of having to reinvent behavioral expressions for every new application.

1.2 Design Patterns as a Basis for Standardized HRI Design

Design patterns have been first introduced by Alexander [2] to organize architecture design knowledge in an structured and accessible way and have been picked up by different disciplines in software development [6] and human-computer interaction [3, 17][3]. Patterns offer a way to structure design knowledge and document design standards in a way that is practical and accessible. While some attempts have been made to apply patterns in HRI [7–9], those approaches have focused on rather narrow use cases or methodological aspects, and thus none of them has made a lasting impact for establishing HRI design standards. However, they have demonstrated that generating HRI patterns requires a different methodological approach and a different systematization than other disciplines.

2 A Pattern Approach to Multimodal, Human-centered HRI

I developed the SHAPE_HRI Approach (**S**ocial be**HA**vioral **P**att**E**rn Design for **H**uman-**R**obot **I**nteraction) that proposes a methodology and framework for establishing design standards for multimodal robot communication based on the concept of behavioral design patterns [11]. The approach focusses specifically on non-verbal robot expressions to leverage communicative strengths of robots rather than mimicking human communication. Thus, it contributes directly to establishing a universal robot communication language. Moreover, standardization facilitates knowledge transfer within the interdisciplinary HRI community, allowing practitioners to build upon proven solutions rather than starting from scratch. SHAPE_HRI is meant to provide a guideline for HRI practitioners from different disciplines for creating reusable robot behaviors. Ultimately, this approach leads to higher quality, more efficient, and more user-centered robot applications.

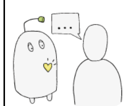
2.1 Components for Pattern Generation, Documentation and Organization

SHAPE_HRI consists of three main components that specify how patterns are generated, documented and related among each other to form a pattern database.

2.1.1 Design Process

The SHAPE_HRI design process provides a structured, human-centered methodology for generating reusable behavioral expressions for social robots. It guides designers through analysis, creation, and evaluation steps that ensure behaviors are grounded in user needs and backed by research. It uses a combination of structured user research and creative participatory design methods to help HRI designers to systematically create high-quality, multimodal robot behaviors that are comprehensible and pleasant for users. In brief, recurring interaction situations that may benefit from pattern generation are derived through Essential Use Cases Analysis [5]. For each recurring interaction situation potential communication mechanisms and modalities are gathered based on a review of the existing body of research in HRI as well as best practices from interactions between humans and humans, animals and technology. The results are documented and visually prepared to be further elaborated on in an Ideation Workshop with the goal of noting down specifications for each communication modality used by the pattern. This process is guided by the Modality Card

Table 1. Pattern "Listening", documented with the formal pattern notation format.

Name	**Listening**
Preamble	
Type	Composed Pattern, Having a conversation
Ranking	** (initially validated through user study)
Version	1
Author	Kathrin Pollmann
Design Challenge	
Interaction Situation	The user is telling the robot something. The robot provides feedback that it is aware of the user talking and that it is recording the spoken information.
CG	"I am listening to what you are currently saying."
Design Solution	
Solution Illustration 	Let the robot express that: • It directs its attention towards the user using a user-oriented positioning in the room, letting robot and user face each other. (Attentive) • It is currently recording spoken information from the user by highlighting the recording sensory organ (microphones or ears). (Speech recording) • Its operation mode is <on> setting the status light to permanently glowing. (Operation mode on)
Rationale	The act of listening to an interaction partner requires two sub-actions: facing the interaction partner to signal that your attention is focused on them and indicating that you are taking in and processing the provided information. In human-human interaction the latter is often expressed implicitly by signals of attention (eye contact, tilting the head to the side). In HRI it is especially important to make the information processing explicit, as a) the user otherwise cannot know that their speech input is processes at the moment and b) speech recording might require special informed consent, causing potential privacy issues. Further details about these two sub-actions can be found in the descriptions of the Atomic Patterns "Attentive" and "Speech recording".
Examples	Videos available at Robot Pattern Wiki
References and Context	`needs` Attentive AND Speech recording AND Operation mode on `is opposed to` all other Composed Patterns – two Composed Patterns can never be executed at the same time

Deck [10] that supports the selection, specification and documentation of multimodal behavioral expressions for robots.

2.1.2 Formal Pattern Notation

The formal pattern notation defines a standardized, text-based template for documenting the behavioral patterns created with the SHAPE_HRI design process. It organizes each pattern into clear components such as interaction situation, communication goal (CG), solution, and examples (compare table 1). Thus, patterns become accessible to interdisciplinary teams and can be transferred across robots and use cases.

2.1.3 Pattern Language

The pattern language establishes a systematization through a modular network that links patterns with each other, showing how they may be combined to create more complex robot behaviors. The whole pattern collection currently contains 60 patterns from nine categories which are all made available online at the Fraunhofer IAO Robot Behavioral Pattern Wiki (https://websites.fraunhofer.de/robot-pattern-wiki/). The categories are: navigating public spaces, indicating safety distance, having a conversation, communicating background activity, communicating robot state, managing user's attention, expressing emotions, giving feedback, playing a game, and dealing with robot bullying.

2.2 Empirical Pattern Validation

Out of the 60 patterns 19 have been validated in two empirical end user studies. Each study followed the process for pattern evaluation described by the SHAPE_HRI design process: First participants observed or experienced a complete scenario of a human-robot encounter, in which the robot used multimodal SHAPE_HRI patterns for communication. Next, they were presented with video snippets of each pattern in the order of their occurrence and provided their written or oral evaluation based on the following questions:

- *How comprehensible was the behavior of the robot to you?* (rated on a Likert scale ranging from "not at all" to "very")
- *As how pleasant did you experience the behavior of the robot?* (rated on a Likert scale ranging from "not at all" to "very")
- In your opinion, what does the robot shown in this video snippet want to express?
- Which behavior of the robot led to your interpretation?

The following paragraphs present an excerpt from the study results to provide insight into the empirical validation of patterns generated by the SHAPE_HRI approach. Comprehensibility and pleasantness were analyzed by calculating the mean values of the ratings for each pattern. Based on these means, pattern can be rendered as "validated" or "in need of revision". The two open questions were analyzed using content analysis. However, the results are not presented in this paper.

2.1.1 Use Case 1: Robot-Based Playful Brain Training

In the first user study, a scenario was evaluated where older adults played a brain training quiz game with a robotic quiz master [14, 15]. To enrich the behavior of the quiz master and include intuitive non-verbal cues to support verbal bidirectional exchange, 13 patterns were implemented on the MiRo robot [4], providing communication of robot state and background activity, expressing emotions, and facilitating a motivating quiz game experience.

The study was conducted as a video-based online study with 21 participants who qualified as *older adults* (over 60 years old; M_{age} = 66.5, SD = 7.9; all retired; 11 females, 10 males) and followed the the procedure and data analysis described above.

The results indicate that, overall, patterns were rated as comprehensible (M = 3.04) and neither pleasant nor unpleasant (M = 2.60). The overview of the ratings for the individual patterns (table 2) shows a broad range of mean ratings. Six out of 13 patterns reached mean ratings above 2.5 for both dimensions, which renders their status "initially validated". The other seven patterns show room for improvement.

2.2.2 Use case 2: Encounter with a Cleaning Robot

In the second study, 31 participants (M_{age} = 40.90, SD = 8.90; 15 female, 16 male) encountered the Adlatus CR700 sweeping robot [1] while it was cleaning an underpass close to a train station [12]. Patterns were implemented on the robot to enhance its ability to communicate its state, movement direction and safety distance. After the live encounter they were presented with the video snippets of the patterns and asked for their feedback.

Patterns were, on average, rated as comprehensible (M = 4.322) and pleasant (M = 3.92). An overview of ratings for the individual patterns is provided in table 2. All seven patterns received mean ratings that are clearly on the positive side of the scale (above 3) and can thus be labeled as "initially validated".

2.2.3 Synthesis and Discussion of Results

Taken together, the user studies showed that 12 out of 19 patterns received positive ratings for both dimensions and 7 patterns need revision. While these findings suggest that SHAPE_HRI is suitable to produce patterns that are perceived as comprehensible and pleasant on the first go, it also highlights the importance of the iterative nature of human-centered design and the SHAPE_HRI design processes. After their initial generation, patterns need empirical validation, sometimes repeatedly. In addition, to be considered truly validated, patterns have to be implemented on different robotic platforms and tested in different contexts.

The answers to the two open questions can shed light on how to improve patterns that fail initial validation, as they reveal misinterpretations of communication cues. They can also be used to objectify the subjective ratings, by matching participants' interpretation of the robot's behavior with the intended communication goal, resulting in the calculation of a comprehension rate for each pattern.

1.3. Outlook: Context- and User-Adaptive HRI

When establishing a universal robot communication language, we have to consider that the way a robot communicates non-verbally should be context- and user adaptive. For example, when lighting conditions change, the light signal of the robot may also have to be adjusted, or people with visual impairments may require other communication modalities than people with impaired hearing.

Table 2. Overview of the patterns that were evaluated in the two empirical studies and mean ratings for comprehensibility and pleasantness. Patterns that qualified as "initially validated" are marked in blue. * Ratings were provided on a 4-point Likert scale (1 = not at all; 4 = very); ** Ratings were provided on a 5-point Likert scale (1 = not at all; 5 = very);

Pattern Name	Communication Goal	Comprehensibility		Pleasantness	
		M	SD	M	SD
Use Case 1: Robot-based playful brain training*					
Inactive	„I am currently not engaged in interaction and not capable of initiating an interaction with you."	3.14	0.71	2.23	0.79
Becoming active	"I am changing from a passive, inactive to an active state where I am fully ready to interaction with you."	3.71	0.45	3.14	0.55
Active	„I am ready to interact with you and capable of initiating interaction by myself."	2.48	0.91	2.41	0.72
Listening	"I am listening to what you are currently saying."	3.67	0.56	3.14	0.55
Explaining	"I am explaining something to you. Stay focused on me!"	2.95	0.84	2.36	0.71
Getting Ready to play	"I am loading the game. I am looking forward to playing it with you. I will let you know when I am ready."	2.57	1.05	2.09	0.90
Showing	"I am showing you information. Please pay attention to me and to this information."	3.67	0.47	3.14	0.62
Processing	„I am processing what I just learned from you. This will take some time – I will tell you when I am ready."	2.95	1.00	2.91	0.90
Not Understanding	„I know that you just provided some information, but I do not understand. "	3.19	0.85	2.64	0.83
Joyful positive feedback	"I can tell you that what you did / said is correct and I feel happy for you."	2.27	0.98	2.09	0.85
Empathic negative feedback	"I have to tell you that what you did / said is incorrect. I assume that you are disappointed, and I feel with you."	3.05	0.85	2.82	0.89
Encouraging good performance	"I believe in you and I'll support you to show good performance in the upcoming action!"	2.62	1.13	2.00	0.74
Becoming inactive	"I am changing from a ready, active to a passive, inactive state where I am still available, but not capable of initiating interactions with you."	3.24	0.81	2.91	0.90
Use Case 2: Encounter with a cleaning robot **					
Attentive towards surroundings	"I am scanning my surroundings."	4.26	0.68	4.00	0.77
Working	"I am working."	4.45	0.62	4.03	0.84
Slowing down	"I am slowing down."	3.71	1.13	3.90	0.79
Moving backwards	"I am moving backwards."	4.35	1.05	3.39	1.09
Turning on the spot	"I am turning on spot."	4.26	0.73	4.06	0.77
Turning left / right	"I am turning left / right."	4.55	0.68	4.16	0.78
Safety zone	"This is my safety zone. I'm keeping my distance from you."	4.68	0.54	3.87	1.02

envision a fully adaptive system where the robot is extended by an Artificial Intelligence (AI)-based module that can interpret context data that is detected via the robot's sensors and selecting a matching pattern from the implemented collection of patterns. This pattern is then executed on the robot (Figure 1).

A multimodal, adaptive, pattern-based communication framework can significantly improve intuitiveness of robot communication and enhances the individual interaction experience. Most importantly, it strengthens the importance of multimodal, non-verbal communication as not only an add-on to speech-based interaction, but as a communication strategy on its own that will be essential for creating effective, user-centered HRI applications in the future.

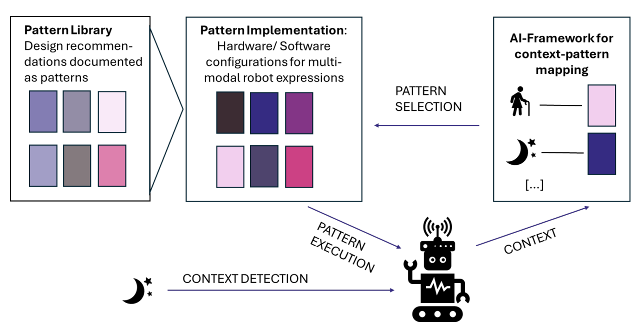


Figure 1. Overview of a potential framework for multimodal, pattern-based adaptive HRI

ACKNOWLEDGMENTS

This research was conducted as part of the projects NIKA and ZEN-MRI, funded by the German Federal Ministry of Research, Technology and Space ((BMBF 16SV7941, BMFTR 16SV8932)

REFERENCES

[1] Adlatus Robotics GmbH. *Reinigungsroboter CR700C/D*. Retrieved December 2, 2025 from https://www.adlatus.eu/adlatus-robotics/.

[2] Christopher Alexander. 1977. A pattern language: towns, buildings, construction. Oxford University Press.

[3] Jan O. Borchers. 2000. A pattern approach to interaction design. In Proceedings of the 3rd International Conference on Designing interactive systems: processes, practices, methods, and techniques. ACM, New York, 369–378.

[4] Consequential Robotics. The MIRO Robot. Retrieved August 14, 2019 from https://consequential.bitbucket.io/.

[5] Larry L. Constantine and Lucy A. D. Lockwood. 1999. Software for use: a practical guide to the models and methods of usage-centered design. Pearson Education.

[6] Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides. 1995. Design Patterns: Elements of Reusable Object-Oriented Software Addison-Wesley. Reading, MA, 1995.

[7] Peter H. Kahn, Nathan G. Freier, Takayuki Kanda, Hiroshi Ishiguro, Jolina H. Ruckert, Rachel L. Severson, and Shaun K. Kane. 2008. Design patterns for sociality in human-robot interaction. In Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction. ACM, 97–104.

[8] Manja Lohse, Frederic Siepmann, and Sven Wachsmuth. 2014. A modeling framework for user-driven iterative design of autonomous systems. International Journal of Social Robotics 6, 1, 121–139.

[9] Julia Petalson. 2013. Modeling Human-Robot-Interaction Based on Generic Interaction Patterns. Dissertation. University of Bielefeld, Bielefeld.

[10] Kathrin Pollmann. 2021. The Modality Card Deck: Co-Creating Multi-Modal Behavioral Expressions for Social Robots with Older Adults. Multimodal Technologies and Interaction 5, 7. DOI: https://doi.org/10.3390/mti5070033.

[11] Kathrin Pollmann. 2024. A Human-Centered Pattern Approach to Comprehensible and Pleasant Behavioral Expressions for Social Robots. Dissertation. University of Stuttgart, Stuttgart.

[12] Kathrin Pollmann, Selina Layer, and Katharina Hölzle. 2025. Mensch und Roboter Hand in Hand/Humans and robots teaming up – Interaction design patterns for enhancing the communication skills of autonomous robots. wt Werkstattstechnik online 115, 06, 411–418.

[13] Kathrin Pollmann, Christopher Ruff, Kevin Vetter, and Gottfried Zimmermann. 2020. Robot vs. Voice Assistant: Is Playing with Pepper More Fun than Playing with Alexa? In Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. HRI '20. Association for Computing Machinery, New York, NY, USA, 395–397. DOI: https://doi.org/10.1145/3371382.3378251.

[14] Kathrin Pollmann and Daniel Ziegler. 2020. Personal Quizmaster: A Pattern Approach to Personalized Interaction Experiences with the MiRo Robot. In Proceedings of the Conference on Mensch und Computer, Bernhard Preim, Andreas Nürnberger and Christian Hansen, Eds. MuC'20: Mensch und Computer 2020. Association for Computing Machinery, New York, NY, USA, 485–489. DOI: https://doi.org/10.1145/3404983.3410414.

[15] Kathrin Pollmann and Daniel Ziegler. 2021. A Pattern Approach to Comprehensible and Pleasant Human–Robot Interaction. Multimodal Technologies and Interaction 5, 9. DOI: https://doi.org/10.3390/mti5090049.

[16] Elaheh S. Shourmasti, Ricardo Colomo-Palacios, Harald Holone, and Selina Demi. 2021. User Experience in Social Robots. Sensors 21, 15. DOI: https://doi.org/10.3390/s21155052.

[17] Jenifer Tidwell. 2010. Designing Interfaces: Patterns for effective Interaction Design. 2nd edition. O'Reilly Media, Inc., Canada.