Beyond the Drop-Off: AI-Driven Communication Strategies for Mobile Delivery Robots

Volker Göhler*
TU Bergakademie Freiberg
Freiberg, Germany
volker.goehler@informatik.tu-freiberg.de

Felix Heisel*
TU Bergakademie Freiberg
Freiberg, Germany
felix.heisel@informatik.tu-freiberg.de

Bastian Pfleging
TU Bergakademie Freiberg
Freiberg, Germany
bastian.pfleging@informatik.tu-freiberg.de

(a) A delivery robot handing over a parcel

(b) A delivery robot communicates with parcel recipients

Figure 1: Figures generated by ChatGPT with the following prompts: a) "Draw a delivery robot handing over a parcel to a human at a family home" and b) "Draw a delivery robot that communicates human-like with parcel recipients".

Abstract

The rapid rise of automated delivery robots promises faster, contact-free parcel handovers, yet the final interaction between the robot and the end customer remains a fragile touchpoint. Misaligned expectations, unclear instructions, or inappropriate communication can turn a convenient delivery into a source of frustration. We propose an AI-driven handover coordinator that ingests dispatch metadata, real-time environmental conditions (weather, location, traffic), the recipient's historical interaction profile, and live feedback during the handover. By synthesizing these signals, the AI generates a tailored communication strategy, adjusting tone, movement, and content, to maximize the recipient's satisfaction. At the same time, this raises critical questions about data privacy, robustness, ethical persuasion, and real-time performance that must be addressed for the solution to be viable. We hope to spark a critical discussion about the possibilities of using AI in this context in order to pave the way for a design space for delivery robots in the long term.

*Both authors contributed equally to this research.

The 3rd InterAI Workshop at CHI 2026, Barcelona, Spain

ACM ISBN 978-1-4503-XXXX-X/2026/04
https://doi.org/XXXXXXX.XXXXXXX

CCS Concepts

• **Human-centered computing → Interactive systems and tools**; *HCI theory, concepts and models*; **Ubiquitous and mobile computing systems and tools**.

Keywords

Human-Robot-Interaction, Context-Awareness, Last-Mile Delivery Robot, Parcel Handover, Inclusive Design

ACM Reference Format:
Volker Göhler, Felix Heisel, and Bastian Pfleging. 2026. Beyond the Drop-Off: AI-Driven Communication Strategies for Mobile Delivery Robots. In *The 3rd InterAI Workshop: "Interactive AI for Human-Centered Robotics at ACM CHI 2026, April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 4 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

As automated delivery robots become more integrated into society, it is essential to design the quality of the final parcel handover for user satisfaction. This brief encounter constitutes a complex, coordinated activity in which the robot's communication abilities shape the user's entire experience. For these interactions to be successful, the robot must do more than simply complete its task. It must communicate in a way that is not only clear and efficient but also socially and emotionally intelligent. We propose a conceptual handover system that serves as a basis for discussing synchronized verbal and non-verbal communication strategies aimed at fostering positive user experiences. Building on this framework, we highlight key challenges and open questions to stimulate critical discussion on affective communication, user perception, and responsible deployment.

2 Background

2.1 Communication Modalities

Research on social robotics shows that users expect verbal communication while non-verbal cues are often more effective for conveying intent and ensuring a fluent handover. Although users anticipate verbal interaction from a robot, non-verbal channels are frequently highlighted as more positive and trust-building [16]. The integration of both verbal and non-verbal modalities is crucial for creating a socially competent robot that can manage interactions effectively to promote natural and engaging communication [3]. A multimodal approach that synchronizes speech with clear, unambiguous physical cues like gaze and gesture is most effective. This combination enhances user engagement and social interaction, making the interaction feel more natural and reliable [13].

2.2 Mood and User Experience in Social Robotics

In human-robot interaction (HRI), user experience (UX) is often defined by the user's perception of the robot's warmth, competence, and likeability [1]. This positive perception can be cultivated through mood contagion, an automatic mechanism where observing an emotional expression induces a similar mood in the observer [8]. Studies confirm that a robot expressing a positive mood through its body language can cause users to report a matching positive mood, especially in low-stress tasks [18]. By modulating the parameters of functional movements, such as posture and motion dynamics, a robot can express mood. And when recognized by the user, this can enhance collaboration, improve satisfaction with the process, and lead to better performance outcomes [2].

2.3 Conversational and Behavioural Strategies for Positive Interaction

A delivery robot can foster a positive handover by combining empathetic verbal strategies with expressive, context-aware non-verbal behaviours. To achieve this, the robot must first attempt to recognize the user's emotional state through cues like body language [15]. Based on this assessment, it can apply empathetic regulation, a process of producing an appropriate affective response to influence the user's perception and mood positively [5, 10]. This bidirectional communication is most effective when verbal approaches are synchronized with expressive physical motions [3]. By modulating functional movements, the robot can convey a positive disposition and enhance user engagement without disrupting the delivery task [4]. Table 1 provides examples of how to apply these multimodal strategies.

3 Methodology

Building on the insights from the related work, it becomes evident that many adaptive communication frameworks are evaluated in highly embodied settings, most notably with humanoid robots. In contrast, the present work focuses on a delivery robot platform that is only weakly embodied. While the integration of manipulators is technically conceivable, most delivery robots do not possess limbs and therefore cannot rely on expressive gestures or body language comparable to humanoid systems. Consequently, interaction is restricted to a limited set of communication channels, primarily visual displays and audio output. Non-verbal cues and gestural behaviour must therefore be approximated through alternative modalities, such as subtle movement patterns, orientation changes, or short micro-driving sequences. These elements can function as proxies for expressive behaviour and allow the system to convey communicative intent despite the robot's minimal embodiment [12, 17].

3.1 General System Overview

We propose a system in which an AI component dynamically adapts communication strategies based on heterogeneous contextual inputs, as illustrated in Figure 2.

The system integrates three main categories of input data. First, *sender data* represent relatively stable background information about the initiator of the communication. This may include whether the sender is a private individual or an organization, as well as contextual attributes such as industry affiliation or service type that can be derived from publicly available information. Second, *environmental data* capture situational context through sensor data and external information sources. These inputs include factors such as weather conditions, seasonal context, public holidays, and characteristics of the local surroundings. Such information reflects the immediate interaction context and allows the AI to adapt its communication behaviour dynamically. Third, *recipient history* is stored in dedicated databases and provides longitudinal information about previous interactions with the recipient. This historical perspective enables the system to personalize communication strategies over time and to adapt interaction styles according to observed user preferences and prior experiences.

Based on these inputs, the AI selects appropriate verbal and non-verbal communication strategies and mediates the interaction between the robot and the user. During the interaction, observable user feedback signals are continuously analysed, enabling iterative refinement of communication behaviour and allowing the system to adjust its strategy in real time.

3.2 Evaluation Measures

To evaluate user perception and interaction quality, we propose a combination of subjective and behavioural measurements.

First, subjective feedback will be collected through established questionnaire instruments administered after the interaction. These include the Perceived Safety Questionnaire (PSQ)[9], the System Usability Scale (SUS)[6], and the User Experience Questionnaire Short (UEQ-S)[14]. These instruments provide standardized measures for perceived safety, usability, and overall user experience, enabling comparisons between different communication strategies. Second, the system collects behavioural indicators related to user stress and emotional state during the interaction. These signals are obtained through multimodal sensing approaches, including facial expression recognition, posture and gait analysis, and voice stress

Table 1: Weakly embodied Robot Communication Strategies Based on User State

Scenario	Verbal Strategy	Non-Verbal Strategy
Positive/Neutral User	Use simple, positive greetings ("Hello! I have your delivery.").	Employ smooth motion, and direct "gaze" to build rapport and appear lively [7].
Negative/Stressed User	Offer empathetic regulation ("No rush, I can wait.").	Use slower, more deliberate movements and avert "gaze" to reduce social pressure and signal patience [5, 11].

analysis. From these inputs, arousal–valence estimates as well as discrete emotion classifications can be derived. Finally, to assess the perceived social characteristics of the robot, the Robotic Social Attributes Scale (RoSAS)[1] will be employed. This instrument measures dimensions such as perceived warmth, competence, and discomfort, thereby providing additional insight into how adaptive communication strategies influence the perceived social qualities of the robotic system.

4 Discussion

Although we present a specific solution, its design and deployment raise a range of interconnected technical, ethical, and societal challenges that warrant careful discussion. In the following, we reflect on several key dimensions that are central to the responsible and effective deployment of such systems.

With respect to data protection, considerations must be carefully balanced against the expected benefits of the system. In particular, it is necessary to assess whether potential improvements in user experience and usability justify the collection and processing of personal data in both scope and volume. Furthermore, a systematic evaluation of the contribution of individual data elements is necessary, as this may enable the reduction or elimination of less influential inputs.

Regarding model robustness and interpretability, the reliability of the AI's context and affect inference is critical, as errors directly propagate into the system's adaptive communication behaviour. Misclassifications of emotional states, situational cues, or user intent may result in inappropriate responses, thereby significantly deteriorating UX and trust in the system. Restricting the system to well-defined "safe" communication strategies may help mitigate such risks, particularly in ambiguous situations. In addition, cultural and individual differences must be considered to avoid systematic misinterpretations.

Concerning data reliability, environmental and contextual information may be inaccurate or unavailable due to technical limitations or external service failures. Similar limitations may affect other data sources. Such disruptions can negatively impact the system's ability to adapt its communication behaviour. Consequently, redundancies and fallback mechanisms must be considered during system design and development to ensure robust and reliable operation under imperfect data conditions.

From an ethical perspective, attempts to influence the mood of package recipients raise important questions about acceptable boundaries and potential misuse. To mitigate such risks, clear operational boundaries should be defined in advance, for instance, restricting the robot's communicative behaviour strictly to the context of package delivery. In addition, publicly accessible guidelines are advisable in order to promote transparency and foster user acceptance of the system. Recipients should be informed that the delivery robot may adapt its communication style based on contextual and behavioural information. Providing clear transparency mechanisms, such as brief notifications on the robot's display or accessible explanations of the system's decision logic, may help users understand why certain communication styles are employed. In addition, opt-out mechanisms should be considered that allow recipients to disable personalized or adaptive communication if they prefer a neutral interaction.

Finally, with regard to infrastructural constraints, limitations in mobile connectivity and on-device processing can negatively affect system performance and user experience. At the same time, substantial data exchange is essential for realizing the proposed concept. Therefore, an appropriate balance must be established between latency, data volume, and quality of AI-generated results.

5 Conclusion

In this workshop paper, we introduced a concept for context-dependent adaptation of the communication style of a delivery robot during parcel handover to recipients. The described scenario leads to special communication cases and channels. In addition to the potential input variables that the AI can use for inference, critical challenges were examined. This includes data protection and privacy concerns, the risk of hallucinations in AI models, inaccuracies or incompleteness in input data, and potential disruptions in data transmission.

References

[1] Colleen M. Carpinella, Alisa B. Wyman, Michael A. Perez, and Steven J. Stroessner. 2017. The Robotic Social Attributes Scale (RoSAS): Development and Validation. *ACM/IEEE International Conference on Human-Robot Interaction* Part F127194 (3 2017), 254–262. doi:10.1145/2909824.3020208

[2] Alwin de Rooij, Simone van den Broek, Michelle Bouw, and Jan de Wit. 2024. Co-Creating with a Robot Facilitator: Robot Expressions Cause Mood Contagion Enhancing Collaboration, Satisfaction, and Performance. *International Journal of Social Robotics 2024 16:11* 16, 11 (10 2024), 2133–2152. doi:10.1007/s12369-024-01177-3

[3] Alexander Hong, Nolan Lunscher, Tianhao Hu, Yuma Tsuboi, Xinyi Zhang, Silas Franco Dos Reis Alves, Goldie Nejat, and Beno Benhabib. 2021. A Multimodal Emotional Human-Robot Interaction Architecture for Social Robots Engaged in Bidirectional Communication. *IEEE Transactions on Cybernetics* 51, 12 (12 2021), 5954–5968. doi:10.1109/TCYB.2020.2974688

[4] Yuhan Hu, Peide Huang, Mouli Sivapurapu, and Jian Zhang. 2025. ELEGNT: Expressive and Functional Movement Design for Non-anthropomorphic Robot. *Proceedings of Designing Interactive Systems (Conference DIS'25)* 1 (1 2025). http://arxiv.org/abs/2501.12493

[5] Lesong Jia, Jackie Ayoub, Danyang Tian, Miao Song, Ehsan Moradi Pari, and Na Du. 2025. Emotion Regulation in Human-Robot Interaction: Insights from Video and Game Simulations. *Proceedings of the Human Factors and Ergonomics Society*

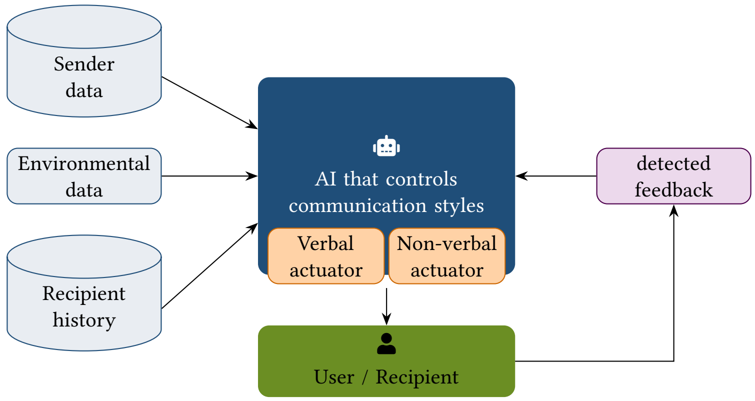

Figure 2: Schematic overview of the input data sources and processing flow used by the AI to derive a communication style tailored to the user.

Annual Meeting 69, 1 (9 2025), 1025–1027. doi:10.1177/10711813251368825
[6] James R. Lewis. 2018. The System Usability Scale: Past, Present, and Future. *International Journal of Human-Computer Interaction* 34, 7 (7 2018), 577–590. doi:10.1080/10447318.2018.1455307
[7] Raveesh Meena, Kristiina Jokinen, and Graham Wilcock. 2012. Integration of gestures and speech in human-robot interaction. *3rd IEEE International Conference on Cognitive Infocommunications, CogInfoCom 2012 - Proceedings* (2012), 673–678. doi:10.1109/CogInfoCom.2012.6421936
[8] Roland Neumann and Fritz Strack. 2000. "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality and Social Psychology* 79, 2 (2000), 211–223. doi:10.1037/0022-3514.79.2.211
[9] Sebastian Osswald, Daniela Wurhofer, Sandra Trösterer, Elke Beck, and Manfred Tscheligi. 2012. Predicting information technology usage in the car: Towards a car technology acceptance model. In *AutomotiveUI 2012 - 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, In-cooperation with ACM SIGCHI - Proceedings*. 51–58. doi:10.1145/2390256.2390264
[10] Sung Park and Mincheol Whang. 2022. Empathy in Human–Robot Interaction: Designing for Social Robots. *International Journal of Environmental Research and Public Health 2022, Vol. 19, Page 1889* 19, 3 (2 2022), 1889. doi:10.3390/ijerph19031889
[11] Bjorn Petrak, Julia G. Stapels, Katharina Weitz, Friederike Eyssel, and Elisabeth Andre. 2021. To move or not to move? Social acceptability of robot proxemics behavior depending on user emotion. *2021 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021* (8 2021), 975–982. doi:10.1109/RO-MAN50785.2021.9515502
[12] Liam Roy, Elizabeth A. Croft, Alex Ramirez, and Dana Kulic. 2025. GPT-Driven Gestures: Leveraging Large Language Models to Generate Expressive Robot Motion for Enhanced Human-Robot Interaction. *IEEE Robotics and Automation Letters* 10, 5 (2025), 4172–4179. doi:10.1109/LRA.2025.3547631
[13] Tim Schreiter, Lucas Morillo-Mendez, Ravi T. Chadalavada, Andrey Rudenko, Erik Billing, Martin Magnusson, Kai O. Arras, and Achim J. Lilienthal. 2023. Advantages of Multimodal versus Verbal-Only Robot-to-Human Communication with an Anthropomorphic Robotic Mock Driver. *IEEE International Workshop on Robot and Human Communication, RO-MAN* (2023), 293–300. doi:10.1109/RO-MAN57019.2023.10309629
[14] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 6 (2017), 103. doi:10.9781/ijimai.2017.09.001
[15] Chapa Sirithunge, A. G.Buddhika P. Jayasekara, and D. P. Chandima. 2019. Proactive Robots with the Perception of Nonverbal Human Behavior: A Review. *IEEE Access* 7 (2019), 77308–77327. doi:10.1109/ACCESS.2019.2921986
[16] Nicole Strutz, Luis Perotti, Anika Heimann-Steinert, and Robert Klebbe. 2024. Older adults' communication with an interactive humanoid robot. *Zeitschrift für Gerontologie und Geriatrie 2024 57:5* 57, 5 (1 2024), 371–375. doi:10.1007/s00391-023-02268-y
[17] Jacqueline Urakami and Katie Seaborn. 2023. Nonverbal Cues in Human-Robot Interaction: A Communication Studies Perspective. *ACM Transactions on Human-Robot Interaction* 12, 2 (3 2023). doi:10.1145/3570169
[18] Junchao Xu, Joost Broekens, Koen Hindriks, and Mark A. Neerincx. 2015. Mood contagion of robot body language in human robot interaction. *Autonomous Agents and Multi-Agent Systems 2015 29:6* 29, 6 (8 2015), 1216–1248. doi:10.1007/s10458-015-9307-3

A Disclaimer

Generative artificial intelligence (AI) tools were used exclusively for the purpose of reviewing and refining grammar and vocabulary in the preparation of this manuscript. At no point were AI tools used to generate original content or free-form text. All conceptualization, analysis, and writing were conducted by the authors.