

# Use of Irony and Sarcasm for Uncertainty in HRI

Mario Barbato<sup>1</sup>, Alessandra Rossi<sup>1</sup> and Silvia Rossi<sup>1</sup>

<sup>1</sup>Department of Electrical Engineering and Information Technologies, University of Naples "Federico II", Piazzale Tecchio 80, 80125, Naples, Italy

## Abstract

Autonomous robots are being used in human-centred environments, such as offices, restaurants, hospitals and private homes, for carrying out collaborative and cooperative tasks. These activities require that robots engage people in socially acceptable ways, even when they make errors. It is very common that robots make communication failures due to technical or environmental limitations, such as mismatch of multimodal observations. While these errors cannot be entirely avoided, it is still necessary to minimize them. In this paper, we want to use sarcasm by using contrasting multiple cues, both verbal and non-verbal, for allowing a robot to hide its uncertainty of the interaction signals. The results indicate some differences between the two attitudes, such as in the robot's independence and assertiveness.

## Keywords

HRI, uncertainty, humour

## 1. Introduction

Social robotics is a rapidly developing field. Thanks to advancements in hardware and software, encountering a robot is becoming an increasingly common event: from hospitals, where they interact with both children and older people, to museums as guides, and even in restaurants to serve customers [1, 2]. Once placed in these unsupervised scenarios, the likelihood that a robot may make errors increases, in particular when robots need to interpret the interactions with humans via multiple signals. Disruptions such as ambient noise, poor lighting, or the higher dynamism of a real-world context often lead to contrasting or uncertain signals while, as a consequence, produce social failures during a human-robot interaction (HRI). Social failures are errors that violate social norms and can degrade the perception of the robot's social and affective abilities [3]: not listening to the interlocutor, interrupting while they are speaking, or changing the subject without reason are just a few examples of social failures. In the field of HRI, it has been necessary to study behavioural techniques to mitigate the problem. One of these techniques is the use of humour by the robot, inspired by human-human interactions. Humour is pervasive in social relationships, being one of the most common ways to produce a positive influence on others: it has been shown that the use of spontaneous humour makes individuals more likeable and attractive in the eyes of others [4, 5], making them more friendly and improving the trust conveyed [6]. In situations free from specific tasks [7], such as making icebreaker jokes, telling puns and then apologizing, contrasting serious topics with jokes to ease tension, and being self-deprecating, humour generate laughs and empathy towards the agent. Even in more structured scenarios, such as vaccination in a hospital or reception in a hotel, there have been advantages noted in personalities endowed with humour compared to neutral ones, specifically in terms of engagement, likeability, ease of interaction and empathy [8, 9]. Exploring the vast field of humour, we particularly focus on irony and sarcasm. Some HRI researchers [10] have suggested that these can bring benefits to interaction, but it is not easy to understand how to effectively incorporate them into the robot's personality, since not everyone has the same humour or sarcasm. In this work we decided to adopt the Incongruity Theory

---

ALTRUIST, BAILAR, SCRITA, WARN 2024: Workshop on social robots for personalized, continuous and adaptive assistance, Workshop on Behavior Adaptation and Learning for Assistive Robotics, Workshop on Trust, Acceptance and Social Cues in Human-Robot Interaction, and Workshop on Weighing the benefits of Autonomous Robot personalisation. August 26, 2024, Pasadena, USA

✉ mario.barbato5unina@gmail.com (M. Barbato); alessandra.rossi@unina.it (A. Rossi); silvia.rossi@unina.it (S. Rossi)

ORCID 0000-0003-1362-8799 (A. Rossi); 0000-0002-3379-1756 (S. Rossi)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

[11], according to which humour is described as a process related to the experience of inconsistency, focusing on unexpectedness and inappropriateness. The approach proposed in this study is based on handling a delicate management of episodes where multimodal user feedback is deemed unreliable. In these cases, the robot reacts sarcastically by contrasting verbal cue polarity (i.e., the spoken phrase) with non-verbal cues, which include voice pitch and speed, facial expression (i.e., colour of LEDs), and gestures. The goal is to elicit a positive, particularly amused, reaction from the user, avoiding the unpleasant scenario where the interlocutor realizes that the robot did not actually understand, leading to a poorer perception of the robot's social and affective abilities.

## 2. The Scenario

The incongruity-based behaviour approach has been integrated into *BRILLO* (Bartending Robot for Interactive Long Lasting Operations), a three-year national project aimed at creating an autonomous robotic system capable of performing bartender tasks and interacting naturally with customers. The typical BRILLO scenario involves a user and three interaction systems: a kiosk where the user authenticates/registers to order a cocktail, the bartender robot that prepares the drinks, and optionally a waiter robot tasked with serving the customer if they are at a table. Focusing on the bartender robot, which is part of the system equipped with the proposed approach, it consists physically of a head, represented by Furhat, a torso, and two robotic arms for drink preparation. From an interaction standpoint, a key element is personalization and recommendation of both the drinks and the interaction. The robot adapts its behaviour based on the context, classifying the current customer into one of several profiles (e.g., with a worker on a lunch break, the bartender will converse in a way that relaxes them, while with a curious person, it will try to discuss various topics).

## 3. The Use Case

In this section, we present the acquisition and processing of user input, the decision algorithm for the behaviour to be adopted, and non-verbal signals configurations, and the results of use case scenario will be presented.

### 3.1. Input Acquisition and Processing

During the dialogue, two types of input are taken from the robot's sensors: voice and face. The first input is immediately processed by the robot's speech-to-text module, which transmits the phrase to the cloud service *LUIS*<sup>1</sup> (Language Understanding Intelligent Service) to perform intent recognition, necessary to understand the user's will, highlighting the involved entities, and sentiment analysis to calculate the sentence emotional polarity (positive, neutral or negative). The second input, in the form of video, is sent to the *Affectiva* tool<sup>2</sup> to recognize, by processing frame by frame on a cloud-based Docker container, the facial expression (positive: happiness, surprise; neutral; negative: sadness, anger, contempt, disgust and fear).

### 3.2. Incoherent Behaviour Decision

Once the inputs are processed, the core of the process begins: the decision on which behaviour the robot should adopt, based on the user feedback polarity. Given the user's facial expression and speech polarities:

- If the user's facial expression polarity is neutral, the robot will interact coherently, using verbal and non-verbal signals with the same polarity as the user's speech.

---

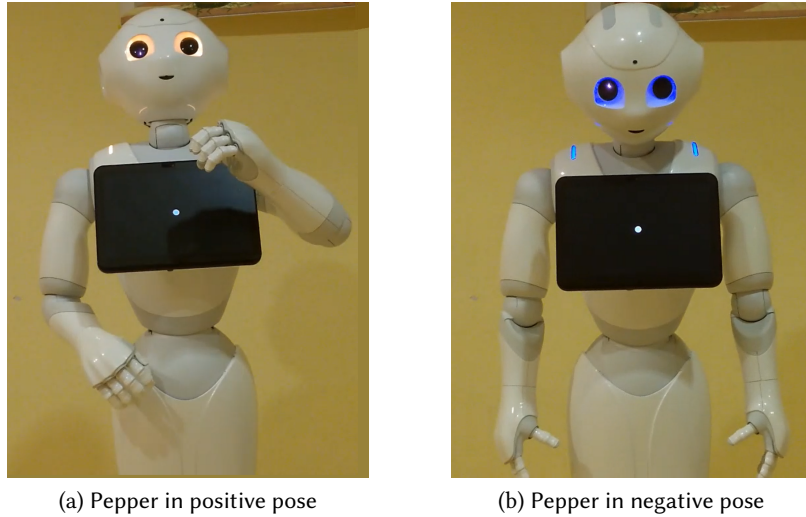
<sup>1</sup>LUIS <https://www.luis.ai>

<sup>2</sup>Affectiva SDK:

<https://www.affectiva.com/science-resource/affdex-sdk-a-cross-platform-realtime-multi-face-expression-recognition-toolkit/>

- If the user’s facial expression polarity is not neutral, it is compared with the user’s speech polarity:
  - If the two user polarities are the same, a coherent behaviour will be chosen.
  - If the two user polarities are different, an incoherent behaviour will be chosen: the robot’s non-verbal signals will be opposite to its speech, simulating sarcasm.

We observed that the *LUIS* output was more robust compared to that of *Affectiva SDK*. Example of Pepper non-verbal cues are shown in Figure 1, while the full detailed non-verbal cues defined by polarity are reported in Table 1.



**Figure 1:** Non-verbal cues examples by polarity

**Table 1**  
Non-verbal cues configurations by polarity

Polarity	LEDs	Animation	Tone/Speed
Positive	Yellow	Excited_1, Happy_4, Hysterical_1, Enthusiastic_4, Enthusiastic_5, Yes_1	1.2/120
Negative	Blue	Desperate_2, Desperate_4, Desperate_5, Embarrassed_1	0.96/90

### 3.3. Experimental Design

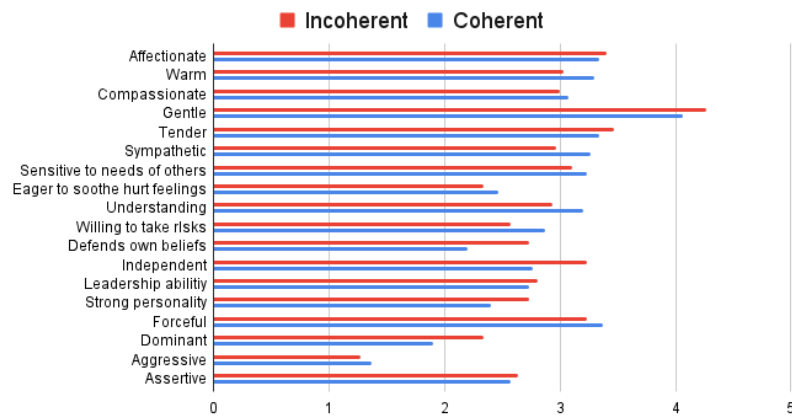
To evaluate the proposed approach, an online study was conducted as a within-subject counterbalanced, repeated measures study. The study was organised in three phases. Firstly, we collected demographic information (i.e., age, gender, nationality, and previous experiences with robots), then we asked them to complete the Italian adaptation of the Humor Styles Questionnaire [12], with the aim of finding the participant humour style closest to the among:

- **Affiliative:** focuses on everyday life events, creating a sense of bonding with the listener.
- **Self-enhancing:** involves laughing at oneself and one’s abilities, often being perceived as humble.
- **Aggressive:** includes insults and anything aimed at putting someone else down, typical of bullying.

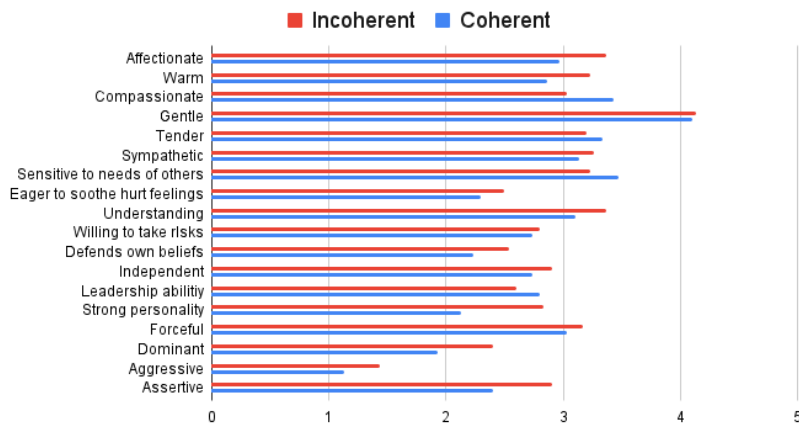
- **Self-defeating:** involves putting oneself down aggressively, often ending up ridiculing oneself.

Participants were tested either with Coherent or Incoherent robot’s behaviours. We asked participants to watch two videos of Pepper welcoming a customer and entertaining them in small talk, such as asking how they were doing. In both videos, the robot adopts the user sentence polarity for its verbal reply; in the first video it is positive, in the second one it is negative. The distinction between the two groups of participants occurs in the non-verbal signals: in the two videos shown, half of the participants witness the robot’s coherent behaviour (non-verbal polarity matches verbal polarity), and the other half see an incoherent behaviour (non-verbal polarity opposite to verbal polarity), hence sarcastic.

At the end of each video, participants are asked to rate the robot using the Short Form Bem Sex Role Inventory questionnaire [13], to evaluate its character traits, such as kindness, understanding, aggressiveness. We used a 5-point Likert scale (from 1 - totally disagree, to 5 - totally agree).



(a) Incoherent vs. Coherent (Positive Response)



(b) Incoherent vs. Coherent (Negative Response)

**Figure 2:** Comparison of Incoherent and Coherent Responses

### 3.4. Results

We collected responses of 63 respondents (50 male, 10 female, no non-binary), average age was 24.3 years. We analysed only 60 questionnaires, with 3 discarded due to incorrect completion. Only 23%

had previous experiences with robots, mostly as observers. Regarding humour style: 83.3% fell into the Affiliative category, 10% into Self-enhancing, and 6.7% into Self-defeating.

The two behaviours, coherent and incoherent, were compared given a certain polarity of the customer's response, positive or negative. Full means are reported in Figure 2. Regarding the positive incoherent behaviours, we observed higher results in terms of Defence of one's beliefs (2.70 vs. 2.20), Independence (3.23 vs. 2.76), and Dominance (2.30 vs. 1.90), while the coherent mode was better in Warmth (3.30 vs. 3.00), Sympathetic (3.26 vs. 2.96), and Understanding (3.20 vs. 2.93). Regarding the negative incoherent behaviours, we observed higher averages in Strong personality (2.83 vs. 2.13), Dominance (2.40 vs. 1.93), and Assertiveness (2.90 vs. 2.40), whereas the coherent behaviour stood out in Sensitivity to the needs of others (3.46 vs. 2.96) and Compassion (3.43 vs. 3.03). In both polarities, there was a tendency to perceive the incoherent robot as more self-confident, probably because participants felt more surprised and "threatened" by the sarcastic reaction. For the incoherent configuration group alone, the evaluations between the two polarities were also compared. No significant differences were noted. A T-Test was conducted to check for statistically significant differences between the two behaviours for each item of the Short Form BSRI with relevant variance. We did not find any statistically significant difference.

## 4. Conclusion

In this study, we investigated people's perceptions of a robot's sarcastic behaviour, which have been created by contrasting incoherent behaviours. The incoherent behaviours have been presented with verbal and non-verbal cues communicating positive and negative affective expressions. Our results showed that the robot's incoherent behaviour was perceived as more self-confident and assertive compared to the coherent modality, which was rated as more warm and gentle. The differences were probably determined by the contrasts between the two attitudes. However, these results were not supported by statistical significance. Future works will test this incoherent approach in a real-bar interaction, by developing incongruency through facial emotions, and by exploring different humour styles.

## 5. Acknowledgments

This work has been supported by Italian PON R&I 2014-2020 - REACT-EU Azione IV.4 (CUP E65F21002920003, and Italian PON I&C 2014-2020 within the BRILLO research project "Bartending Robot for Interactive Long-Lasting Operations", no. F/190066/01-02/X44

## References

- [1] Claudia Di Napoli, Giovanni Ercolano, and Silvia Rossi. Personalized home-care support for the elderly: a field experience with a social robot at home. *User Modeling and User-Adapted Interaction*, 33(2):405–440, 2022.
- [2] Deirdre E Logan, Cynthia Breazeal, Matthew S Goodwin, Sooyeon Jeong, Brianna O'Connell, Duncan Smith-Freedman, James Heathers, and Peter Weinstock. Social robots for hospitalized children. *Pediatrics*, 144(1), 2019.
- [3] Leimin Tian and Sharon Oviatt. A taxonomy of social errors in human-robot interaction. *J. Hum.-Robot Interact.*, 10(2), 2021.
- [4] Charles P. Wilson. *Jokes: Form, Content, Use, and Function*. European monographs in social psychology. European Association of Experimental Social Psychology by Academic Press, 1979.
- [5] Arnie Cann, Lawrence Calhoun, and Janet Banks. On the role of humor appreciation in interpersonal attraction: It's no joking matter. *Humor-international Journal of Humor Research - HUMOR*, 10:77–90, 1997.
- [6] William P. Hampes. The relationship between humor and trust. *HUMOR*, 12(3):253–260, 1999.

- [7] Peter H. Kahn, Jolina H. Ruckert, Takayuki Kanda, Hiroshi Ishiguro, Heather E. Gary, and Solace Shen. No joking aside: using humor to establish sociality in hri. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction, HRI '14*, page 188–189, New York, NY, USA, 2014. Association for Computing Machinery.
- [8] Deborah L. Johanson, Ho Seok Ahn, JongYoon Lim, Christopher Lee, Gabrielle Sebaratnam, Bruce A. MacDonald, and Elizabeth Broadbent. Use of Humor by a Healthcare Robot Positively Affects User Perceptions and Behavior. *Technology, Mind, and Behavior*, 1(2), 2020.
- [9] Andreea Niculescu, Betsy Dijk, Anton Nijholt, Haizhou Li, and Si See. Making social robots more attractive: The effects of voice pitch, humor and empathy. *International Journal of Social Robotics*, 5:171–191, 2013.
- [10] Tony Veale. A massive sarcastic robot: What a great idea! two approaches to the computational generation of irony. In *Proceedings of the 9th International Conference on Computational Creativity, ICCO 2018*, pages 120–127, 2018.
- [11] Elizabeth E. Graham. The involvement of sense of humor in the development of social relationships. *Communication Reports*, 8(2):158–169, 1995.
- [12] Ilaria Penzo, Enrichetta Giannetti, Cristina Stefanile, and Saulo Sirigatti. Stili umoristici e possibili relazioni con il benessere psicologico secondo una versione italiana dello humor styles questionnaire (hsq) [humor styles and possible relationship with psychological well-being according to an italian version of the humor styles questionnaire (hsq)]. *Psicologia della Salute*, 2:49–68, 2011.
- [13] Namok Choi, Dale R. Fuqua, and Jody L. Newman. Exploratory and confirmatory studies of the structure of the bem sex role inventory short form with two divergent samples. *Educational and Psychological Measurement*, 69(4):696–705, 2009.