Semantic Gesture in Robot Humor: A New Dimension to Enhance the Humor Expression

Heng Zhang
Autonomous Systems and Robotics Lab, U2IS, ENSTA Paris, Institut Polytechnique de Paris
Palaiseau, France
heng.zhang@ensta-paris.fr

Chuang Yu
UCL Interaction Centre, Computer Science Department, University College London
London, United Kingdom
chuang.yu@ucl.ac.uk

Xiaoxuan Hei
Autonomous Systems and Robotics Lab, U2IS, ENSTA Paris, Institut Polytechnique de Paris
Palaiseau, France
xiaoxuan.hei@ensta-paris.fr

Adriana Tapus
Autonomous Systems and Robotics Lab, U2IS, ENSTA Paris Institut Polytechnique de Paris
Palaiseau, France
adriana.tapus@ensta-paris.fr

ABSTRACT

Humor is pervasive in our daily life. It serves not only to build rapport but also to ease tension during interactions and create stronger social interactions. In the realm of Human-Robot Interaction (HRI), humor also plays a vital role in fostering engaging and positive interactions. However, endowing robots with the ability to express humor appropriately is still a challenge. Drawing inspiration from pantomime and sign language humor, our research focuses on the role of semantic gestures in the social robot's expression of humor. In this work, we conducted an experiment in which NAO robot made humorous performances by using a series of semantic gestures. The results of the online survey show that the semantic gesture can significantly enhance the degree of funniness of the robot humor performance. Furthermore, the impact of the semantic gesture is closely tied to both the clarity of its expression and the appropriateness of the chosen semantic words.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design**; • **Computer systems organization** → **Robotics**;

KEYWORDS

Robot humor, Humor style, Semantic gesture, Nonverbal behavior

ACM Reference Format:
Heng Zhang, Chuang Yu, Xiaoxuan Hei, and Adriana Tapus. 2024. Semantic Gesture in Robot Humor: A New Dimension to Enhance the Humor Expression. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion), March 11–14, 2024, Boulder, CO, USA.* ACM, 5 pages. https://doi.org/10.1145/3610978.3640710

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA

ACM ISBN 979-8-4007-0323-2/24/03…$15.00
https://doi.org/10.1145/3610978.3640710

1 INTRODUCTION

Humor is magic. Appropriate humor in Human-Human Interactions has many benefits. It can relieve tension in a conversation[3], reduce the social distance between strangers [1], attract others' attention[10], and so on. As social robots progressively integrate into our society, there is a natural expectation for these robotic companions to acquire the capability to convey humor, thereby enriching the authenticity and long-term human-robot interactions.

However, generating meaningful and humorous content is a significant challenge, especially for a robot. Conversely, humor surpasses mere joke-telling; it is a nuanced art that demands the speaker to navigate various elements such as nonverbal cues, comedic content, context, and timing [4].

Regarding the significance of nonverbal behavior in conveying humor, valuable insights can be acquired from Charlie Chaplin's silent films. Recognized for his exaggerated body language and facial expressions, Chaplin skillfully crafted characters that could evoke hearty laughter. Upon analyzing Chaplin's silent films, his proficiency in employing descriptive gestures to clarify actions and intentions, enhancing commonplace content with a humorous dimension can be noticed [2]. Additionally, inspiration is drawn from the humor embedded in sign language within the deaf community [11]. In the domain of sign language humor, performers use not only conventional sign language but also gestures to convey unique semantics, including orientation, shape, and even vivid imitation of animal characteristics. These examples emphasize that, within the context of comedic performances, thoughtful utilization of expressive body movements contributes significantly to the overall attractiveness of the presentation.

Drawing inspiration from Chaplin's silent films and sign language humor, this paper explores a potential approach to enhance the humor performance of humanoid robots. Moreover, by appropriately employing semantic gestures, the humor expression capabilities of robots can be effectively enhanced. To test this, we designed a series of corresponding semantic gestures based on the selected jokes. We posit that embodying robots with semantic gestures will enhance the humor expression capabilities of social robots.

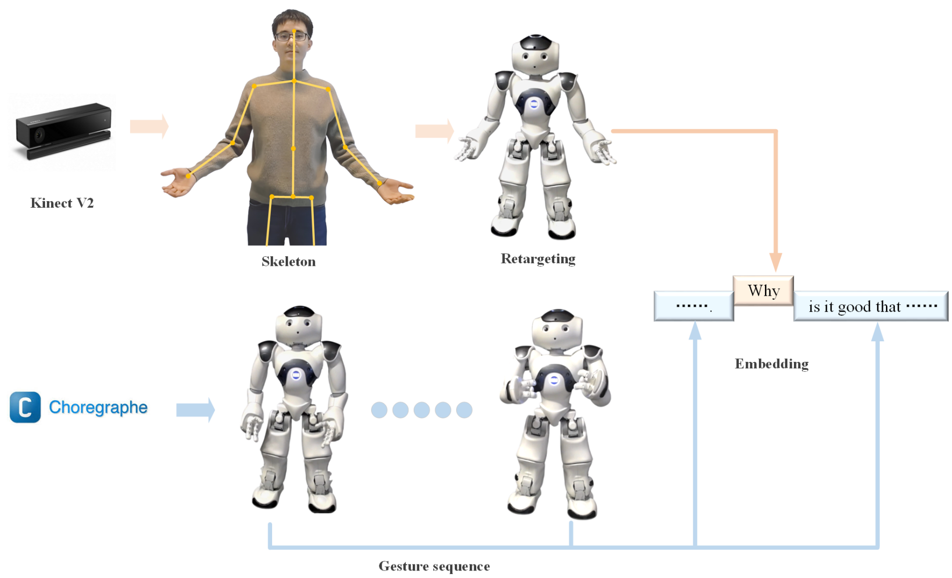

Figure 1: Pipeline of robot gestures generation and gesture embedding

2 RELATED WORKS

Humor is ubiquitous, yet it is considered the most complex cognitive attribute of humans [8]. Consequently, empowering robots with the ability to express humor is always a challenge. Stock et al. [9] argue that a system designed with humor should be able to *"recognize situations appropriate for humor, choose a suitable kind of humor for the situation, generate an appropriate humorous output, and, if there is some form of interaction or control, evaluate the feedback"*. For social robots, it is essential be able to recognize and generate humour, and adapt its behaviour based on the situation.

In[12], Weber, et al. proposed a learning-based human modeling method to shape robot humor. This method not only enables the social robot to tell jokes but also to adapt the humor style according to the user's reaction. The authors in [16] introduced a method based on big data that can recognize and generate humor for robots. As this approach can acquire extensive data from the internet, it enhances the effectiveness of existing humor generation methods, particularly in situational adaptation.

Apart from studying methods for generating humor in robots, there is also research focusing on underlying mechanisms of why and how robot humor is perceived. In [15], Heng et al. investigated how four humor styles (Affiliative, Aggressive, Self-enhancing, and Self-defeating) influence human perception of robot humor. Their findings revealed that an increase in Self-defeating humor level led to a higher perception of funniness in all types of humor content. Additionally, in [7], Niculescu et al. explored the impact of verbal features, such as voice pitch, on humor expression.

In the robotic humor literature, it is evident that the predominant focus of current research lies in the generation of humorous speech or text, often overlooking a comprehensive examination of the robot's appearance and body. Nevertheless, humor inherently involves multiple modalities and necessitates the effective integration of nonverbal cues. Although some studies [13, 14] have explored the role of nonverbal communication and the influence of gestures on perceived humor in human-robot interactions, there remains a lot of more nuanced research on the specific types of nonverbal behavior that can enhance the expression of robot humor. The present study investigates whether semantic gestures, a form of nonverbal behavior, can indeed amplify the humor expression of robots, thereby introducing a new dimension to the field of social robot humor research.

3 EXPERIMENTAL DESIGN

To investigate the influence of semantic gestures on robot humor, we used NAO robot to perform eight jokes representing four different humor types (including Affiliative, Aggressive, Self-enhancing, and Self-defeating [5]) accompanied by a set of semantic gestures. The robot's humorous performances were video recorded and integrated into an online questionnaire.

3.1 Hypotheses

For this study, the following hypotheses were formulated:

H1: Using semantic gestures by a social robot may significantly enhance the expression of humor;

H2: There is a clear correlation between the frequency of semantic gestures and the degree of amusement in humor performance;

H3: As the clarity of semantic gesture expression increases, so does the amusement derived from humor performances;

H4: The appropriateness of semantic word choices directly influences the degree of funniness in robot humor performances;

H5: The impact of employing semantic gestures in Aggressive and Self-defeating jokes surpasses their usage in the other two joke types in robot humor performance;

H6: Individuals who believe that semantic gestures enhance humor are likely to rate the robot performance significantly higher than those who hold a different opinion

3.2 Humor Performance

3.2.1 Jokes. In this study, the jokes employed were drawn from our earlier study [15], where 20 jokes were selected from the Jester joke dataset or talk-show performances. In our earlier study, we distributed an online questionnaire containing all 20 jokes, and 77 participants rated these jokes on a scale from 1 (Not funny at all) to 10 (Very funny).

For the current research, we narrowed down the selection to 8 jokes out of the original 20. The selection aimed to include 2 jokes from each humor type, while maintaining an equivalent level of funniness for each type. The average ratings of each type of joke are shown in Table 1.

Table 1: Average Ratings for Each Type of Joke

Humor Style	Joke 1	Joke 2	Average ratings
Affiliative	4.85	4.92	4.89
Self-enhancing	5.19	4.79	4.99
Self-defeating	5.03	4.96	5.00
Aggressive	5.44	4.87	5.15

3.2.2 Audio. Following the selection of jokes, we used Microsoft Azure TTS (Text to Speech) to convert the joke text to audio. All 8 jokes were written in SSML(Speech Synthesis Markup Language). By changing the parameters in the SSML script, speech attributes such as speed, pause, pitch, and timbre can be easily tuned.

3.2.3 Robot Platform. In this study, NAO robot was used to perform all 8 jokes by using designed gestures. Its physical attributes makes it well-suited for humor performance using anthropomorphic gestures.

3.3 Robot Gesture

According to David McNeill's theory[6], semantic gestures refer to gestures closely linked to speech content, serving as a complement to verbal expression. Beat gestures are rhythmic hand movements used to emphasize or synchronize with the prosody and rhythm of speech. While this research primarily focuses on semantic gestures, we also incorporated some designed beat gestures into the humor performance. This enhances the overall fluency and naturalness of the robot's performance. The next part outlines the design process for both semantic and beat gestures.

The design method is based on our previous work[14]. To ensure the clarity of semantic gestures, we opted to design the semantic gesture and beat gesture separately in this work. Specifically, we designed 8 complete beat gestures for each of the 8 joke audios. Subsequently, we designed the semantic gestures for the corresponding selected semantic words. These semantic gestures were then embedded into the appropriate positions of the 8 beat gestures.

3.3.1 Semantic Gesture. Before designing the semantic gestures, we carefully selected appropriate semantic words (or phrases) from the 8 jokes, adhering to the following criteria:

(1) The chosen words (or phrases) should be easy to describe with semantic gestures and ambiguity should be minimized;
(2) Two neighboring semantic words (or phrases) within the same joke should not be too close to each other, to avoid the robot's failure to make two complete semantic gestures in a relatively short time frame.

Following these criteria, we identified 24 semantic words and phrases. Subsequently, a volunteer performed all the corresponding semantic gestures, which were then retargeted onto the robot.

As shown in Figure 1, all 24 semantic gestures were performed by a volunteer one by one. A Kinect camera was used to capture the volunteer's actions. Subsequently, we employed Blender to edit and process the recorded FBX file. Each semantic gesture was exported as a BVH file, encompassing the positional data of all joints. Finally, we computed the joint rotation angles (pitch, roll, and yaw) of the upper body in each frame. This process enabled us to acquire the rotation angle data for the 24 semantic gestures, each stored in a separate CSV file. Some examples of semantic gestures are provided in Figure 2.

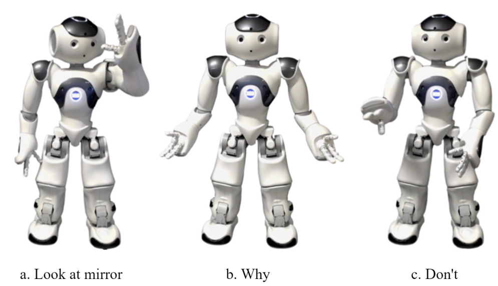

Figure 2: Examples of semantic gestures

3.3.2 Beat gesture. In this study, beat gestures were designed mainly according to the rhythm of robot's speech. To make the robot's gestures more natural, Choregraphe was used to design the robot's beat gestures.

Within Choregraphe, there exists a gesture library featuring various predefined gestures. For instance, the gesture titled *Both Arms Bump In Front* prompts the robot to raise its two small arms and wave them up and down. Additional predefined gestures suitable for use as beat gestures are available, each lasting approximately 1 to 2 seconds. In synchronization with the rhythm of each audio, we carefully selected several appropriate gestures from the library and integrated them to form a cohesive gesture sequence.

To accommodate pauses in the speech, we introduced the delay module between two predefined gestures. This adjustment enhances the overall co-speech, aligning it more effectively with the rhythm of the audio. Employing this method, we designed 8 gesture combinations for the 8 audio segments.

3.3.3 Gesture embedding. Following the acquisition of data for both semantic gestures and beat gestures, we incorporated each of the 24 semantic gestures into the corresponding positions within the 8-segment beat gestures. The embedding process is depicted in Figure 1. In order to optimize the overall fluidity of the gesture sequence, we employed interpolation on the gesture data at the embedded articulations.

3.4 Questionnaire

Upon completing the above tasks, we made the robot sequentially deliver eight jokes. Simultaneously, we recorded the entire performance and later segmented it into eight clips through editing. Subtitles were configured for each video clip. In instances where phrases were conveyed through semantic gestures, we applied a red font for emphasis. This was implemented to facilitate participants in identifying phrases aligned with the robot's semantic gestures.

The 8 clips were arranged in a randomized order within the questionnaire. Following each video clip, participants were asked to respond 6 questions corresponding to six aspects:

Q1: the rating of the funniness level of this joke;

Q2: the rating of the funniness level of the robot performance (excluding the robot gestures);

Q3: the rating of the funniness level of the robot performance (including the robot gestures);

Q4: the rating of the clarity of the semantic gesture expression;

Q5: the rating of the appropriateness of the semantic word choices

Q6: the judgment of whether or not the semantic gestures enhanced the robot's humorous expression.

With the exception of the last question, the first five questions used a 1 to 10 Likert scale for participants to express their responses.

4 RESULTS

We conducted the online questionnaire, and 40 subjects participated. The demographic profile of the participants is outlined as follows: ages ranging from 18 to 25 (9 participants), 26 to 30 (26 participants), and 31 to 40 (5 participants); 12 females and 28 males; 32 individuals reported having interacted with the social robot. We organized all the responses and extracted relevant data. Subsequently, we used the collected data to validate our six hypotheses.

As the data from Q1, Q2, and Q3 does not conform to a normal distribution, we employed the Kruskal-Wallis test for data analysis. The outcome suggests that there is no significant difference between the ratings of the joke and the ratings of the robot's performance without gestures. However, the ratings for the robot's performance with gestures are significantly higher than the first two ($H = 101.5$, $p < 0.05$). Therefore, H1 is validated.

Following that, we calculated the frequency of the semantic gesture in each performance. The Kruskal-Wallis test was used to analyze the correlation between the frequency and people's ratings of the performance (Q3). However, we did not find that these two were significantly related or conformed to a certain common trend. H2 was not validated.

For H3 and H4, we used Spearman's rank correlation coefficient to analyze the correlation between the data from Q3 and Q4 as well as Q3 and Q5. The results show that both the ratings of the clarity of the semantic gesture expression and the ratings of the appropriateness of the semantic word choices are positively correlated with ratings of robot performance ($\rho = 0.774$, $p < 0.05$; $\rho = 0.755$, $p < 0.05$). consequently, both H3 and H4 are therefore validated.

As we mentioned earlier, four humor types of jokes were used in this research. According to the previous survey, the degree of funniness of these four types jokes was rated very close to each other. However, there are significant differences in participants' ratings of these four types of humor after they were performed by the robot using semantic gestures. The results of the Kruskal-Wallis test show that the ratings of the Affiliative humor significantly higher than ratings of the other three types of humor($p < 0.05$); The ratings of the Aggressive and Self-defeating humor significantly higher than ratings of the Self-enhancing humor($p < 0.05$); No significant difference between the Aggressive and Self-defeating humor($p = 0.395 > 0.05$) was found. Hypothesis H5 was partially validated. The distribution of ratings on all four humor types of performance is shown in Figure 3.

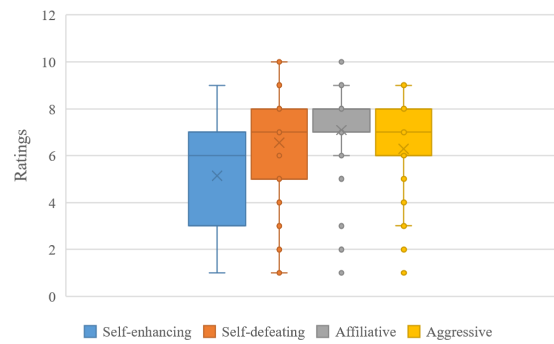

Figure 3: Ratings on different humor types of performances

We divided participants' ratings of the robot's humor performance(Q3) into two categories based on their responses in Q6. The result of the Kruskal-Wallis test show that people who think that semantic gestures enhance humor rate the robot performance significantly higher than those with the opposing view ($H = 117.7$, $p < 0.05$). Consequently, H6 is validated.

5 CONCLUSION AND FUTURE WORKS

This study focused on the influence of semantic gestures on the humor expression of a social robot. Employing the humanoid robot NAO, we designed and conducted an experiment to deliver jokes accompanied by semantic gestures. The results indicate a substantial enhancement in the perceived funniness of the robot's performance when incorporating semantic gestures.

In addition, the clarity of semantic gesture expression as well as the appropriateness of semantic word choices significantly impact the perceived funniness of the robot humor performance when using semantic gestures. Consideration of these two factors is crucial when designing a robot with humorous content. Another intriguing finding, wherein the ratings of the Affiliative humor performance significantly surpass those of the other three humor types, despite the Affiliative-type jokes receiving the lowest ratings. This pattern may not be an isolated occurrence. When jokes are executed by entities with embodiment, like humans or robots, humor types of Aggressive and Self-defeating might potentially be perceived as offensive. The above is still our conjecture, and this interesting phenomenon needs to be further investigated.

Future work will focus on the generation of robot co-speech gestures with clear and appropriate semantic gestures. The objective is for this framework to be applied in robot humor performances across various contexts.

REFERENCES

[1] Elizabeth E Graham. 1995. The involvement of sense of humor in the development of social relationships. *Communication Reports* 8, 2 (1995), 158–169.

[2] Dan Kamin. 2008. *The comedy of Charlie Chaplin: artistry in motion.* Scarecrow Press.

[3] Simon A Lei, Jillian L Cohen, and Kristen M Russler. 2010. Humor on learning in the college classroom: Evaluating benefits and drawbacks from instructors' perspectives. *Journal of instructional Psychology* 37, 4 (2010), 326–332.

[4] Rod A Martin and Thomas Ford. 2018. *The psychology of humor: An integrative approach.* Academic press.

[5] Rod A Martin, Patricia Puhlik-Doris, Gwen Larsen, Jeanette Gray, and Kelly Weir. 2003. Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of research in personality* 37, 1 (2003), 48–75.

[6] David McNeill. 2019. *Gesture and thought.* University of Chicago press.

[7] Andreea Niculescu, Betsy van Dijk, Anton Nijholt, Haizhou Li, and Swee Lan See. 2013. Making social robots more attractive: the effects of voice pitch, humor and empathy. *International journal of social robotics* 5 (2013), 171–191.

[8] Joseph Polimeni and Jeffrey P Reiss. 2006. The first joke: Exploring the evolutionary origins of humor. *Evolutionary psychology* 4, 1 (2006), 147470490600400129.

[9] Oliviero Stock and Carlo Strapparava. 2003. HAHAcronym: Humorous agents for humorous acronyms. (2003).

[10] Madelijn Strick, Rob W Holland, Rick Van Baaren, and Ad Van Knippenberg. 2009. Humor in the eye tracker: Attention capture and distraction from context cues. *The Journal of General Psychology: Experimental, Psychological, and Comparative Psychology* 137, 1 (2009), 37–48.

[11] Rachel Sutton-Spence and Donna Jo Napoli. 2012. Deaf jokes and sign language humor. *Humor* 25, 3 (2012), 311–337.

[12] Klaus Weber, Hannes Ritschel, Ilhan Aslan, Florian Lingenfelser, and Elisabeth André. 2018. How to shape the humor of a robot-social behavior adaptation based on reinforcement learning. In *Proceedings of the 20th ACM international conference on multimodal interaction.* 154–162.

[13] Cornelia S Wendt and Guy Berg. 2009. Nonverbal humor as a new dimension of HRI. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication.* IEEE, 183–188.

[14] Heng Zhang and Adriana Tapus. 2022. Preliminary Study: Robot Telling Jokes and Human Humor Factor. In *In International Workshop on Context-Awareness in HRI: Approaches and Challenges, part of HRI 2022.*

[15] Heng Zhang, Chuang Yu, and Adriana Tapus. 2022. Why do you think this joke told by robot is funny? The humor style matters. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN).* IEEE, 572–577.

[16] Dingju Zhu. 2019. Humor robot and humor generation method based on big data search through IOT. *Cluster Computing* 22 (2019), 9169–9175.