

Title	Analysis of body gestures in anger expression and evaluation in android robot
Author(s)	Ajibo, Chinenye Augustine; Ishi, Carlos Toshinori; Mikata, Ryusuke et al.
Citation	Advanced Robotics. 2020, 34(24), p. 1581-1590
Version Type	AM
URL	<a href="https://hdl.handle.net/11094/78300">https://hdl.handle.net/11094/78300</a>
rights	
Note	

***Osaka University Knowledge Archive : OUKA***

<https://ir.library.osaka-u.ac.jp/>

Osaka University



# Analysis of body gestures in anger expression and evaluation in android robot

Chinenye Augustine Ajibo , Carlos Toshinori Ishi , Ryusuke Mikata , Chaoran Liu & Hiroshi Ishiguro

To cite this article: Chinenye Augustine Ajibo , Carlos Toshinori Ishi , Ryusuke Mikata , Chaoran Liu & Hiroshi Ishiguro (2020) Analysis of body gestures in anger expression and evaluation in android robot, Advanced Robotics, 34:24, 1581-1590, DOI: [10.1080/01691864.2020.1855244](https://doi.org/10.1080/01691864.2020.1855244)

To link to this article: <https://doi.org/10.1080/01691864.2020.1855244>



View supplementary material [↗](#)



Published online: 08 Dec 2020.



Submit your article to this journal [↗](#)



Article views: 48



View related articles [↗](#)




View Crossmark data [↗](#)

FULL PAPER



## Analysis of body gestures in anger expression and evaluation in android robot

Chinenye Augustine Ajibo <sup>a,b</sup>, Carlos Toshinori Ishi<sup>a,c</sup>, Ryusuke Mikata<sup>a</sup>, Chaoran Liu<sup>a</sup> and Hiroshi Ishiguro<sup>a,b</sup>

<sup>a</sup>Hiroshi Ishiguro Laboratories, ATR, Kyoto, Japan; <sup>b</sup>Intelligent Robotics Laboratory, Osaka University, Osaka, Japan; <sup>c</sup>Robotics Project, RIKEN, Kyoto, Japan

### ABSTRACT

Recent studies in human–human interaction (HHI) have revealed the propensity of negative emotional expression to initiate affiliating functions which are beneficial to the expresser and also help fostering cordiality and closeness amongst interlocutors during conversation. Effort in human–robot interaction has also been devoted to furnish robots with the expression of both positive and negative emotions. However, only a few have considered body gestures in context with the dialogue act functions conveyed by the emotional utterances. This study aims on furnishing robots with human-like negative emotional expression, specifically anger-based body gestures roused by the utterance context. In this regard, we adopted a multimodal HHI corpus for the study, and then analyzed and established predominant gestures types and dialogue acts associated with anger-based utterances in HHI. Based on the analysis results, we implemented these gesture types in an android robot, and carried out a subjective evaluation to investigate their effects on the perception of anger expression in utterances with different dialogue act functions. Results showed significant effects of the presence of gesture on the anger degree perception. Findings from this study also revealed that the functional content of anger-based utterances plays a significant role in the choice of the gesture accompanying such utterances.

### ARTICLE HISTORY

Received 31 January 2020  
Revised 24 June 2020 and  
14 October 2020  
Accepted 12 November 2020

### KEYWORDS

Anger emotion; gesture and speech; android robot; human–robot interaction

## 1. Introduction

Human–robot interaction (HRI) is an emanating domain in robotics and has as one of its primary objectives the quest to integrate robots into human's social environment. Recent studies on HRI, have been geared towards improving the efficiency, safety and naturalness of the interaction between humans and robots [1,2]. This interaction is usually expected to take social, emotive and cognitive form. However, the dynamism in human behavior born out of human varying social–cultural orientation and concealed intentions, have successively inspired complex social-emotive behaviors in humans.

Recent studies in social sciences have shown that emotional expression serves as a vital underpin in coordinating social interaction and in shaping people's responses to their social environment [3]. Similarly, studies have also revealed that emotional expression has social effects which are dependent on the way these emotions are expressed and also on the peculiar features of the social and cultural context of the environment in which the interaction takes place [4,5].

During human–human interaction (HHI), the six basic emotions (happiness, fear, anger, disgust, surprise and sadness) [6] are usually conveyed via a synchrony of facial expressions, body gestures and speech modalities [7]. These emotional expressions have been established to aid in the formation and sustenance of long-term and intimate relationship through fostering closeness and cordiality, and also by offering consolation and eschewing social segregation [8]. According to [9], emotional expressions have both symmetric and asymmetric implications. In the 'symmetric' sense, a positive emotional expression usually results in advantageous aftermath on the expresser, while a negative expression yields unfavorable consequences on the expresser. On the other hand, in the case of 'asymmetric', negative emotional expression results in favorable outcomes on the expresser, while positive expressions have a deleterious implication on the expresser.

From the aforementioned, it is apparent that expressing negative emotion has the tendency to communicate valuable information such as concerns and situation

appraisal [10,11]. Also, expressing negative emotions could help alert other persons that a person is in danger thus evoking help and support [12]. In addition, expressing negative emotion could help strengthen relationship with others in cases where one acknowledges his/her wrongdoing and expresses regret, shame or guilt [13].

Regarding anger expression, while anger is seen as a negative emotion that facilitates the production of aggressive and contrary behavior, and is said to be destructive and detrimental when expressed during social encounters [14], it also has its symmetric and asymmetric consequences which are based on the context of interaction. Its social functions can also be adjudged advantageous or detrimental depending on the context of the expression [15].

Several efforts have been made in the bid to furnish robots (or CG agents) with the emotive expression for improved HRI, as outlined in Section 2. Most of these studies have been towards equipping social agents (robotic and non-robotic agents) with multimodal emotional expressions (positive and negative emotions inclusive) by focusing on facial expression, head movement, and utterance content [16–23], while fewer have accounted for body gestures (including hand gestures and upper-body motion) in the expression of emotions by the robot [24–29].

Numerous advances have also been made in a bid to improve on the expressive behavior of android robots, which have a high humanlike appearance. These include furnishing them with nodding, head tilting and gazing motions in synchrony with the dialogue acts [30,31], lip motion in synchrony with the phonetic contents [32], trunk motion in synchrony with the speech prosody [33] and hand gestures driven by the spoken contents and speech prosody [34,35]. Also, efforts have been made to furnish these robots with facial expression, head and torso motions in synchrony with emotive vocalized expressions like laughter [36] and surprise [37].

From the reviewed studies, it is pertinent to mention that none of the studies so far have accounted for body gestures with respect to the spoken context (e.g. the dialogue act function conveyed by the utterances) for robots.

Thus, in this study, we seek to develop humanlike anger expressions for robots, by accounting for body gestures in tandem with the dialogue acts conveyed by the robot's utterances. In order to realize this objective, we foremost identified predominant body gestures associated with anger-based utterances in HHI, and carried out analyses of the observed gestures, accounting for the perceived intensity of the expressed emotion, the predominant motion types, and the dialogue acts of the accompanying utterances. Then we implemented these gestures in

an android robot, and conducted a within-subject experiment in order to evaluate their effects on the perceived degree of anger expression and the appropriateness of the different gesture types relative to the anger-based utterances with different dialogue act categories.

## 2. Related studies

Several studies have been carried out in the last decade that are fundamentally geared towards developing emotional-based systems. In this section, we focus on previous studies including anger expression by robots.

Efforts have been made to advance the applications of emotional social agents (robotic and non-robotic) for varying HRI scenario. For instance, a robot furnished with anger-based facial expression during a collaborative task was found to have a regulatory effect in human action, and to induce human attention when an unclear condition arises [16]. Similarly, the effects of robot facial characteristics in persuasive HRI are investigated in [17,18]. It is reported that a social robot presenting more social cues will cause higher reactance and this effect is stronger when the user feels involved in the task [17]. Furthermore, a facial expression with upturned eyebrows and lips is more persuasive, evokes more trust, and less psychological reactance compared to one displaying eyebrows pointing down and lips curled downwards at the edges (facial characteristics typically not trusted in humans) [18]. Also, the influence of the robot's mood on that of the interlocutor showed that negative mood of the robot has a positive impact on the performance of participants on a given task [24].

Regarding the modalities for anger expression in robots, most of the previous studies have focused on facial expressions. The control of facial expression is found to be effective to anger expression by humanoid robots which can properly control facial motions. For example, the CG agent in [19], the GolemX-1 robotic face with 8 DOFs for the face [16], the Sociobot robot which uses an animated face through back-projection [17,18], the WE-4 robot which has 22 DOFs for the face [20], are able to express anger facial expressions which can be recognized with high accuracy, close to 100%. In contrast, other robots with less DOFs show lower recognition rates for anger expression, such as Felix robot with 4 DOFs [21], and KOBIAN robot with 7 DOFs (only 1 for the eyebrows, and 1 for the eyelids) [20]. The control of ears and comb for animal-like robots also increase the recognition of anger expression, such as the Kismet robot with 15 DOFs [22], and the robot head EDDIE with 26 DOFs [23].

Body language is exploited mostly in robots that have no or very few DOFs for facial motion, such as NAO

robot [24] and Pepper robot [25,26]. However, anger is not evaluated in these studies. Five emotions, including anger, are expressed through body gestures in the Alpha 2 robot [27]. However, among the 4 implemented gesture types for anger expression, only the pointing gesture achieved a moderate recognition rate of 75%. For the humanoids WE-4II and KOBIAN, prototypical body gestures for the basic emotions were implemented [28,29]. Particularly for KOBIAN, which have very few DOFs for the face, the inclusion of body gestures largely improved the anger recognition rate.

In the light of the reviewed studies, few or none have tried to equip robots with anger emotion by considering body gestures in context with the spoken contents. Thus, in this study, we investigate how different gesture types appear during anger expression in HHI by considering the dialogue acts conveyed by the utterances (Section 3), and evaluate how these gestures are effective when expressed by an android robot (Section 4).

### 3. Analysis of vocalized anger expression in HHI

In this section, we present analysis results on vocalized anger expression in HHI, in terms of the degree of anger and their predominant accompanying gesture types. In that regard, we adopted the MELD corpus [38] which is a multi-speaker English-based dataset characterized by daily interactions between individuals who have a close relationship. In the dataset, clips are annotated into seven different emotional classes. For the purpose of this study, we adopted the Anger class and then reviewed these samples since there were samples where the speaker expressing the anger utterance does not appear in the video. Samples that had clear audio and visual anger expressions by the speaker were selected. These selected samples (i.e. over 140 samples) formed the base dataset for anger expressions in HHI.

#### 3.1. Data annotation: predominant anger conveyor

We annotated the perceived intensity of anger conveyed in each sample and the predominant conveyor for the anger emotion expression foremost from the audio-visual clip for each sample. The intent was to establish the predominant conveyor of anger expression from visual cues (facial expression, head motion or upper-body motions).

For this study, we adopted three levels of anger degrees: Cold, Intermediate and Hot. While Cold anger have been identified to be characterized by passive anger utterances with no body gestures in most cases, Intermediate anger on the other hand are synonymous with active anger utterances and body gestures. In contrast, Hot

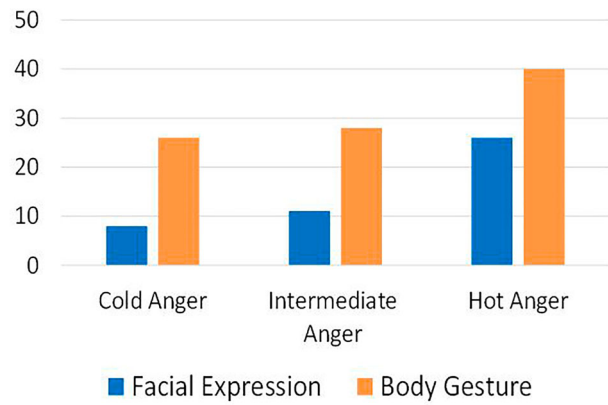


Figure 1. Perceived predominant anger conveyor.

anger are marked with very obvious aggressive utterances and body gestures.

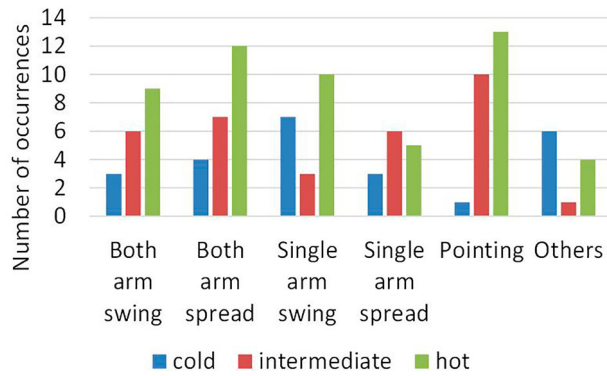
In addition, we identified body motion types associated with anger expression in HHI. From these motions, we adopted five predominant gesture types: Pointing, Single-arm Swing, Single-arm Spread, Both-arm Swing and Both-arm Spread, which have been identified in [27,34] to convey significant information during a conversation between interlocutors for this study.

We evaluated the visual clips in the dataset to establish predominant visual modality that best convey the emotion under investigation. This annotation was done by two annotators with Kappa inter-rater agreement of 0.74. Figure 1 shows the occurrences of the predominant anger conveyor for different degrees of anger. Results from the analysis revealed that the audio information accounts for about (57%) of the perceived anger from the base samples while visual information accounted for about (43%). Further analysis of the visual information conveyed in the samples showed that body motion (including hand gestures) accounts for (69%) of the observed visual information, while facial expression accounts for the remaining (31%) of the visual information conveying anger emotion. These results indicate the importance of body gestures in the conveyance of anger expression in all anger levels.

#### 3.2. Analysis of motion types for expressed anger in HHI

In this section, we analyzed the motions of the body gestures of the speakers in our base dataset. From this analysis, we established predominant gesture types and the perceived degree of anger. Analysis of the samples relative to the intensity of anger revealed the following ratio: 20%, 32% and 48% for Cold, Intermediate and Hot anger, respectively. We then analyzed each class of anger and established predominant gestures. The distributions of





**Figure 2.** Predominant motion types for varying anger intensity.

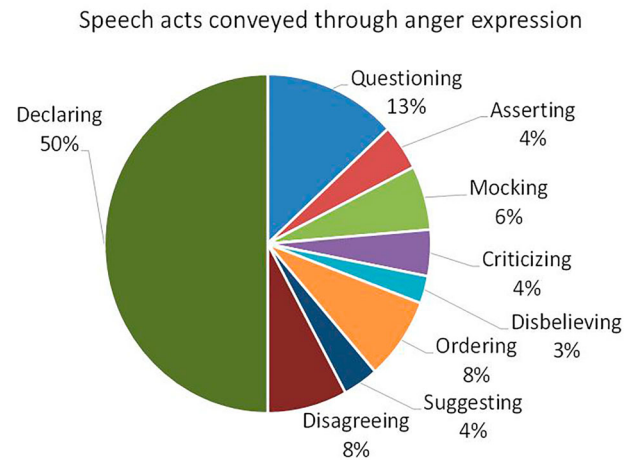
the predominant gesture types for different anger intensities are shown in Figure 2. Results from the analysis revealed that Both-arm swing and Both-arm spread gestures appear with higher occurrences in higher anger intensities (Hot anger), while the single counterparts (Single-arm swing and Single-arm spread) have more distributed occurrences among the different anger intensities. The pointing gestures appear with higher occurrences in Intermediate and Hot anger, so that one can expect it to be effective in expressing high levels of anger in HRI.

### 3.3. Analysis of expressed anger based on dialogue acts

We elaborated a list of dialogue acts associated with anger expression, as described in Table 1, by taking the studies by [39,40] as reference. On this premise, we annotated the different dialogue acts associated with expressed anger in the samples of the base dataset. We also allowed the annotator to freely include other dialogue act categories, if none of the items in Table 1 are suitable. As the number of utterances for the newly included dialogue act categories was small, they are omitted from the analysis. Figure 3 shows the distributions of the dialogue act categories conveyed through the anger expressions found in our dataset. It can be observed that the majority of the dialogue acts in the set were declarative anger-based utterances, followed by questioning, disagreeing and ordering.

## 4. Evaluation of body gestures on anger expression in an android robot

By accounting for the observations from HHI anger behavior analyses in Section 3, we designed and conducted subjective evaluation experiments using the android robot ERICA [41] as testbed, in order to investigate the effects of body gestures on anger expression. We adopted an android for the evaluation experiments, since



**Figure 3.** Distributions of dialogue act categories conveyed through anger expressions.

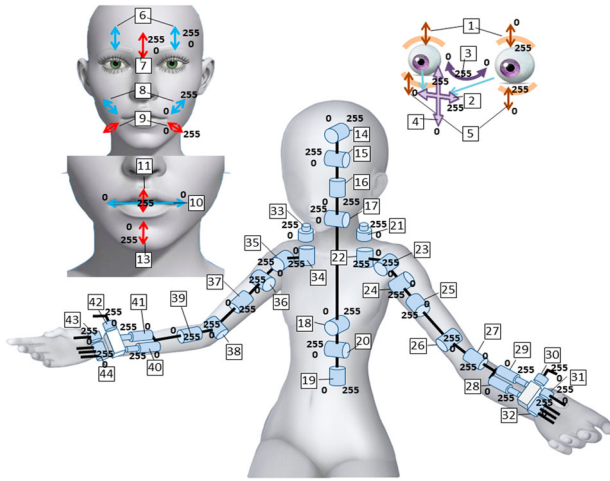
**Table 1.** Anger-based dialogue acts associated with HHI.

Dialogue acts	Description
Questioning	Entails asking someone questions or questioning something in an angry way (both tone and gesture)
Criticizing	Involves expressing judgment over someone or something in an angry manner
Declaring	Saying something in a solemn and emphatic angry manner
Asserting	Entails stating a fact or belief confidently and forcefully in an aggressive manner
Commanding	Involves dominating in an aggressive manner
Mocking	To ridicule someone in a scornful or contemptuous way
Disbelieving	Expressing disbelief in an angry manner
Ordering	Giving an authoritative instruction to do something in an aggressive manner
Swearing	Use of offensive language in an angry manner
Refusing	Indicating or showing that one is not willing to do something in an angry manner
Warning	cautionary advice in an aggressive manner
Disagreeing	Expressing a different opinion in an aggressive manner
Suggesting	To put forward for consideration in angry manner

androids have high humanlike appearance, and people expect more humanlike behaviors in comparison to other robots.

### 4.1. Description of the anger-based motion generation in the android ERICA

The android robot ERICA was adopted for this study and her motions for anger expression were manually generated with identified stroke points (emphasis/stress point) in her utterances synchronized with her respective joint actuators responsible for the desired motion pattern in each motion type. ERICA has 13 Degree of Freedoms (DoFs) for the face, 3 DoFs for the head motion, 3 DoFs for the upper-body motion, and 12 DoFs for each arm/hand as shown in Figure 4.



**Figure 4.** Joint actuators for the android ERICA.

As the purpose of this experiment is to verify how properly anger can be expressed by the different gesture types in the robot, the stroke points for each utterance were manually annotated at the beginning of the focused words. The hand/arm motions for each of the gesture types (pointing, single/both-arm spread, single/both-arm swing) were manually edited in order to reproduce the desired motions. The actuators for facial expression were adjusted to give the robot an angry look, as much as possible, under the actuator constraints. However, as there is no actuator to knit (frown) the eyebrows in ERICA, it is difficult to express anger through facial expression only. The lip motion and head motion were automatically generated based on the speech-driven methods proposed in [30,32]. The torso actuators were controlled by an idle motion generator [33]. The hand/arm actuators were controlled according to the method proposed in [35].

In the gesture generation process, the motions were generated by synchronizing the start point of the gesture stroke to the focus word starting time. In the preparation phase, the hands are moved to a high position (around

the chest/shoulder height), and the torso pitch actuator is controlled to move the upper body (a few degrees) in the backward direction. In the stroke phase, the hands are moved downward and forward stopping at a lower position (around the abdomen height), and the torso pitch is controlled to move the upper body (a few degrees) in the forward direction. Then, a holding time of 1 second is inserted after the gesture stroke, and if two focused words are closer than 1 second, the next stroke starts from the holding position, as in [35]. After the holding duration, the hands and torso are moved back to the rest position. The spread-type motions were generated by moving the hands/arms in the vertical direction, while the swing-type motions were generated by moving the hands/arms from inside to outside direction. Figure 5 shows snapshots of the created motions at the stroke phase, for each gesture type. The torso pitch actuator was additionally coordinated with the gesture stroke, so that the body slightly moves backward in the preparation phase, forward in the stroke phase, and back to the rest position in the retraction phase.

#### 4.2. Experimental setup

In this experiment, we seek to establish subjectively, the relationship between different gesture types and different dialogue acts in anger expression. Five (5) anger-based utterances from five distinct anger-based dialogue acts, namely Suggesting, Disagreeing, Questioning, Asserting and Declaring, which appeared with high occurrence rates in our base dataset, were selected for this study. Table 2 shows the transcript for these utterances. Five (5) motion types: Both-arm Spread (BSp), Both-arm Swing (BSw), Pointing (P), Single-arm Spread (SSp) and Single-arm Swing (SSw), which were found to often occur during anger expression in our base dataset, were generated for each of the utterances.

For the purpose of subjective evaluation, a total of 30 video clips were made out of which 5 had no gestures (for



**Figure 5.** Appearances of the generated gestures in the android ERICA, at the stroke phases.

**Table 2.** Transcripts and dialogue acts of the utterances used in the experiment.

Transcript	Dialogue act
I think she is trying to make something happen with you to get back at me!	Suggesting
They are not Ross, they are just saving them for the important people! What if I was the President!	Disagreeing
So, is there something fundamentally unmarriageable about me?	Questioning
David's heart will be broken; it will be too hard for them to recover from and then Phoebe will end up alone again!	Asserting
They have only been going out for a few weeks and Phoebe is a complete hung-up on Mike! She will say No!	Declaring

use as baseline for comparison) while the remaining 25 clips had both utterances and gestures. For each utterance, the 'no gesture' condition is shown in the beginning to serve as baseline for the subject judgments, followed by the 5 gesture types in a randomized order. Evaluation was conducted through an online form where subjects can view remotely and appraise the video clips based on their perception. A total of 50 subjects of American nationality recruited using Amazon mechanical Turk (AMT) (35 male and 14 female, ages between 21 and 58 years old,  $M = 35.4$ ,  $SD = 11.3$ ) took part in this experiment. The participant population was made homogeneous, since generation and perception of emotional behaviors are culture-dependent [26,29]. As the objective of the study is to verify if gestures are effective to express anger emotion, the experiment was conducted in a within-subject manner, so that all subjects watched and evaluated all video clips. This is to make it easier for participants to establish a uniform benchmark while evaluating the effects of the different gesture types relative to the no gesture condition. For the avoidance of any form of bias due to sequence effect (i.e. the judgment of participants being influenced by previously seen videos), we randomized the order of the videos during the experiment.

After watching to each clip, subjects answered to a questionnaire to grade their perceived impressions on the robot's behaviors. Regarding the assessment of the robot's negative emotion, a four-item tool (irritated, angry, annoyed, and aggravated) is used for assessing an individual's perception of the feelings of anger expressed [42], the Godspeed scale is used for assessing the personality of robots (anthropomorphism, animacy, likability, perceived intelligence and safety) in HRI [43], and a modified version of this scale is proposed to include dimensions such as emotion, extraversion and social intelligence [44]. However, as the purpose of this study is to verify the naturalness and appropriateness of gestures accompanying the spoken contents, in contrast to the effects of emotional expressions in HRI in previous

studies, we developed a questionnaire with four questions using the grading scores as described below.

- Q1. What is your perception of the degree of anger emotion conveyed in this video? (4-point scale: 0: No anger; 1: Low anger; 2: Normal anger; 3: Hot anger).
- Q2. Rate the level of synchronization between the gesture and the utterance in this clip (7-point scale: 1: Not Synchronized; 4: difficult to decide; 7: Well Synchronized).
- Q3. What was the overall degree of naturalness of the android's motion in the clip? (7-point scale: 1: Unnatural; 4: difficult to decide; 7: Very Natural).
- Q4. Rate the appropriateness of this gesture relative to the utterance (7-point scale: 1: Not appropriate; 4: difficult to decide; 7: Appropriate).

The purposes of Q2 and Q3 are for verifying if the created motions are acceptable in terms of synchrony with the utterances and human-likeness, while Q1 and Q4 are intended to compare the effects of different gesture types on anger expression degree and the appropriateness relative to the utterance type.

Finally, it is worth to mention that this experiment was reviewed and approved by the Ethical Committee of our research institute (ethical review number 20-605). Also, all the recruited subjects for this experiment gave their formal consent in accordance with the proscribed procedure of the Ethical Committee.

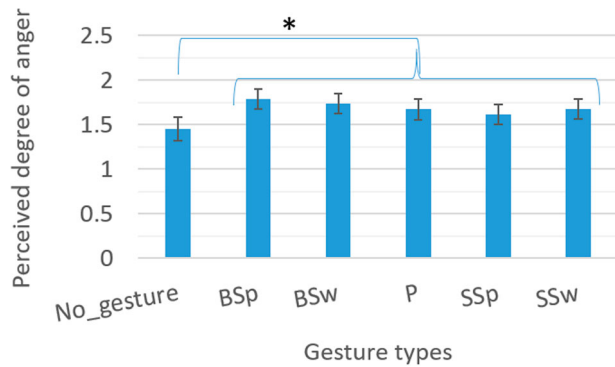
### 4.3. Evaluation results

Statistical significance tests were conducted through one-way repeated-measures ANOVA for the subjective scores. Significant differences after multiple comparisons (through Ryan's method) are shown by \* ( $p < 0.05$ ) in the figures of all results in this section.

Analysis of subjective responses for Q1 with respect to the perceived degree of anger are shown in Figure 6. It can be observed that the average scores of the generated motions are significantly higher than the one in No gesture (which is based on the impression by the voice only). These results indicate that the presence of gestures accompanied by the anger utterances are effective to increase the perceived degree of anger, for all gesture types.

The subjective analysis results for Q2 and Q3 with respect to speech-gesture synchronization and gesture naturalness (human-likeness), all average scores were between 4 and 5 (i.e. slightly synchronized and slightly natural), and no significant differences were found among the different gesture types. These results indicate that all created motions have acceptable human-likeness.





**Figure 6.** Subjective scores (average and standard errors) for the perceived degree of anger for each gesture type.

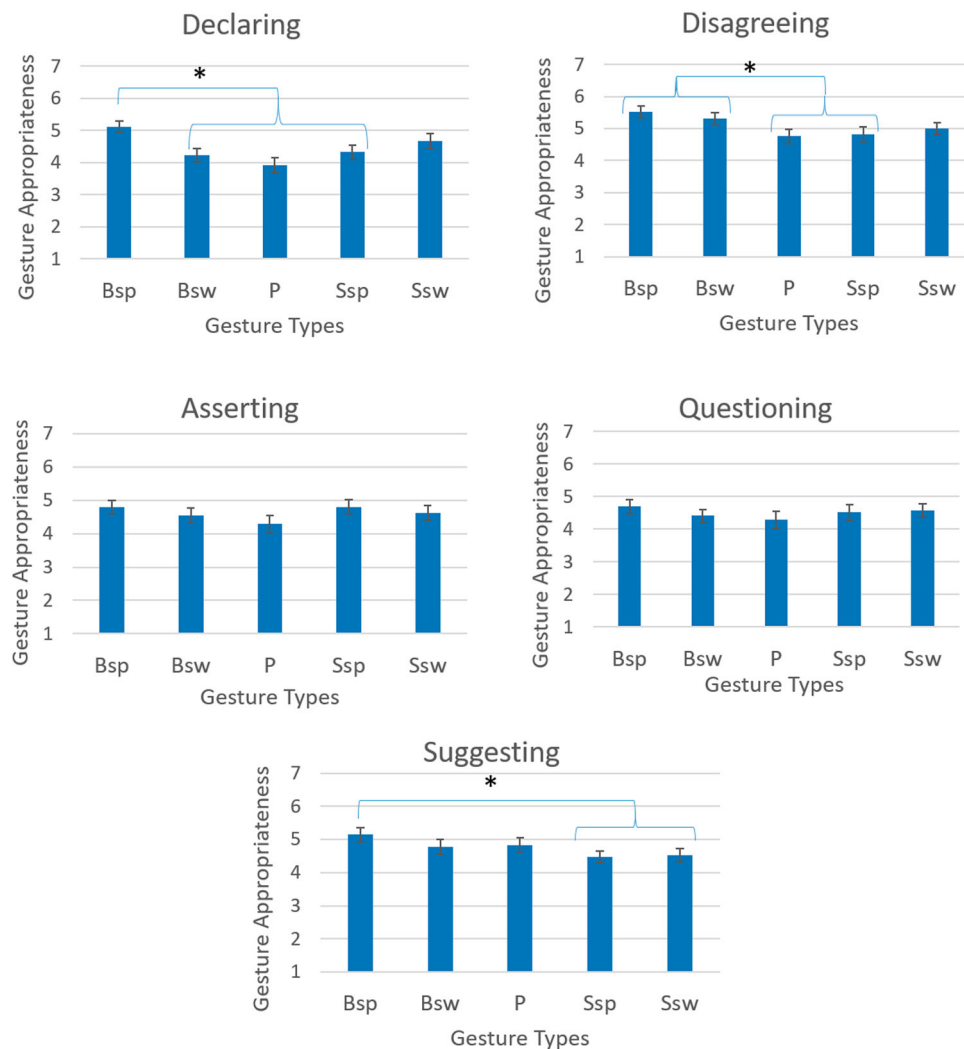
Figure 7 shows the subjective analysis results for Q4 with respect to gesture appropriateness, for each dialogue act (disagreeing, questioning, asserting, declaring and suggesting). For this analysis, statistical significance tests

were conducted through two-way repeated-measures ANOVA. Significant differences after multiple comparison (through Ryan's method) are shown by \* ( $p < 0.05$ ) in the figures.

The results in Figure 7 show different trends in the appropriateness of the gesture types for different dialogue acts.

For the Declaring utterance, a significant difference was found between Both-arm Spread and the other gesture types (Both-arm Swing, Pointing and Single-arm Spread) ( $p < 0.05$ ). These results indicate that both-arm spread gestures may be considered more appropriate for the robot when making a declaration in an angry manner.

Similarly, with regards to Disagreeing, it can be observed that Both-arm Spread and Both arm Swing are adjudged to be more appropriate than Pointing and Single-arm Spread ( $p < 0.05$ ). This basically implies that either both-arm spread or both-arm swing may also be



**Figure 7.** Subjective scores (average and standard errors) for perceived gesture appropriateness for each dialogue act.

regarded as being appropriate for the robot while expressing disagreement in an angry manner.

No significant differences have been found between the gesture types for Asserting and Questioning utterances. Finally, with regard to Suggesting utterance, the Both-arm Spread had a significant difference from Single-arm Spread and Single-arm Swing gestures ( $p < 0.05$ ). These results indicate that the both-arm spread gestures may be considered more appropriate, while single-arm gestures may be less appropriate for the robot when making a suggestion in an angry manner.

## 5. Discussion

Results from the subjective experiments conducted in Section 4 firstly revealed that the introduction of body gestures significantly improves the human perception of the degree of anger being expressed by the android robot as shown in Figure 6. It is worthy emphasizing that even under the constraints on facial expression by the current android, the body gestures were effective to increase the perception of anger degree. Nonetheless, one can expect that the inclusion of angry facial expression will reinforce the anger degree perception.

Then, we seek to establish a relationship between gesture types and dialogue acts associated with anger expression. In this regard, subjective responses showed that the both-arm spread gesture is considered the most appropriate for the majority of the anger-based utterances (declaring, disagreeing and suggesting). Differently from other gesture types, the both-arm swing gesture is also considered to be appropriate in anger-based disagreement expression. The current result suggests that when robots make anger-based utterances accompanied with both-arm spread motion, they can be easily well previewed by the human subject during HRI. Overall, the single-arm gestures were perceived as less appropriate, or with lower anger degree, in comparison to the both-arm gestures. This suggests that the level of emotion expression by the robot may be partly controlled by choice of single or both-arm gestures. The results from this experiment are thus significant to HRI, as they reveal that humans would better evaluate or be able to perceive the robot's degree of anger when the anger utterances are accompanied with appropriate gesture types. Another factor for future investigation on emotion intensity is the amplitude of the gesture motion.

Finally, the results obtained in this study may be constrained by the effects of language and culture. Both analysis and evaluation in the present study targeted Americans, while the impressions on specific motions may differ for other languages [26,29]. Nonetheless the methodology in this study can be applied to any language.

The same can be stated for the robot type. Although we conducted the evaluations using the android ERICA, the method can be applied for other robot types that can reproduce similar body gestures.

## 6. Conclusion

In this study, we were able to establish from HHI the importance of body gesture in the perception of anger expression during the conversation. We also established predominant anger gesture types and dialogue acts from HHI. On this premise, we seek to establish if the same holds for HRI.

Results from our analysis revealed that not only does gesture improve anger perception for robots but also the appropriateness of anger gesture is dependent on the functional content of the anger utterances. The results from the subjective experiments showed that the functional content (dialogue act/intention) of an utterance to an extent determines the appropriateness of some gestures over others. Findings from the study show that the both-arm spread gesture is considered more appropriate for the robot when making declaration and suggestion in an angry mood, while both-arm spread and swing gestures were considered more appropriate when expressing disagreement by the robot.

With the current drive toward integrating robots in every sphere of the society, it is essential that robots are furnished with balanced emotions (positive and negative) in order to foster better acceptance in human-dominated environment. The current results from this study is hoped to serve as a template for subsequent studies on how to furnish robots with appropriate body gestures to express negative emotion that are geared towards creating better intimacy and improved HRI.

## Acknowledgments

We thank Yuka Nakayama and Megumi Taniguchi for contributions in the data analysis, and Takashi Minato for assistance in the android hardware settings.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was partly supported by Grant-in-Aid for Scientific Research on Innovative Areas JP20H05576, and Japan Science and Technology Agency (JST), ERATO [grant number JPMJER1401].

## Notes on contributors

**Chinenye Augustine Ajib** is currently a graduate student at the Department of Systems Innovation, Graduate School of Engineering Science, Osaka University, Japan. He is also an Intern at the Hiroshi Ishiguro Laboratory at the Advanced Telecommunication Research Institute (ATR) Kyoto, Japan.


**Carlos Toshinori Ishi** received the PhD degree in Engineering from The University of Tokyo, Japan, in 2001. He worked at the JST/CREST Expressive Speech Processing Project from 2002 to 2004 at ATR. He joined ATR Intelligent Robotics and Communication Labs, since 2005, and is the group leader of the Department of Sound Environment Intelligence at ATR Hiroshi Ishiguro Labs, since 2013. He joined RIKEN Robotics Project from 2020.

**Ryusuke Mikata** received his MS degree from the Graduate School of Engineering Science Osaka University, Japan, in 2020. He made his internship at the Hiroshi Ishiguro Laboratory at the Advanced Telecommunication Research Institute (ATR) Kyoto, Japan.

**Chaoran Liu** received his PhD degree from the Graduate School of Engineering Science, Osaka University, Japan, in 2015. He is currently working at Advanced Telecommunications Research Institute. His research interests include sound signal processing and machine learning.

**Hiroshi Ishiguro** received a D.Eng. in Systems Engineering from Osaka University, Japan in 1991. He is currently Professor of the Department of Systems Innovation in the Graduate School of Engineering Science at Osaka University (2009–) and Distinguished Professor of Osaka University (2017–). He is also visiting Director (2014–) (group leader: 2002–2013) of Hiroshi Ishiguro Laboratories at the Advanced Telecommunications Research Institute and an ATR fellow. His research interests include sensor networks, interactive robotics, and android science. In 2015, he received the Prize for Science and Technology (Research Category) by the Minister of Education, Culture, Sports, Science and Technology (MEXT).

## ORCID

Chinenye Augustine Ajibo  <http://orcid.org/0000-0002-9837-9754>

## References

- [1] Lasota PA, Fong T, Shah JA. A survey of methods for safe human–robot interaction. *Found Trends Robot.* 2017;5(3):261–349.
- [2] Sheridan TB. Human–robot interaction: status and challenges. *Hum Factors.* 2016;58(4):525–532.
- [3] Van Kleef GA. The social effects of emotions are functionally equivalent across expressive modalities. *Psychol Inq.* 2017;28(2-3):211–216.
- [4] Algoe SB, Fredrickson BL, Gable SL. The social functions of the emotion of gratitude via expression. *Emotion.* 2013;13(4):605–609.
- [5] Van Kleef GA, Van Doorn EA, Heerdink MW, et al. Emotion is for influence. *Eur Rev Soc Psychol.* 2011;22(1):114–163.
- [6] Ekman P. Are there basic emotions? *Psychol Rev.* 1992;99(3):5509–553.
- [7] Levenson R. Human emotions: a functional view. In: Ekman P, Davidson R, editors. *The nature of emotion*; New York, NY: Oxford University Press 1994. p. 123–126.
- [8] Wong E, Tschann F, Messerli L, et al. Expressing and amplifying positive emotions facilitate goal attainment in workplace interactions. *Front Psychol.* 2013;4(188):1–15.
- [9] Van Kleef GA. Understanding the positive and negative effects of emotional expressions in organizations: EASI does it. *Hum Relat.* 2014;67(9):1145–1164.
- [10] Campellone TR, Kring AM. Who do you trust? The impact of facial emotion and behaviour on decision making. *Cogn Emot.* 2013;27(4):603–620.
- [11] Angie AD, Connelly S, Waples EP, et al. The influence of discrete emotions on judgement and decision-making: a meta-analytic review. *Cogn Emot.* 2011;25(8):1393–1422.
- [12] Rime B. Emotion elicits the social sharing of emotion: theory and empirical review. *Emot Rev.* 2009;1(1):60–85.
- [13] Zeelenberg M, Van der Pligt J, Manstead AS. Undoing regret on Dutch television: apologizing for interpersonal regrets involving actions or inactions. *Pers Soc Psychol B.* 1998;24(10):1113–1119.
- [14] Shahsavarani AM, Noohi S. Explaining the bases and fundamentals of anger: a literature review. *Int J Med Rev.* 2015;1(4):143–149.
- [15] Leite I, McCoy M, Ullman D, et al. Comparing models of disengagement in individual and group interactions. In: 10th ACM/IEEE International Conference on Human–Robot Interaction; Portland USA, March 2015. p. 99–105.
- [16] Reyes ME, Meza IV, Pineda LA. Robotics facial expression of anger in collaborative human–robot interaction. *Int J Adv Robot Syst.* 2019;16(1):1–13.
- [17] Ghazali AS, Ham J, Barakova E, et al. The influence of social cues in persuasive social robots on psychological reactance and compliance. *Comput Hum Behav.* 2018;87:58–65.
- [18] Ghazali AS, Ham J, Barakova E, et al. Effects of robot facial characteristics and gender in persuasive human–robot interaction. *Front Robot AI.* 2018;5(73):1–16.
- [19] De Melo CM, Carnevale P, Gratch J. The effect of expression of anger and happiness in computer agents on negotiations with humans. In: *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems*; Taipei Taiwan, May 2011. p. 937–944.
- [20] Endo N, Momoki S, Zecca M, et al. Development of whole-body emotion expression humanoid robot. In: *Proceedings of the IEEE International Conference on Robotics and Automation*; California, USA, May 2008. p. 2140–2145.
- [21] Canamero LD, Fredslund J. How does it feel? Emotional interaction with a humanoid LEGO robot. In: *Proceedings of the American Association for Artificial Intelligence Fall Symposium*, FS-00-04; Menlo Park, California, 2000. p. 7–16.
- [22] Breazeal C. Emotion and sociable humanoid robots. *Int J Hum Comput Stud.* 2003;59(1-2):119–155.
- [23] Kühnlenz K, Sosnowski S, Buss M. Impact of animal-like features on emotion expression of robot head Eddie. *Adv Robot.* 2010;24(8-9):1239–1255.

- [24] Xu J, Broekens J, Hindriks K, et al. Robot mood is contagious: effects of robot body language in the imitation game. In: *Proceedings of the International Conference on Autonomous Agents and Multi-agent Systems*; Paris France, May 2014. p. 973–980.
- [25] Tsiourti C, Weiss A, Wac K, et al. Multimodal integration of emotional signals from voice, body, and context: effects of (in) congruence on emotion recognition and attitudes towards robots. *Int J Soc Robot*. 2019;11(4):555–573.
- [26] Viet Tuyen NT, Jeong S, Chong NY. Emotional bodily expressions for culturally competent robots through long term human–robot interaction. In: *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*; Madrid, Spain, October 2018. p. 2008–2013.
- [27] Guo F, Li M, Qu Q, et al. The effect of a humanoid robot's emotional behaviors on users' emotional responses: evidence from pupillometry and electroencephalography measures. *Int J Hum-Comput Interact*. 2019;35(20):1947–1959.
- [28] Itoh K, Miwa H, Nukariya Y, et al. Mechanisms and functions for a humanoid robot to express human-like emotions. In: *Proceedings of the IEEE International Conference on Robotics and Automation*; Orlando, FL, USA, May 2006. p. 4390–4392.
- [29] Zecca M, Endo N, Momoki S, et al. Design of the humanoid robot KOBIAN-preliminary analysis of facial and whole-body emotion expression capabilities. In: *Proceedings of the 8th IEEE-RAS International Conference on Humanoid Robots*; Daejeon, Korea, August 2008. p. 487–492.
- [30] Liu C, Ishi C, Ishiguro H. Probabilistic nod generation model based on speech and estimated utterance categories. *Adv Robot*. 2019;33(15–16):731–741.
- [31] Liu C, Ishi C, Ishiguro H, et al. Generation of nodding, head tilting and gazing for human–robot speech interaction. *Int J Hum Robot*. 2013;10(1):1350009.
- [32] Ishi CT, Liu C, Ishiguro H, et al. Evaluation of formant-based lip motion generation in tele-operated humanoid robots. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*; Vilamoura, Portugal, October 2012. p. 2377–2382.
- [33] Sakai K, Minato T, Ishi CT, et al. Speech driven trunk motion generating system based on physical constraint. In *Proceedings 25th IEEE International Symposium on Robot and Human Interactive Communication*; Columbia University, New York City, USA, August 2016. p. 232–239.
- [34] Kondo Y, Takemura K, Takamatsu J, et al. Planning body gesture of android for multi-person human–robot interaction. In: *Proceedings IEEE International Conference on Robotics and Automation*; MN, USA, May 2012. p. 3897–3902.
- [35] Ishi CT, Machiyashiki D, Mikata R, et al. A speech-driven hand gesture generation method and evaluation in android robots. *IEEE Robot Autom Lett*. 2018;3(4):3757–3764.
- [36] Ishi CT, Minato T, Ishiguro H. Analysis and generation of laughter motions, and evaluation in an android robot. *APSIPA Trans Signal Inf Process*. 2019;8(6):1–10.
- [37] Ishi CT, Minato T, Ishiguro H. Motion analysis in vocalized surprise expressions and motion generation in android robots. *IEEE Robot Autom Lett*. 2017;2(3):1748–1754.
- [38] Poria S, Hazarika D, Majumder N, et al. MELD: a multimodal multi-party dataset for emotion recognition in conversations; 2018. Preprint arXiv:1810.02508.
- [39] Indrianingsih E, Ariatmi SZ, Hum M. A pragmalinguistics analysis of anger expressions in the crime movies [PhD thesis]. Universitas Muhammadiyah Surakarta; 2017.
- [40] Scarantino A. How to do things with emotional expressions: the theory of affective pragmatics. *Psychol Inq*. 2017;28(2-3):165–185.
- [41] Glas DF, Minato T, Ishi CT, et al. Erica: the ERATO intelligent conversational android. In: *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication*; Columbia University, New York City, USA, August 2016. p. 22–29.
- [42] Rains SA, Turner MM. Psychological reactance and persuasive health communication: a test and extension of the intertwined model. *Hum Commun Res*. 2007;33(2):241–269.
- [43] Bartneck C, Croft E, Kulic D. Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. In: *Proceedings of the Metrics for Human–Robot Interaction Workshop in Affiliation with the 3rd ACM/IEEE International Conference on Human–Robot Interaction*; New York United States, March 2008. p. 37–44.
- [44] Mileounis A, Cuijpers RH, Barakova EI. Creating robots with personality: the effect of personality on social intelligence. In: *International Work-Conference on the Interplay Between Natural and Artificial Computation*; Elche, Spain, June 2015. p. 119–132.