

ORIGINAL RESEARCH

Nonverbal behaviors toward an audience and a screen for a presentation by a humanoid robot

Hiroko Kamide^{*1}, Koji Kawabe², Satoshi Shigemi², Tatuo Arai¹

¹Graduate School of Engineering Science, Osaka University, Japan

²Honda R&D Co., Ltd., Wako, Saitama, Japan

Received: May 15, 2014

DOI: 10.5430/air.v3n2p57

Accepted: July 9, 2014

URL: <http://dx.doi.org/10.5430/air.v3n2p57>

Online Published: July 15, 2014

Abstract

Objective: We propose a model which predicts patterns of nonverbal behaviors for a successful presentation by a humanoid robot, especially focusing on two types of the behaviors. One is the nonverbal behaviors of eye-contact with its face and open postures with arms to keep an attention of an audience who are listening to the presentation. The other is pointing with its hand and approaching to a screen with a step to emphasize important points on the screen which is used in the presentation.

Methods: We tested the hypothesis that both types of nonverbal behaviors are effective for ensuring better understanding of the presentation. We prepared four conditions which show high or low tendency in each type of the nonverbal behaviors. A total of 139 participants observed a presentation by a humanoid robot in a between-subject design and then completed a surprised test. They also evaluated general impressions of the robot based on a psychological scale which is developed for an evaluation of humanoids.

Results: We found that both approaches are related to higher scores regarding the audience's correct understanding of the presentation, with higher psychological impressions given in relation to utility of the robot and the clearness of the voice of the robot. Additionally, we found that the behaviors toward a screen is more effective than the behaviors toward an audience in this case of the presentation by a humanoid robot.

Conclusions: We concluded that both types of the nonverbal behaviors are important for the audience's correct understanding and also the behaviors which emphasize the key points in the screen is crucial rather than the behaviors which keep the attentions of the audience when a humanoid robot gives a presentation. Finally, we discuss the universality of the proposed model for use with other humanoid robots.

Key Words: Presentation, Nonverbal behaviors, Psychological evaluation, Humanoid robot

1 Introduction

1.1 Presentation by service robots

Service robots are increasingly required to conduct social activities owing to an aging population and a decreasing labor force, particularly in Japan. To realize a robotized society in which humans and robots coexist in harmony, we need to develop new robot technologies and evaluate a robot's performance based on human impressions or feelings. This article focuses on presentation as one of the most feasible activities by robots because to give information to anonymous audience doesn't need to do flexible interaction with

humans and can be realized by pre-programmed motions.

In this study we regard the robot as a medium to convey the message of a presentation with the audience listening to the presentation. The details of our hypothesis are given in the following sections; however, active promotion emphasizing both the important words on a screen used in the presentation and the attention paid by the audience will successfully enhance the audience's understanding of the presentation. As Figure 1 shows, we focus on two important aspects for a successful presentation of Targets and Contents, therefore we propose this model as a TC model. By testing the hy-

^{*}**Correspondence:** Hiroko Kamide; Email: kamide@arai-lab.sys.es.osaka-u.ac.jp; Address: 1-3 Machikaneyamacho, Toyonaka, Osaka 560-8531, Japan

pothesis regarding the types of behaviors that allow humans to understand a core message correctly, we propose this theoretical model for effective presentation behaviors in a humanoid especially which has a head, two arms and two legs.

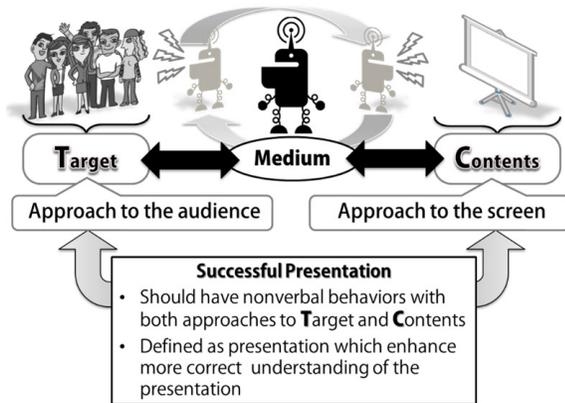


Figure 1: Nonverbal Behaviors for successful presentation by a humanoid

1.2 Related works in an engineering field

In order to realize a friendly interface, there is a widespread attention to put faces, heads, arms, and vocal systems on robots.^[1-3] Ref. 4 developed as system for an attentional control and pattern categorization for constructing attentional maps of the unknown environment and also attentional behaviors of eyes of a robot. In order to find a target from an entire image, an attention mechanism have been proposed by using a maximum-likelihood strategy.^[5] Modeling of a visual attention of robots is one of core factors to realize the effective making of decisions by robots through extracting necessary information from the environment.^[6,7] To draw human's attention to the same environment can make robots generate a situated utterance.^[8] Humans can communicate with robots which express the attentions appropriately to targets with motions of faces and perform pointing gestures.^[9] A virtual agent using speech and gestures was developed to interact with humans in a multimodal way.^[10] Ref. 11 proposed a model which make robot give a presentation using speech, gaze, and gestures during narration based on behaviors of human's presentation. From single modal to multimodal way, many technologies have been developed that can make robot convey information to humans.

Based on these developments, some robots have been introduced in our real social situation such as museums, shopping malls, and so on. The role of the robots is mainly to communicate information which humans want to know. It is also necessary to communicate with humans flexibly depending on the situation but this time it is still very difficult for robots to do it. Therefore, currently one of important aims is to realize the robot which can convey information successfully to humans. Previous works as above demon-

strated that the attention from the humans and detecting the targets in the environment are two of most significant factors in human-robot interaction. This study focuses on the combinations of these two factors. In addition, previous findings suggest that human-like behaviors generated by human-like appearance help humans to understand what robots intend to do. Therefore we focus on a humanoid robot's nonverbal behaviors during the presentation based on previous findings, especially in relation to a correct understanding and psychological impressions of an audience.

1.3 Related works in a psychological field

Teaching behavior among humans may serve as a useful reference for presentation from a humanoid to humans. Social psychological studies^[12,13] have proposed several points related to the effective teaching of nonverbal behaviors: the appearances of teachers and students, gestures and movements, facial expressions, eye contact, voice, proxemics, touching, seating position, music, and other environmental conditions. Ambady and Rosenthal^[13] claim that teachers who are evaluated highly in terms of positive impression, including familiarity, competence, and expertise, use more nonverbal behaviors such as walking around, smiling, looking at the students, nodding, and touching. In addition, they insist that the primary function of these nonverbal behaviors is to enhance the students' emotional likability toward the theme of the class and the teachers themselves. This means that teachers need to enhance their students' attention toward their presentations and point out the key points in the class using effective nonverbal behaviors.

In terms of effective concrete nonverbal behaviors, there are many factors related to teaching, as mentioned above. Here, we focus on certain factors that are adoptable to the presentations of humanoid robots. Humanoid robots have two arms and two legs just like humans, and, therefore, we can try to realize as many nonverbal human behaviors as possible. As stated in ref. 12, 13, eye contact with an audience seems to be important in involving an audience into a presentation. An opening of the speaker's arms to the audience is also useful in keeping their attention.^[12] To emphasize key points in the presentation, we make the robot come closer to a screen and point the important words on the screen. In the context of this study, touching a listener is physically dangerous, and we therefore do not consider such behavior at this time. In addition, facial expressions are not manipulated as the humanoid robot used in this study cannot show such expressions. Other conditions such as voice, music, and environmental issues are controlled. Finally, the seating position of the audience is taken into consideration in this analysis.

Finally, we describe below the effective nonverbal behaviors for successfully conveying a main message to an audience. We view the robot as a medium to combine the message on the screen with the audience itself. We then focus on two

aspects: an approach for maintaining audience attention and involving them in the presentation, and an approach for emphasizing an important message displayed on a screen. We manipulated both approaches and hypothesized that taking both approaches enhances correct understanding in the audience more than taking no approach (see Figure 2, Cond. *TC* vs. Cond. *tc*). We then explored which approach is more effective for a correct understanding (Cond. *Tc* vs. Cond. *tC*).

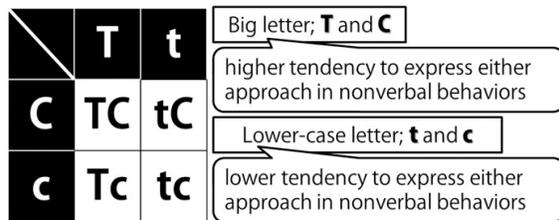


Figure 2: Two sides of presentation behavior (*TC*, *Tc*, *tC*, and *tc* means a combination of Target-oriented/Contents-oriented behaviors) and a Proposed TC model and four comparable conditions in this study

1.4 Effect of presentation order

It is well known in the social psychological field that there are two biases by orders of presentation that are effective in helping an audience memorize information. One order utilizes a primary effect in which people memorize more information appearing in the initial phase of the presentation than in the latter phases.^[14–16] Conversely, the recency effect states that people retain more information from the most recent phase than from the previous phases. Although the time span of the presentation is not specified, the beginning and end sections of the presentation tend to be better remembered overall than the middle phases. We deemed it necessary to consider this order effect in this study, and we divided the presentation into 3 phases.

2 Methods

2.1 Participants and presentation

A total of 139 people from their 10s to 60s (mean age, 39.57; SD = 15.85; 67 male and 72 female) participated in this experiment. We used ASIMO as an example of a humanoid robot. The presentation for phase 1 is the self-introduction of the robot, phase 2 is the presentation of a “Beach Clean Activity” by Honda regarding the use of machines to clean a beach, and phase 3 is made up of other social activities by Honda regarding supporting of child education and environment conservation. The overall length of the presentation was about 10 min, and each phase lasted around 3 min. The contents of the presentation are from Honda’s standard robot demonstration.

2.2 Measurements

In this experiment, we measured the audience’s understanding of the presentation, their impressions of the robot after observing the presentation, and their buccal secretions, which were taken three times (before the presentation, after observing of the presentation, and after the final phase of the presentation). However, buccal secretion is not the main issue of this study, and we therefore present these results in another report. Finally, the experiment was videotaped. We will additionally investigate eye-contact toward the robot, the screen, and elsewhere in relation to the understanding of the presentation.

2.2.1 A surprised test for correctness of understanding of an audience

To test correct understandings of participants, it is important to control residual variables as much as possible. We didn’t make the announcement to participants that they would be tested about the contents of the presentation. They were just told to observe the performance of a robot and to complete some questions about subjective impressions of the robot. To give a surprised test after the observation without any notice, we measure the effect of nonverbal behaviors on their correct understanding of the presentation comparing four conditions.

The test used for measuring the audience’s understanding of the presentation is composed of 11 questions, 2 questions for phase 1, 7 questions for phase 2, and 2 questions for phase 3, which are as follows: (1) the reason for the size of the robot, (2) sensor placement for avoiding obstacles, (3) start time of the Beach Clean Activity, (4) number of beaches cleaned under the Beach Clean Activity, (5) persons or groups executing the Beach Clean Activity, (6) activity name, (7) how to perform the Beach Clean Activity, (8) the reason for conducting the Beach Clean Activity, (9) Honda’s hopes regarding the outcome of the Beach Clean Activity, (10) reasons for saving the environment, and (11) Honda’s hopes regarding supporting of child education. (6) and (9) were free descriptions and others are selective questions from 3 to 5 items.

2.2.2 Psychological impressions to a robot

Although the main variable of this study is audience understanding, their psychological impressions were also measured to determine their perception of the robot’s presentation. We used a PHIT-24,^[17] which is composed of 9 dimensions, to evaluate the general impressions of the humanoid robots from the viewpoint of an ordinary user. Ref. 17 revealed basic perspectives which are used by ordinary people to evaluate humanoid robots then developed the humanoid-oriented psychological scale to quantify general impressions based on the perspectives. The basic dimensions of the scale are Familiarity, Repulsion, Utility, Performance, Motion,

Voice, Sound, Humanness, and Agency (see Table 1). A PHIT-24 has 24 items and uses a 7-point Likert scale of 1 (strongly disagree) to 7 (strongly agree).

Table 1: Nine factors of the psychological impressions

Category	Contents of descriptions
Familiarity	General likes of robots, motions, and designs
Repulsion	Anxiety or sense of aversion about humans being replaced by robots in terms of work or existence itself
Utility	Clear aims of robots' work, usefulness, and cost
Performance	High performance including interactions, sensations, and intelligence
Motion	Smooth, natural, and dynamic motions or stability and balance of motions
Voice	Clearness of voice
Sound	Soft sound of motors
Humanness	General humanness including designs, motions, and voice
Agency	Own will and mind

We asked about both the audience interest in the content of the presentation and their advance knowledge regarding the content, as both are thought to have a relationship with an understanding of the presentation. The question regarding audience interest in the content focused on how much they usually consider environmental issues such as beach cleaning activities. The question regarding their advance knowledge of the content focused on how much they are already familiar with Honda's Beach Clean Activity.

2.2.3 Manipulation check

To test manipulation of nonverbal behaviors (explained below), based on a 7-point scale, we asked the participants (1) their perception of the extent of the robot's eye contact with them, (2) their perception of the extent of the robot's eye contact with the audience in general, (3) extent of robot's eye contact with the screen, and (4) accuracy of the robot in pointing important words on the screen. Questions (1) and (2) were used to determine the manipulation approach with the audience, and (3) and (4) were used to determine such approach for the screen.

2.3 Manipulation of nonverbal behaviors

To realize four conditions (*TC*, *Tc*, *tC*, and *tc*) which reflect high and/or low tendency in both behavioral approaches toward the audience of the Target and the Content of the presentation, we manipulated the nonverbal behaviors as follows.

We manipulated the high tendency of the behaviors oriented to the target (*T*) as to indicate eye contact with the audience and an open posture. The behaviors were distributed to six participants evenly (see Figure 3). During the explanation of the contents, the robot looked at the participants in sequence and extended its arm and hand then spread it

wide toward the all participants (1 and 2 in Figure 4). The high tendency of the behaviors oriented to the content (*C*) was deemed as pointing to important words on the screen and making eye contact with these words and also going a single step toward the words on the screen. The robot took another stride toward the screen and showed the direction of the key words with the hand looking at the screen (3 and 4 in Figure 4).



Figure 3: Observation location used for the robot's presentation

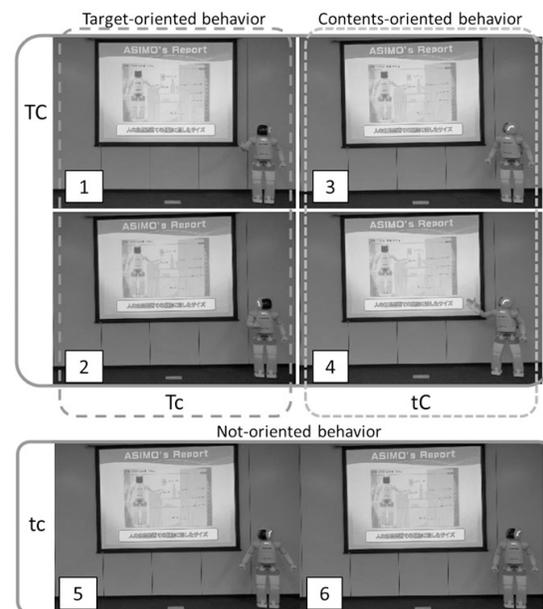


Figure 4: Nonverbal behaviors in each condition

In the *TC* condition, the robot did all behaviors from 1 to 4 in Figure 4. In case of the *Tc* condition, the robot did 1 and 2 in Figure 4 and during the time the robot of the *TC* condition would do 3 and 4, the robot faced forward but only repeated motions such as head rolls to the right or left, opening and closing of straight arms right next to the body, rolling the waist to the right or left. The robot of the *tC* condition did 3 and 4 in Figure 4 and during the time the robot of the *TC* condition would do 1 and 2, the robot faced at the screen but only repeated motions of head rolls, opening

and closing of straight arms, rolling the waist, as in the *Tc* condition.

The point for the comparison is not the difference in the quantity of the behaviors, but in the quality of behaviors. Therefore, to make behaviors of the *tc* condition which should reflect no orientations either to the target and the content, we use the same parts of the body. The motions are made not to reflect meaning such as 1, 2, 3 and 4 in Figure 4. The robot looked between the audience and the screen and repeated motions such as head rolls to the right or left, opening and closing of straight arms right next to the body, rolling the waist to the right or left, and opening and closing of the legs (5 and 6 in Figure 4).

2.4 Procedure

One experiment included 6 participants, each sitting in a box separated by partitions (see Figure 3). The facilitator of the experiment explained the entire procedure, and the participants signed a letter of consent. After signing, they gave a sample of their buccal secretion using a special tip. This took less than one minute. They observed the robot's presentation and gave another buccal secretion. They then completed a surprised test of 11 questions regarding the presentation. The time given for answering this test was limited to four minutes. After answering the test, they completed a questionnaire, which included the PHIT-24 and their interests in the content of the presentation and advance knowledge regarding the content. Finally, they gave a third buccal secretion sample.

3 Results

3.1 Manipulation check

Two items for manipulation check, approaching the audience and the screen, were averaged. To investigate the manipulation of nonverbal behaviors among the conditions one-way ANOVA was conducted for both variables. A significant difference among the conditions for approaching the audience was found ($F(3, 135) = 34.34, p < .001$).

P value is less than .001 and F value is large enough. It was concluded that there is a significant difference among the four patterns of presentation. Then in order to reveal which pair of four patterns significantly differ from each other, we conducted Tukey's Honest Significant Difference test. The results are described in Figure 5.

Cond. *TC* is recognized as a greater approach toward the audience than Cond. *tC* and Cond. *Tc*, while Cond. *Tc* is recognized as a greater approach toward the audience than Cond. *tC* and Cond. *tc*. This means that the manipulation is successful for approaching the audience. Cond. *tc* is recognized as a greater approach toward the audience than Cond. *tC*. The repeated motions of the robot's face in Cond. *tc* could be viewed as relative eye contact.

A significant difference among the conditions for the manipulation in approaching the screen was also found ($F(3, 135) = 20.49, p < .001$). Tukey's Honest Significant Difference test also showed significant differences, which are given in Figure 6. Cond. *TC* is recognized as a greater approach toward the screen than Cond. *Tc* and Cond. *tc*. Cond. *tC* is recognized as a greater approach toward the screen than Cond. *Tc* and Cond. *tc*. This means that the manipulation used for approaching the screen was also successful.

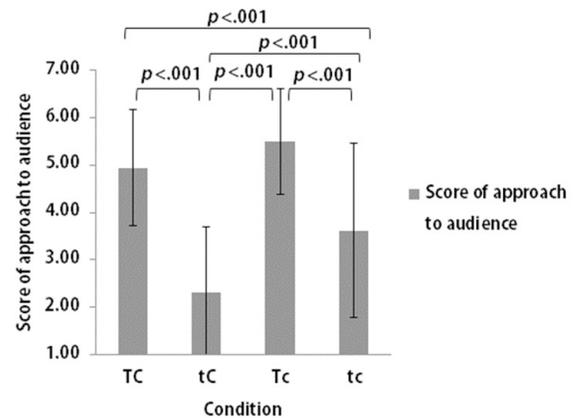


Figure 5: Manipulation check for approaching audience

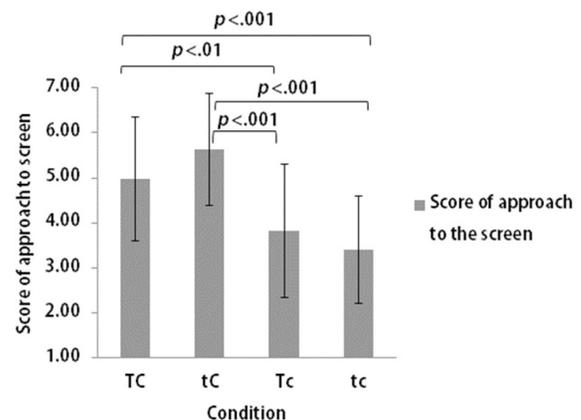


Figure 6: Manipulation check for approaching screen

3.2 Correlation between test regarding audience understanding of presentation and other variables

Four variables were used for the test: the number of correct answers for all 11 questions, the correct rate of phase 1, correct rate of phase 2, and correct rate of phase 3. The placement of the seat was also considered in the analysis. The mean score and SD are shown in Table 2, and the correlation was calculated as shown in Table 3.

Table 2: Mean score and SD

Variables	M	SD
Placement of seat	1.96	0.81
Number of correct answers	6.94	1.48
Correct rate of phase 1	0.83	0.26
Correct rate of phase 2	0.50	0.17
Correct rate of phase 3	0.89	0.22
Interest in the content	3.34	1.82
Knowledge about the content	3.55	1.71

Note. Score of seat placement ranging from 1 to 3, where a higher score indicates lesser distance from the center of the 6 seats.

Table 3 shows that older people had fewer correct answers, and participants who were sitting on the relative center of six seats had more correct rate in phase 2. The significant correlation between knowledge regarding the content and the number of correct answers or correct rates of phases 2 and 3 indicates that participants who were more familiar with the content in advance paid less attention to the presentation.

Table 3: Correlation among variables

	Age	PS	N of CA	CR of P1	CR of P2	CR of P3	I	K
Sex	.00	-.08	-.12	-.03	-.08	-.15 [†]	.08	-.13
Age		-.08	-.15 [†]	.02	-.21 ^{**}	.03	.09	-.07
PS			.12	.00	.16 [†]	-.02	.06	.02
N of CA				.49 ^{***}	.88 ^{***}	.42 ^{***}	-.10	-.20 [*]
CR of P1					.14 [†]	.09	-.13	.01
CR of P2						.13	-.06	-.19 [*]
CR of P3							-.01	-.15 [†]
I								-.28 ^{**}

Note. [†]p<.10, *p<.05, **p<.01, PS is "Placement of seat," N of CA is "Number of correct answers," CR of P* is "Correct rate of Phase*," I is "Interest in the content," and K is "Knowledge about the content."

3.4 Test of audience understanding of the presentation among the conditions used

We found significant correlations between the number of correct answers and age and knowledge of the content, and therefore to control these variables, an analysis of covariance was performed for the number of correct answers. We found no significant difference.

An analysis of covariance was performed for the correct rate of phase 1, which also showed no significant difference.

3.3 Order effect

We conducted an analysis of covariance for the correct rate to investigate the effect of phase control in terms of sex, age, seat placement, and knowledge about the content, which are correlated with the correct rate. We found a significant order effect ($F(1.85, 248.10) = 4.91, p<.01$) and the differences in Fisher's Least Significant Difference Test are indicated in Figure 7.

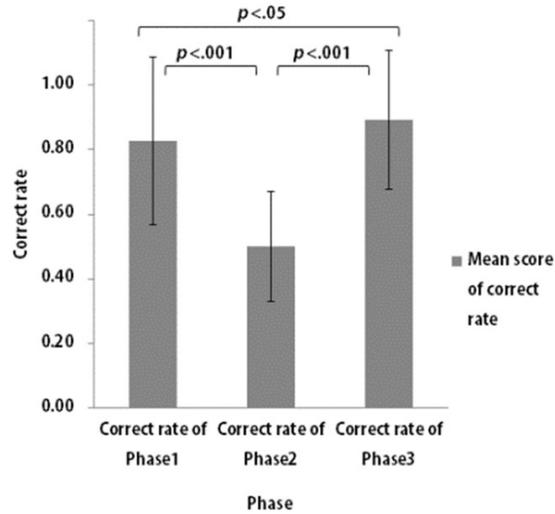


Figure 7: Correct rate during each phase

We found significant correlations between the correct rate in phase 3 with sex and knowledge about the content, and therefore to control these variables an analysis of covariance was performed for the correct rate of phase 3. We found no significant difference.

For the correct rate of phase 2, there is a significant correlation with age, seat placement, and knowledge about the content. Controlling these variables, we performed an analysis of covariance for the correct rate of phase 2 and found

a significant difference among the conditions ($F(3, 132) = 3.19, p < .05$). The differences from Fisher's Least Significant Difference Test are shown in Figure 8.

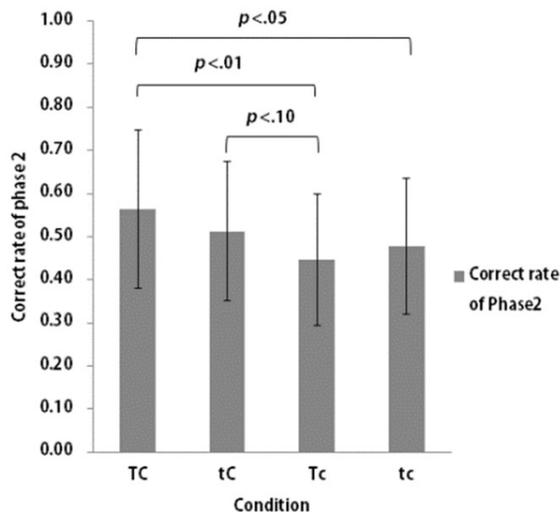


Figure 8: Correct rate for phase 2

Figure 8 shows that Cond. *TC* had a significantly higher rate of correct answers than Cond. *tc*. This means that approaching both the audience and the screen are significantly effective for audience understanding during the middle of the presentation compared with when neither approach was used. Cond. *TC* then had a more correct rate than Cond. *Tc*. This means approaching the screen and emphasizing important points has an effect on the audience's understanding of the presentation when a robot gives a presentation. The difference between Cond. *tC* and Cond. *Tc* suggests that approaching the screen is more important than approaching the audience in terms of making the audience understand the presentation correctly.

As psychological studies predicted,^[14-16] there are two types of order effects and participants tended to maintain memories of both beginning and ending of the presentation. The significant difference in phase 2 means that the both approaches toward the audience and the screen have effects on the middle time of the presentation when people tend to lose their attention to the presentation.

3.5 Correlation among psychological impressions, sex and age

There are nine dimensions regarding the psychological impressions of the audience. The placement of the seat is also considered in the analysis. We assumed that a subjective evaluation of the presentation would affect the viewer's psychological impressions. The mean score and SD of the 9 dimensions and subjective evaluation are shown in Table 4, while the correlation was calculated as shown in Table 5.

Table 4: Mean score and SD

Variables	<i>M</i>	<i>SD</i>
Familiarity	5.75	1.08
Anxiety	2.56	1.41
Utility	5.62	1.27
Performance	5.70	1.02
Motion	4.60	1.27
Voice	4.88	1.52
Sound	4.61	1.68
Humanness	3.54	1.42
Agency	3.16	1.62

Table 5: Correlation among variables

	Sex	Age
Familiarity	.13	.003
Anxiety	-.04	-.32***
Utility	.16 [†]	.04
Performance	.12	-.21*
Motion	.14	.03
Voice	.15 [†]	.06
Sound	.24**	.22**
Humanness	.01	.11
Agency	.30***	-.20*

Note. [†] $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 5 shows that sex and age are significantly related to psychological impressions, as well as to the audience's understanding of the presentation. In following analyses, we controlled variables which correlate with factors of impressions.

3.6 Psychological impressions

We controlled the variables that are significantly related to the psychological impression of each dimension according to the results in Table 5, and conducted an analysis of covariance for the scores of psychological impressions for the nine dimensions. We found a significant tendency toward a difference among the conditions for Utility and Voice, and no significant differences in the other dimensions.

For Utility, there is a significant tendency toward a difference among the conditions ($F(3, 133) = 2.439, p < .10$). The results of a Fisher's Least Significant Difference Test, shown in Figure 9, highlight these differences.

Utility means how much the robot is useful as an instrument from user's perspective. Cond. *TC* was evaluated higher in terms of Utility than Cond. *tc*. This means that approaching both the audience and screen are significantly effective for impressions of Utility compared with when neither approach is used. Thus, Cond. *TC* has a higher evaluation of Utility than Cond. *Tc*, and Cond. *tC* has a higher evaluation of Utility than Cond. *tc*. This means that both approaching the audience and the screen make the robot evaluated as more useful in terms of subjective impressions. The

difference between Cond. *tC* and Cond. *Tc* suggests that approaching the screen is more important than approaching the audience in terms of a higher evaluation of Utility.

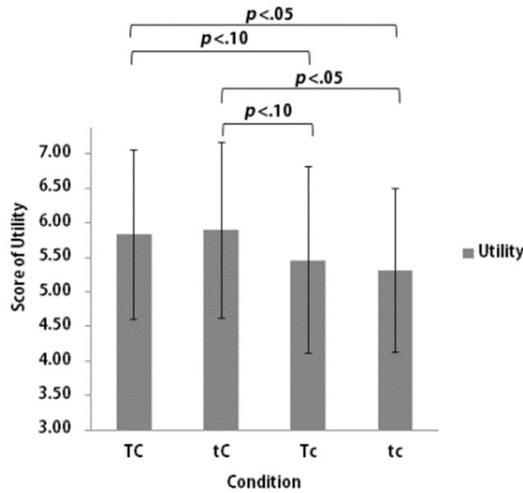


Figure 9: Psychological impression of Utility for each condition

For Voice, there is a significant tendency toward a difference among the conditions ($F(3, 133) = 2.38, p < .10$). The results of a Fisher’s Least Significant Difference Test, shown in Figure 10, highlight these differences. Voice reflects the impression of clearness or understandability of the robot’s voice. Cond. *TC* is evaluated higher than Cond. *tc* in terms of Voice. This means that approaching both the audience and the screen is significantly effective for impressions regarding the Voice of the robot compared with when neither approach is used. Cond. *tC* has a higher evaluation of Voice than Cond. *tc*. This means that approaching the screen was deemed to making it easier for the audience to understand the robot’s voice.

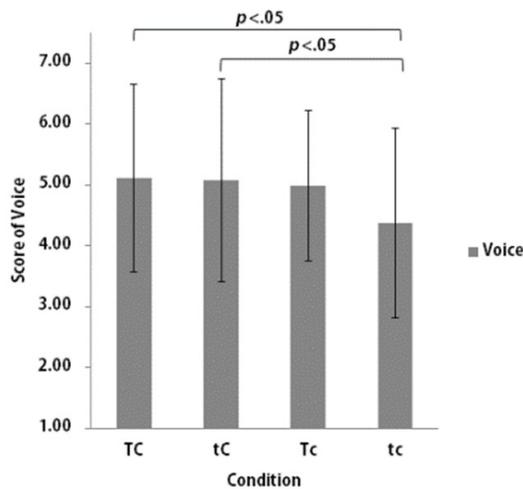


Figure 10: Psychological impression of Voice for each condition

Although we used the same voice for every condition, the audience impression of the robot’s voice is different among the conditions. This implies that an understanding of what the robot is saying is relatively dependent on nonverbal behaviors. In particular, the nonverbal behaviors under Cond. *tc* are simply repetitions of meaningless behaviors, and the timing of these behaviors do not match with what the robot is saying. Considering the Voice results, it seems important to match nonverbal behaviors with the spoken lines for a better understanding of the robot’s voice.

The other dimensions did not show any significant differences; however, we dealt only with the presentation, which was the robot’s task. Considering social or business aspects of the presentation, other dimensions such as familiarity or humanness did not seem to be affected by the conditions used in this study. Performance was thought to be related to the conditions; however, as the robot we used is quite well known, there seems to be no significant difference among the conditions.

3.7 Eye-contact from participants

Additionally, we investigate the time of eye-contact toward the robot, the screen, and elsewhere from the participants. We aim to clarify the effects of conditions on the eye-contact of participants and also relationships between the correct rate of the surprised test and participants’ eye-contact.

We counted the time of eye-contact toward the robot, the screen and elsewhere from the videotape. We found significant correlations between the time of eye-contact and sex, and therefore to control the variable, an analysis of covariance was performed for the time of eye-contact toward the robot, the screen and the elsewhere. We found no significant difference for the time of eye-contact toward the screen and the elsewhere but the robot ($F(3, 36) = 6.25, p < .01$).

As Figure 11 shows, Cond. *TC* and *Tc* have longer time of eye-contact toward the robot from participants than Cond. *tC* and *tc*. It seems that the robot’s nonverbal behaviors related to the approach to the audience enhance the eye-contact from participants toward the robot. However, Figure 8 revealed that Cond. *tC* has higher correct rate of the test than Cond. *Tc*. It is possible that too much nonverbal behaviors of a robot that enhance attentions of the participants gather their attentions not to the screen but to the robot. When a robot gives a presentation, such behaviors might prevent participants from understanding the presentation. This result suggests that nonverbal behaviors related to the screen is more helpful for the audience to understand the presentation in case that a robot convey information to humans.

Then, we analyze the correlation between the eye-contact and the correct rate of the surprised test (see Table 6). The result of correlation didn’t show the significant relationships between the eye-contact to the screen and the correct rate. But the eye-contact to elsewhere tend to decrease the num-

ber of correct answers and the correct rate of phase 1. It seems important to keep attention to the place of presentation and also as seen above, the attention to the screen seems more important to the correct understanding considering the result of Figures 8 and 11.

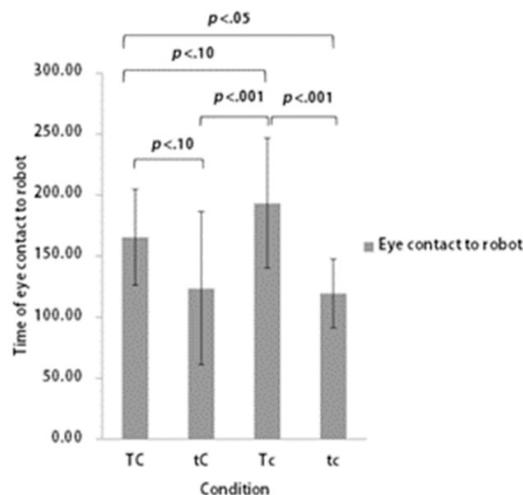


Figure 11: The time of eye-contact toward the robot for each condition

Table 6: Correlation among variables

	Eye contact to the robot	Eye contact to the screen	Eye contact to elsewhere
Number of correct answers	-.02	-.15	-.27 [†]
Correct rate of phase 1	-.06	-.10	-.27 [†]
Correct rate of phase 2	-.02	-.11	-.21
Correct rate of phase 3	.04	-.09	.06

Note. [†] $p < .10$

4 Discussion

This study focused on a presentation given by a humanoid robot as an example of social activity, and tested the hypothesis that approaching both the audience and the screen are effective for understanding of the content of a presentation. Considering the order effect of memories, we found that both approaches are related to a higher rate of correct answers regarding the presentation and a higher psychological evaluation in relation to Utility and Voice.

The results support the effectiveness of the TC model described in this study (see Figure 12). This model sees the robot as a medium connecting the audience with the content of the presentation. It is important for this medium to connect them both smoothly and successfully. Therefore,

we compared approaching the audience with approaching the screen at high and low levels, and determined that both approaches can enhance the correct understanding for the audience. Additionally, both provide a higher evaluation of the robot’s utility for humans and make it is easier for audiences to understand what the robot is saying.

As psychological studies predicted, participants in this study also showed both primacy and recency effects on memories of the presentation. The results revealed that the proposed model in this study is effective in the middle time of the presentation when people tend to divert their concentrations. These order effects are universal phenomenon for humans, therefore it is important that results proved the effectiveness of the model in the time when people decrease their concentrations.

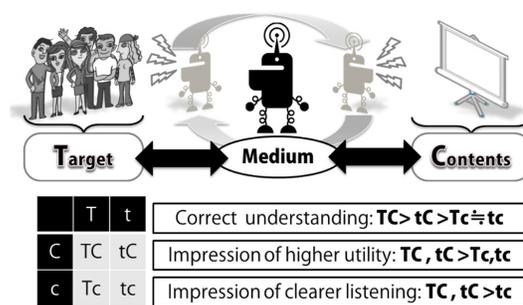


Figure 12: Proposed TC model of nonverbal behaviors in Humanoid

The results between Cond. *tC* and Cond. *Tc* suggest that approaching the screen is more important for a successful presentation than approaching the audience. Humans might pay attentions to a robot than when humans (teachers) give presentations, considering the robot is still new to them. It is supposed to more effective to emphasis the contents on the screen rather than to keep attentions of the audience in case that a robot gives a presentation. In addition, the effectiveness of presentation by robots should be validated through the comparison between robot’s presentation and human’s presentation. This is because this study regards the robot as the replacement of the humans. We haven’t compared the human-human interaction with this study yet but a future research will reveal this difference.

Additionally we used only one robot as the first step. It is necessary that an investigation into the TC model be applicable to other humanoid robots. There is also the room for investigating the possibility to expand this model to other devices such as big arrows on the screen to emphasize the important words or flashing lights to keep the attentions of the audience. This study deals with TC model as a psychological model but it is significant to reveal generality of TC model for such other devices and robots.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 26705008 and 25119505.

References

- [1] T. Kanda, T. Miyashita, T. Osada, Y. Haikawa and H. Ishiguro, Analysis of Humanoid Appearances in Human-robot Interaction, *IEEE Trans. on Robotics*. 2008; 24: 725-735. <http://dx.doi.org/10.1109/TRO.2008.921566>
- [2] T. Kanda, H. Ishiguro, M. Imai, and T. Ono, Development and Evaluation of Interactive Humanoid Robots, in *Proc. the IEEE (Special issue on Human Interactive Robot for Psychological Enrichment)*. 2004; 1839-1850.
- [3] S. Negi, T. Arai, K. Inoue, Y. Ujiie, T. Takubo, Psychological Assessment of humanoid Robot Appearance and Performance Using Virtual Reality, in *Proc. 17th IEEE International Symposium on Robot and Human Interactive*. 2008: 719-724.
- [4] Garcia, L.M., Oliveira, A.A.F., Grupen, R.A., Wheeler, D.S., Fagg, A.H., Tracing patterns and attention: Humanoid robot cognition, *IEEE Intelligent Systems*. 2000; 15(4): 70-77. <http://dx.doi.org/10.1109/5254.867915>
- [5] H. Tagare., K. Toyama., Jonathan G. Wang, A Maximum-Likelihood Strategy for Directing Attention during Visual Search, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2001; 23(5): 490-500. <http://dx.doi.org/10.1109/34.922707>
- [6] L. Itti, C. Koch, e. Niebur, A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1998; 20(11): 1254-1259. <http://dx.doi.org/10.1109/34.730558>
- [7] Vijayakumar, S., Conrad, J., Shibata, T., Schaal, S, Overt visual attention for humanoid robot, *Proc. of Int. Conf. on Intelligence in Robotics and Autonomous System (IROS 2001)*. 2001: 2332-2337.
- [8] Imai, M., Ono, T., and Ishiguro, H. Physical relation and expression: Joint attention for human-robot interaction. *IEEE Transactions on Industrial Electronics*. 2003; 50(4): 636-643. <http://dx.doi.org/10.1109/TIE.2003.814769>
- [9] M. Bennewitz, F. Faber, D. Joho, M. Schreiber, and S. Behnke, Towards a humanoid museum guide robot that interacts with multiple persons, in *Proc. of the IEEE/RSJ Int. Conf. on Humanoid Robots (Humanoids2005)*. 2005: 418-423.
- [10] S. Kopp and I. Wachsmuth. Model-based animation of coverbal gesture. In *Proc. of Computer Animation*, 2002.
- [11] C. M. Huang and B. Mutlu, Learning-Based Modeling of Multimodal Behaviors for Humanlike Robots, *ACM/IEEE Conference on Human-Robot Interaction (HRI2014)*. 2014: 57-64.
- [12] Richmond, V. P. *Nonverbal communication in the classroom*, 2nd ed. Acton, MA: Tapestry. Press. 1997
- [13] Ambady, N., & Rosenthal, R. Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*. 1993; 64(3): 431-441. <http://dx.doi.org/10.1037/0022-3514.64.3.431>
- [14] Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory*. 1968; 2: 89-195. San Diego, CA: Academic Press.
- [15] Glanzer, M. Storage mechanisms in recall. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*. 1972; 5: 129-193. San Diego, CA: Academic Press.
- [16] Murdock, B. B. The serial position effect of free recall. *Journal of Experimental Psychology*. 1962; 64: 482-488. <http://dx.doi.org/10.1037/h0045106>
- [17] H. Kamide, K. Kawabe, S. Shigemi, and T. Arai Development of a Psychological Scale for General Impressions of Humanoid, *Advanced Robotics*. 2013; 27(1): 3-17. <http://dx.doi.org/10.1080/01691864.2013.751159>