Gracefully Mitigating Communication Protocol Reuse Breakdowns

Khaoula Youssef and Michio Okada
Toyohashi University of Technology, Aichi 441-85150, Japan
Email: khaoula.youssef10@gmail.com; okada@tut.jp

Abstract—In our current research, we assessed the feasibility of integrating a dual coded robot's feedback. Dual coded feedback consists of the robot combining it visible behaviors (e.g.: the robot's movement) with auditory icons which we refer to as Inarticulate Utterances (IUs). In our study, we investigate whether using dual coded robot's feedback during the Human-Robot Interaction (HRI) may help maintaining the rules of interaction (protocol) that are previously established in an adaptive way between the human and the robot in past HRI instances and which eventually can be forgotten after a long period of not using the robot. We compare dual coded feedback with other alternatives which are baseline condition in which only the robot's visible behavior is the only afforded feedback for the human and another manipulation condition that consists on changing the IUs version during the second HRI instance when the human is supposed to remember the Previously Established Communication Protocol (PECP). As expected, using dual coded feedback afforded better interaction outcomes in terms of protocol remembrance and subjective evaluation of the HRI in comparison to the baseline and manipulation conditions.

Index Terms—Inarticulate utterance, protocol of communication, dual-coding, recall

I. Introduction

The drive for consistency when interacting with a robot can promote learned helplessness [1]. In fact, altered predicting error signaling while reusing interaction rules that were previously established during past HRI instances between the human and the robot or what we call PECP may contribute to some of the hallmarks of learned helplessness. Specifically, the human may have high confidence about the interaction rules of the PECP that he remembers being established previously between him and the robot while he could have got confused. Because he cannot predict errors and cannot accurately retrieve the rules of the interaction's protocol previously established, possibilities of inconsistency during the HRI increase and the human may feel lost during post HRI instances (when the PECP is supposed to be reused) [2].

Such a scenario can even lead to worse consequences. According to the attribution theory, in order to maintain a positive self-image while performing a task, people tend to attribute failure outcomes to external factors, such as the robot's imperfection [3], [4]. They may feel better about their abilities by shifting the blame for unsuccessful attempts while attempting to recall the PECP from themselves to the robot [5]. One can think that the robot might tell the human through direct speech that he (the human) is the wrong party because he has forgotten the PECP. However, considerable research from HRI and politeness theory shows that people cannot readily accept that technology defeats them by showing rejection for the human's orders [6]. Showing rejection for the human's orders can be considered as a harm especially for self-esteem seekers [7] while a robot may not injure a human being or, through inaction, allow a human being to come to harm according to Isaac Asimov's famous three laws of robotics [8].

Furthermore, research from literature in HRI focuses on PECP recall boosting for only multi-modal expensive robots and no considerable research to the little of our knowledge was conducted in order to investigate the PECP recall challenge when a non-expert user has to remember the PECP previously established between him and a minimally designed robot[1]. In our current research, we are more interested in minimally designed robots that are affordable (cost) for non-expert people. Adding another dimension that it is minimal design paradigm is challenging because rejecting directly the non-expert user's requests may lead in addition to the previously elicited problems that may cause the human's social face harm[2], another problem for the case of minimally designed robots that it is "the adaptation gap". The adaptation gap is related to the differences between the functions of the robot that users expect before starting their interactions which are highly related to the robot's appearance, and the functions they perceive after their interactions. An adaptation gap resulting from the difference between the minimalistic robot appearance and its role as an authority that may dictate to the human how to interact when the human forgets the PECP, may lead to a decrease in the robot's likeability and perceived competence [9].

Manuscript received January 31, 2017; revised March 29, 2017.
Corresponding author email: khaoula.youssef10@gmail.com.
doi:10.12720/jcm.12.3.187-193

[1] A minimal designed robot has small number of sensors and simplified in terms of anthropomorphic features. The design of the robot should be efficient enough to make the minimalistic robot sociable but also affordable (cost) for common non-expert users.

[2] Social face: It is the individual's portrayed identity in a particular situation. It is highly related to self-esteem.

Given the centrality of these issues, we propose to use indirect non deliberative IUs as non-linguistic utterances combined with the minimally designed robot's visible behaviors rather than the direct rejecting speech of the non-expert's requests when the non-expert user cannot remember the PECP. We argue that once IUs are combined with the robot's visible behaviors, the communication protocol can be maintained. Many cartoon films use IUs rather than natural language as a means of communication where viewers will coordinate the cartoon character's behavior with the IU to understand the context, e.g.: Pingu. Thus, we assume that IUs contribute on the context's understanding. By combining, the situation presented in the cartoon with the IU, the human may understand the complete meaning. Linking a visible situation with an auditory icon many times (information encoding phase) may increase the possibility that we remember that particular information (recall phase). This is related to Paivio's dual coding theory that is based primarily on combining visual information with an auditory icon that can be a nonverbal sound to facilitate the information recall in the future [10]. It has been proven that cued recall consisting of presenting the non verbal or pictorial format of the information encoded may lead to better recall results rather than free recall when it is the human's responsibility to retrieve the complete information without us presenting anything to him [11]. Thus, if we assume that the non-expert user combines the robot's different visible behaviors (pictorial format of the information) with the non verbal IUs (non verbal format of the information) during a first HRI's instance (encoding phase), we might have high recall of the PECP if the robot generates the non verbal format (cued recall) of the information before executing the robot's behavior in order to facilitate the recall of the PECP.

II. Related Work

Since the proposed study and its experimental evaluation is motivated by theories from Social Psychology, design concepts and studies from HRI. This section provides an overview on relevant theoretical foundations in human- human interaction and design concepts as well as other HRI related work.

A. *Proposed Solutions to Deal with or Prevent Miscommunication in HRI*

In this subsection, we expose different miscommunication resolution methods presented in the HRI that can be categorized into two types which are the implicit and explicit methods.

1) *Explicit method*

Several studies successfully explored miscommunication arising from users giving instructions in real-time interaction with an artificial agent executing those instructions during the experiment [12], [13], and related error handling is integrated in spoken dialog systems [14]. Error handling through the usage of spoken speech may cause lexical or conceptual difficulties while the robot sometimes cannot cope with the complexity and vagueness of natural language [15]. Argumentation was another alluring solution for the HRI community [16]. Argumentation consists on deriving reasoning semantics by analyzing the supports and defeats [17]. For that purpose, the robot should query the human for more information that may help it get the whole picture during the HRI. That it is why, inquiry and information-seeking dialogues could be employed to resolve interaction errors due to miscommunication [18]. But, again we are putting at risk the HRI, because the non-expert user is not supposed to deal with a robot that may waste their time with argumentation as the user expects total obedience from the robot.

2) *Implicit recall methods*

In order to avoid a situation when a human does not understand the feedback or forgets each instruction's objectives, a few studies use a LED light as an implicit feedback strategy such as Naoki O. *et al.* [19] where the LED light is used to remind silent bystanders to talk during a multiparty conversation. Thus, a miscommunication because of a user's speech prevailing during an interaction can be avoided. Knox W.B. *et al.* [20] used red and green LED lights for TAMER the robot during a demonstration session in order to indirectly remind the user of the incorrect ways of using the pre-programmed robot. Some other studies use a pseudo-implicit method (forewarning) while an instructor explains how to use the robot before the interaction starts [21]. In Ref. [18], a whiteboard near Simon (the robot) is provided as a reminder about the concept representation and the types of sentences that the teacher could articulate.

First, we believe that, a LED light is an implicit communication channel but not sophisticated enough to inform the human about their error without accentuating the frustration (e.g.: the red light indicating error increases the negative feelings). Besides, informing the human before the interaction starts just like in [21], may lead to the human's confusion about the instructions and feeling that the interaction is not quite natural. Moreover, a forewarning is not useful when the amount of instructions organizing the interaction increases. Finally, writing on the board ([18] to remind the person of the concepts taught to the robot, is also inconvenient because it is not a natural communication channel as an important HRI community goal is to make the communication intuitive and natural.

B. *Inarticulate Utterances (IUs)*

We believe that IUs may be considered as a proto-language. IUs are non linguistic and are unable to communicate complex ideas (e.g. "go 5 meters to reach the location") when compared to natural language. This is why we urge caution in thinking about IUs in the sense of a pure language and we also believe that it will not cause social face threatening. In addition to that, it is an implicit

natural communication channel since it is used by children in the child-caregiver interaction context and helps to maintain the communication protocol. Moreover, we argue that people readily attribute meaning to novel IUs as suggested by some HRI studies [22], [23].

III. Architecture of SDT

Knocking is the only input that the robot uses to compute the next move. The robot uses four microphones to detect the knock's sound based on the weighted regression algorithm. It communicates with the host computer through Wi-Fi using a control unit (a macro computer chip (AVR ATMEGA128)) and employs a servomotor that helps to exhibit the different behaviors: right, forward, left and backward, etc. A speaker is used to generate the IUs before executing the future action. Finally, five photo reflectors are utilized to automatically detect the boundaries of the table and to avoid falling (Fig. 1).

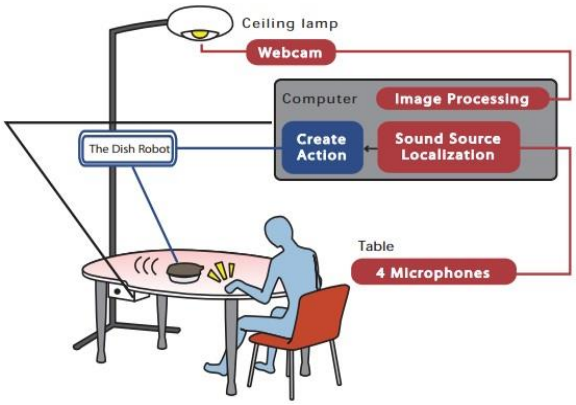


Figure 1. The overall architecture of the SDT: The human's knock is detected by four microphones. The robot (the dish) executes the different behaviors using the servomotor. Using a speaker, the robot can generate the audio output (IU).

IV. ROBOMO Architecture

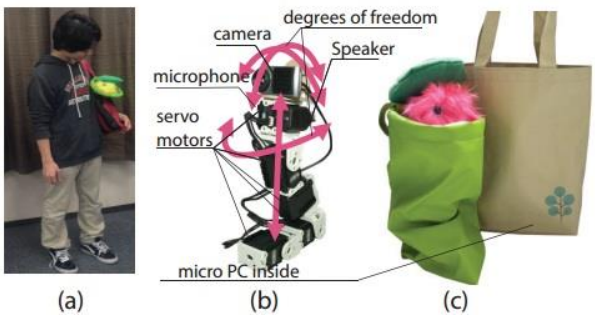


Figure 2. (a) A picture of ROBOMO interacting with a user; (b) A close-up picture showing the inside of ROBOMO the robot; (c) The robot is made of plush material and may emerge from the bag whenever the human interacts with it.

To communicate with ROBOMO, the user has to talk on the microphone so that, the robot can recognize using Julius[3], the meaning of the user's request. ROBOMO tracks the user's face using a Web Camera whilst listening to the human because we believe that face tracking can increase a user's engagement. ROBOMO integrates a micro PC to adapt to the user's request and provides a verbal response through the speaker. ROBOMO uses five servo-motors (AX-12+) to exhibit different gestures such as 'bowing to the left, right, forward or back', 'a confirmation gesture', etc. (Fig. 2)

[3] Julius is a continuous real-time speech recognition decoder for speech-related studies that does not require training.

V. Research Questions

The evaluation of the approach presented in the previous sections included one laboratory study that sought to demonstrate the feasibility of the proposed approach in enabling minimally designed robots to indicate indirectly to non expert users erroneous instructions which are issued by the non-expert users and that do not comply with the communication protocol that was previously established in previous interaction instances.

More specifically, the study sought to answer the question, does combined IU (auditory information) with the minimally designed robot's visible behavior (visual information) enables the minimally designed robot to display appropriate social feedback to the human? Can it minimize changes in the communication protocol that was established in previous interaction instances? If we could validate these two research questions then this may help reducing the information retrieval time in the new interaction instances since the number of erroneous instructions would be reduced.

That it is why, another research question that we may draw is: does our approach facilitate the PECP recall in a shorter time? Also, does it increase the minimally designed robot’s perceived likeability, competence and human's social face support?

VI. Experimental Design and Conditions

To study the research questions above, the study followed a three-by-one, between-participants design. Participants were randomly assigned to one of the three conditions. The three experimental conditions included the following:

- **IUs condition:** The minimally designed robot combined its behaviors with the IUs to facilitate the human's memorization and recall of the PECP (one IU per one robot's behavior; e.g: IU "A" is combined to the left behavior.).
- **Changed IUs during the recall phase condition (manipulated condition):** The minimally designed robot combines IUs with behaviors just like in the IUs condition. However, in the recall phase the IUs used during the encoding phase will be changed. This may help us to validate the importance of IUs usage and maintenance during both phases (encoding and recall) so that the dual-coded feedback could afford the expected communicative outcomes (better recall of the PECP and an amelioration of the non-expert user's perception of the robot's performance).

- **No IUs used (baseline):** The minimally designed robot displayed only its visible behaviors while no IUs will be combined with its behavior.

VII. Hypothesis

The study sought to test the central hypothesis that, by using IUs combined with the minimally designed robot's visible behavior, the minimally designed robot will enable the human to better memorize and recall the PECP in a shorter time which may produce positive outcomes on the objective and subjective levels while the minimally designed robot will be judged more competent, likeable, and supportive for the human's social face. Paragraphs below outline specific instantiations of this central hypothesis.

A. H Ypothesis 1.a

In a given task, IUs generated according to the IUs condition will elicit stronger communicative outcomes such as improved communication protocol recall and a shorter time required for the recall rather than the other two conditions.

B. H Ypothesis 1.b

In a given task, IUs generated according to the IUs condition to hopefully activate the same effect of dual coding theory will improve the participants' perceptions of the minimally designed robot in measures such as likability, competence, and human's social face support more than the other two conditions.

VIII. Setup

We considered in our experiment two minimally designed robots called ROBOMO and SDT which we present in sections 4.3. In all conditions, the minimally designed robots try to collaborate with the human in order to achieve a task. We designed these robots in order that they can be used as service minimally designed robots in the future. Our experiment included two human-robot interaction scenarios:

- Visiting some checkpoints marked on the table with SDT the robot (first scenario: Fig. 3).

- Collaborating with ROBOMO to find a location (second scenario: Fig. 4).

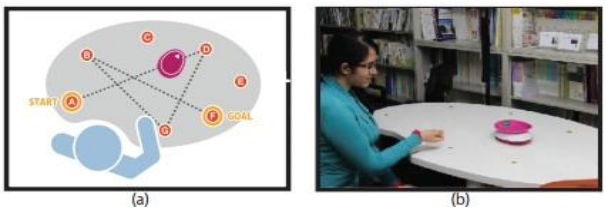


Figure 3. ((a): The experimental setup: The volunteer has to knock on the table so that the robot can translate the composed knocking and chooses an appropriate behavior; (b) a user interacting with SDT the robot.

A. Task Related to SDT

In the first scenario (Fig. 3), the task consists of collaborating with a minimally designed robot that it is SDT to make the dish robot visit different checkpoints marked on the table. In this context, a volunteer has to knock on the table in order to make the robot visit the different checkpoints.

Four microphones positioned under the table help to detect the sound and modulate the knocking in a way that the minimally designed robot can detect whether it is a rapid or a slow knocking and they also help to calculate the number of knocks. As an example, one can imagine a human knocking twice on the table very quickly. The knocking input in this case for the robot is 2 rapid knockings. The robot is capable of showing nine behaviors which are: "going forward", "going left", "going back", "going right", "stop", "vibration", "a whole turn from right to left", "a whole turn from left to right" and "denial" to refuse moving. Our minimally designed robot named SDT will try determine the right behavior to be executed from a composed knocking pattern.

B. Task Related to ROBOMO

In the second scenario (Fig. 4), we setup an indoor space for a navigation task that contained intersections. A volunteer interacts with ROBOMO in order that he/she can find out the directions to visit a particular location. To pick the right behavior, the volunteer is instructed by the robot.

Figure 4. The experimental setup: The volunteer has to hold the robot and ask ROBOMO about the correct direction to finally reach the hidden award.

IX. Experimental Procedure

A. Participants

A total of 32 participants took part in the study. All participants were English speakers from the Toyohashi area with an average age of 22.07 years (SD=3.06), ranging from 18 and 43. Average familiarity with minimally designed robots among the participants was relatively low (M = 2.25, SD =1.7) and average familiarity with the experimental tasks was also low (M = 4.5, SD = 1.8) on a scale of one to seven. Participants were recruited through an email invitation via our JFS database and through personal contacts of the researcher. The majority of participants were students of the Toyohashi University of technology of Japan. Participants each received 1000 yen compensation for their effort.

B. Measurements

The two independent variables in the study were the IUs manipulation (no IUS (condition 3) vs dual coded feedback (condition 1) vs manipulated condition

(condition 2)) that the robot used and participant gender. The dependent vari ables included objective measures of task performance such as communication protocol rules recall (for scenario 1: SDT) and the time needed to recall the rules (for scenario 2: ROBOMO) as well as some subjective measures related to the participants' perceptions of the robot (likeability, competence, social face support).

1) Objective measures

The first measure considered the participant's recall of the communication protocol rules. This measure included a total of nine questions, all related to the rules of the communication protocol established in the first scenario. The questions follow a multi-select format where for each knocking pattern (SDT) a behavior should be combined. The second measure is related to the time it took the participants to finish imitating the robot's gesture correctly after evoking the corresponding command correctly. Specifically, this measurement captured how quickly the participants finished eliciting the correct rules relating to: (1) the directions (right, left, back, forward), (2) traffic lights (go, stop, slow down)(3) the confirmation (yes)/the denial (no).

2) Subjective measures

The post-experiment questionnaire included scales to measure the participants' perceptions of the robot in dimensions of competence of behavior (seven items; Cronbach's alpha= 0.78), social face support (14 items; Cronbach's alpha=0.89), and likeability (5 items; Cronbach's alpha=0.83). The participants rated all questionnaire items using seven-point rating scales.

X. Results

A. Conditions Checks

The analysis of data from condition checks showed that the participants were able to identify the differences across the different videos. The experimental manipulation of the IUs had a significant effect on whether they thought that the behavior which the robot executes and combines with the IUs matched, $F(2,26) = 9.58$, $p < 0.001$, $\eta_p^2 = 0.42$, and whether they found the robot's proposed IUs are always the same when combined with the robot's behaviors or if there is a variation., $F(2,26) = 6.42$, $p=0.006$, $\eta_p^2 = 0.33$.

B. H ypothesis 1.a: Correctly Recalled Rules

As a reminder hypothesis 1. can be elicited as follows: In a given task, IUs generated according to the IUs condition will elicit stronger communicative outcomes such as improved communication protocol recall rather than the other two conditions.

1) Hypothesis 1.a: Number of correctly recalled rules

The data from the information recall measure provided support for this prediction. The number of correct answers out of ten questions in the recall test were on average (mean=4.74 sd,= 1.38), (mean=4.37, sd=1.95), and (mean=7.37, sd=2.66) for the no IUs condition (video3 category), manipulated condition (video 2 category), and condition 1 (video 1 category), respectively. The ANOVA found a significant main effect of the robot's IUs stability (using the same IUs during the encoding phase) on recall accuracy, $F(2,26)=7.18$, $p =0.003$, $\eta_p^2 = 0.35$.

Contrast tests showed that recall performance was significantly higher in the condition 1 (video 1 category) usage than in the manipulated condition (video 2 category), $F(1,26) = 13.71$, $p = 0.001$, $\eta_p^2 = 0.34$, or than in the No IUs used (video 3 category), $F(1,26)=7.87$, $p=0.009$, $\eta_p^2 = 0.23$. Fig. 5 illustrates these results.

2) Hypothesis 1.a: Time Needed to Correctly Recall the Previously Established Rules

Hypothesis 1.a also predicted that there is a reduced time needed to remember the rules, in the condition 1 (video 1 category) in comparison to the other conditions (videos 2 and 3 categories). The analysis of data from this measure partly supported the hypothesis; when the end of the robot's inarticulate utterances (in the case of video 1 or 2 categories) or the instructor's end of spoken question (in the case of video 3 category) was set to zero, the average times in milliseconds that the participants took to remember the corresponding robot's behavior were (mean=457.03, sd= 292.43), (mean=582.14, sd= 612.71), and (mean=-975.22, sd = 405.4) for the No IUs (video 3 category), Highly variated IUs (video2 category), and constant (only one IU per one behavior to facilitate the recall) IUs conditions (video 3 category), respectively. The ANOVA found the main effect of the ex perimental manipulation on the time measure, $F(2,26) = 28.1$, $p < 0.001$, $\eta_p^2 = 0.61$. Contrast tests showed that participants in the IUs (video 1) condition located objects in significantly shorter time than participants in the highly variated IUs (video 2) condition, $F(1,26) = 52.4$, $p < 0.001$, $\eta_p^2 = 0.61$, and No IUs condition (video 3), $F(1,26) = 33.2$, $p < 0.001$, $\eta_p^2 = 0.55$, did (Fig. 5).

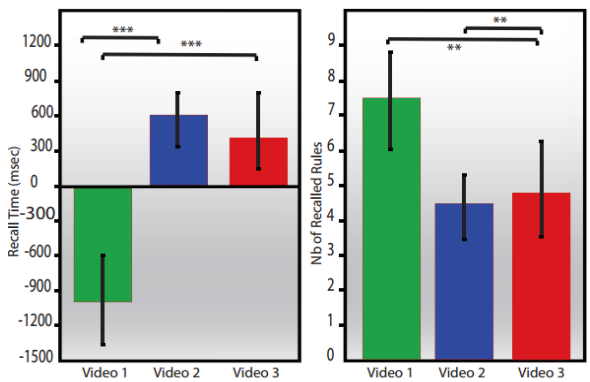


Figure 5. Results on rules correctly recalled and reaction time to retrieve the learned rules watched on the videos. (±), (*), (**), and (***) denote p <0.10, p <0.05, p <0.01, and p <0.001, respectively.

C. Hypothesis 1.b: The Participant's Perception of the Robots

Hypothesis 1.b predicted that the participants would perceive the robot to be more likable, more competent in behavior, and more human social face supportive in the

IUs condition than they would in the other conditions. The data from subjective measures provided partial support for this hypothesis. The analysis showed the main effect of the experimental manipulation on participants' perceptions of the robot's competence, F(2,26) = 4.33, p = 0.024, η_p^2 = 0.250, and human's social face support, F(2,26) = 12.67, p <0.001, η_p^2 =0.49, but not on it likability, F(2,26)=2.25, p=0.125, ηp 2=0.11. In particular, participants in the IUs condition (video 1) rated the robot to be more competent than they did in condition 2 (video 2), F(1,26) = 5.31, p= 0.03, η_p^2 = 0.17, and No IUs condition (video 3), F(1,26) = 7.97, p = 0.009, η_p^2 = 0.24. Similarly, the participants in the IUs condition (video 1) rated the robot to be more human social face supportive than they did in the condition 2 (video 2), F(1,26) = 15.84, p <0.001, η_p^2 = 0.38, and No IUs condition (video 3), F(1,26) = 23.14, p <0.001, η_p^2 = 0.47. Contrast tests with the data from the likeability measure showed that participants rated the robot as marginally more likable in the IUs condition (video 1) than they did in the condition 2 (video 2), F(1,26) = 3.95, p =0.05, η_p^2 = 0.13, and the No IUs condition (video3), F(1,26) = 3.09, p =0.091, η_p^2 = 0.106. Fig. 6 also illustrates these results.

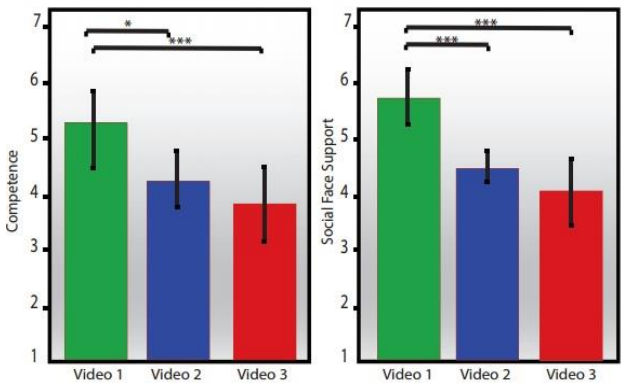


Figure 6. Results on the human's perceptions about the robot in terms of competence, likeability and social face support. (±), (*), (**), and (***) denote p <0.10, p <0.05, p <0.01, and p <0.001, respectively

XI. Discussion

The results provided support for Hypothesis 1.a in measures of information recall. The use of IUs enabled the robot to elicit improved recall of the PECP rules that were presented in the videos. A close look at the participants' behaviors in the no IUs condition (video 3 category) and the condition 2 with the second scenario illustrates why the robot's behaviors in these conditions elicited inferior task outcomes. In fact, the data showed that the participants needed 400-600 milliseconds to recall the rule while barely in condition 1 when the robot starts the IU, the human remembered the activating command and the related behavior before even that the robot finished the IU.

The results also supported Hypothesis 1.b in measures of the robot's perceived competence, social face support, and partially in the likeability measure. The ability to facilitate the communication protocol encoding based on the IUs usage enabled the robot to elicit improved perceptions of the robot in dimensions of competence and the human's social face support, while resulting in marginal improvements in the robot's likeability.

A potential explanation for the lack of significant improvements in the likeability measure is that, while competence and social face support are key qualities for a robot providing a service to a human (like in the case of our robots that it are trained to be used in a restaurant or in the street and as the instructor told to the participants), the participants might not have found likeability to be particularly relevant to the social situations in which they interacted with and evaluated the robot. In fact, as we debriefed the participants to get ideas about their opinions concerning the robot's new design; we remarked that 72% of the participants ascribed positive traits to the robots in condition 1. However, they indicated that robots that need to afford a service for a user should be friendlier enough to be socially accepted by people. Participants afforded many propositions to make the robots friendlier such as adding more degrees of freedom related to the robot's movement, adding an IU related to laughing and mourning, etc. None of the comments indicated that the proposed IUs are not likeable. That it is why, an alternative explanation is that, given the marginal effects in the predicted direction, the study lacked sufficient statistical power to show significant differences due to the small sample size.

XII. Conclusion

This paper presented a novel approach that may help in the future to enable minimally designed robots to indicate to humans without threatening their social faces that the PECP is about to change. We tried to compare three conditions: baseline condition that consists of the robot exhibiting only it visible behavior (no IUs are used), manipulated condition that consists of changing the IUs during the recall phase (video 2 category) and the IUs condition (video 1 category). We measured the differences between the three conditions in terms of rules recall and time needed for rules retrieval. Results indicated that using simple IUs where for each robot's visible behavior, we have one IU that it is combined, ameliorates the human's remembrance of the previously encoded (established) rules in previous interaction instances in a shorter time. We also remarked that changing IUs during the recall phase leads to worse interaction outcomes in terms of recall and the robot's subjective evaluation in comparison to the IUs condition (when the IUs are maintained the same during the recall phase).

Acknowledgment

This research has been supported by Grant-in Aid for scientific research of KIBAN-B (26280134) from the Japan Society for the Promotion of science (JSPS).

REFERENCES

[1] T. Klapwijk, A. L. Goddings, H. S. Burnett, G. Bird, R. M. Viner, and S. J. Blakemore, "Increased functional connectivity with puberty in the mentalising network involved in social emotion processing," presented at the Hormonal Behavior, 2013, pp. 314-322.

[2] E. J. Thomas, S. A. Mark, and I. M. Gerald, "Caught in the crossfire: depression, self-consistency, self-enhancement, and the response of others," *Journal of Social and Clinical Psychology*, vol. 12, no. 2, pp. 113-134, 1993.

[3] T. Thompson, J. A Davidson, and J. G Barber, "Self-worth protection in achievement motivation: Performance effects and attributional behavior," *Journal of Educational Psychology*, vol. 87, no. 4, pp. 598-610, 1995.

[4] W. P. Vispoel and J. R. Austin, "Success and failure in junior high school: A critical incident approach to understanding students' attributional beliefs," *American Educational Research Journal*, pp. 277-312, 1995.

[5] M. Youngme and N. Cliford, "Are computers scapegoats? Attributions of responsibility in human-computer interaction," *Journal of Human-Computer Studies*, vol. 49, no. 1, pp. 79-94, 1998.

[6] T. W. Bickmore and R. W. Picard, "Establishing and maintaining long-term humancomputer relationships," *ACM Transactions Computer-Human Interaction*, vol. 12, no. 2, pp. 293-327, 2005.

[7] L. E. Park and J. Crocker, "Interpersonal consequences of seeking self-esteem," *Personality Social Psychology Bulletin*, vol. 31, no. 11, pp. 1587-1598, 2005.

[8] R. R. Murphy and D. D. Woods, "Beyond Asimov: The three laws of responsible," *IEEE Intelligent Systems*, vol. 24, no. 4, pp. 14-20. 2009.

[9] K. Takanori And Y. Seiji, "Effects of adaptation gap on users' variation of impression about artificial agents," *Transactions of the Japanese Society for Artificial Intelligence*, vol. 24, no. 2, pp. 232-240, 2009.

[10] A. Paivio, "Dual coding theory: Retrospect and current status," *Canadian Journal of Psychology*, vol. 45, no. 3, pp. 255-287, 1991.

[11] E. Marcia, "Cued recall and free recall as a function of the number of items per cue," *Journal of Verbal Learning and Verbal Behavior*, vol. 6, no. 2, pp. 257-263, 1967.

[12] T. Koulouri and S. Lauria, "Exploring mis-communication and collaborative behavior in human-robot interaction," *Proceedings of the SIGDIAL*, pp. 111-119, 2009.

[13] G. Skantze, "Exploring human error recovery strategies: Implications for spoken dialogue systems," *Speech Communication, Special Issue on Error Handling in Spoken Dialogue Systems*, vol. 45, no. 3, pp. 325-341, 2005.

[14] G. G. Skantze, "Error handling in spoken dialogue systems - Managing uncertainty, grounding and miscommunication," Doctoral dissertation, Royal Institute of Technology, 2007.

[15] B. Gonsior, S. Sosnowski, M. Buss, D. Wollherr, and K. Kuhnlenz, "An emotional adaption approach to increase helpfulness towards a robot," in *Proc. International Conference on Intelligent Robots and Systems*, 2012, pp. 2429-2436.

[16] M. Q. Azhar, "Toward an argumentation-based dialogue framework for human-robot collaboration," in *Proc. 14th ACM International Conference on Multimodal Interaction*, 2005, pp. 305-308.

[17] B. Pietro, C. Martin, and G. Massimiliano, "An introduction to argumentation semantics," *The Knowledge Engineering Review*, vol. 26, no. 4, pp. 199-218, 2011.

[18] M. Cakmak and A. L. Thomaz, "Designing robot learners that ask good questions," *Human-Robot Interaction Conference*, 26(4), 17-24.

[19] O. Naoki, Y. Yuta, and M. Okada, "Sociable spotlights: A flock of interactive artifacts," in *Proc. seventh annual ACM/IEEE International Conference on Human-Robot Interaction*, 2011, pp. 321-322.

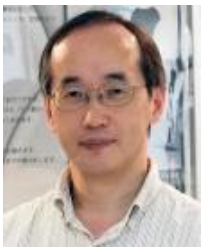
[20] W. B. Knox, P. Stone, and C. Breazeal, "Training a robot via human feedback: A case study," in *Proc. International Conference on Social Robotics*, 2013, pp. 460-470.

[21] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artificial Intelligence*, pp. 716-737, 2008.

[22] R. Read and T. Belpaeme, "Interpreting non-linguistic utterances by robots: Studying the influence of physical appearance," in *Proc. 3rd International Workshop on Affective Interaction in Natural Environments*, 2010, pp. 65-70.

[23] R. Read and T. Belpaeme, "How to use non-linguistic utterances to convey emotion in child-robot interaction," in *Proc. 7th International Conference on Human Robot Interaction*, 2012, pp. 219-220.

Khaoula Youssef was born in Tunisia in 1987. She received the B.S. degree from the Higher Institute of applied sciences and technology of Sousse in 2009 and the M.S. degree from the the Higher Institute of Informatics and Management, in 2012, in smart data systems of information. She is currently pursuing the Ph.D. degree with the Department of Computer Science Engineering, in Toyohashi University of Technology of Japan. Her research interests include human-machine interaction, data mining, big data, and cloud computing.



Michio Okada was born in Japan. He is a professor of Interactions and Communication Design (ICD) Lab at Toyohashi University of Technology (TUT), Japan. His research interests include social robotics, human-robot interaction and cognitive science, especially ecological socio-cultural approach to human cognition.