
Effects of Different Types of Artifacts on Interpretations of Artificial Subtle Expressions (ASEs)

Takanori Komatsu

Shinshu University
3-15-1 Tokida,
Ueda 386-8567, Japan.
tkomat@shinshu-u.ac.jp

Seiji Yamada

National Institute of Informatics
2-1-2 Hitotsubashi,
Tokyo 101-8430, Japan.
seiji@nii.ac.jp

Kazuki Kobayashi

Shinshu University
4-17-1 Wakasato,
Nagano 380-8553, Japan.
kby@shinshu-u.ac.jp

Kotaro Funakoshi

Honda Research Institute Japan.
8-1 Honcho,
Wako 351-0188, Japan.
funakoshi@jp.honda-ri.com

Mikio Nakano

Honda Research Institute Japan.
8-1 Honcho,
Wako 351-0188, Japan.
nakano@jp.honda-ri.com

Abstract

So far, we already confirmed that the artificial subtle expressions (ASEs) from a robot could convey its internal states to participants accurately and intuitively. In this paper, we investigated whether the ASEs from an on-screen artifact could also convey the artifact's internal states to participants in order to confirm whether the ASEs can be interpreted consistently for various types of artifacts. The results clearly showed that the ASEs' interpretations from on-screen artifact were consistent with the ones from robotic agent.

Keywords

Artificial subtle expressions (ASEs), interpretation, robot, on-screen artifact.

ACM Classification Keywords

H5.2 User Interfaces: Evaluation/methodology; J.4 Social and behavioral sciences: Psychology.

General Terms

Experimentation, Human Factors.

Introduction

Up to now, many studies about human communications reported that small changes in the expressions of paralinguistic information (e.g., pitch and power of

Copyright is held by the author/owner(s).
CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

utterances) and nonverbal information (e.g., facial expressions, gaze directions, and gestures) facilitate smooth human communications especially in conveyance of one's internal states to others [1,2], and such simple information is called subtle expressions [3,4]. Some researchers then tried to implement such humans' subtle expressions on the artifacts [5,6]. For example, Sugiyama et al. [6] developed a humanoid robot that can slightly change its behaviors based on its situation recognition. However, it can be easily imagined that these implementation costs were considerably expensive.

On the other hand, we already found that the artifacts' simple expressions like beeping sounds or blinking LED's could play a similar role to such human's subtle expressions [7,8]. Based on the results of these studies, we proposed "Artificial Subtle Expressions (ASEs)" as intuitive notification methodology for artifacts' internal states for users [9]. Specifically, we stipulated that ASEs are artifacts' simple and low-cost expressions that enable humans to estimate the artifacts' internal states accurately and intuitively, and then we could experimentally recognize that such ASEs had succeeded in conveying the robot' internal states (i.e., confidence level of the suggestions) to the participants accurately and intuitively [9, 10]. We are now planning to implement these ASEs in various types of artifacts required in communication with users; not only for robots but also for artifacts appearing on the display.

However, various studies have reported that the different types of artifacts (e.g., robots vs. on-screen agents) evoke users' different attitudes or different impressions toward these artifacts. For example, Wainer et al. [11] reported that a robot was seen as

most helpful, watchful and enjoyable compared to an on-screen agent, while Shinozawa et al. [12] observed that the appropriate types of artifacts depended on its interactive situations. Thus, the issue whether the ASEs can be utilized for various types of artifacts should be investigated because we only confirmed that the ASEs expressed from a MindStorms robot (LEGO Corporation) were effective in our former study [9].

The purpose of this study is then to investigate whether the ASEs expressed from an artifact appearing on the display (we call this artifact "on-screen artifact") could convey its internal states to participants accurately and intuitively, and to compare the results of this experiment with the ones of the former study. If we could observe that the ASEs expressed from an on-screen artifact also succeeded in conveying its internal states to the participants accurately and intuitively like our former study, we could conclude that the ASEs' interpretations from on-screen artifact are consistent with the ones from robotic agent and that the ASEs can be utilized in various types of artifacts.

Experiment

Settings

We used a "driving treasure hunting" video game as an experimental environment to observe the participants' behavior. In this game, a game image scrolls forward on a straight road, like a participant is driving a car with a car navigation system (the left bottom of the Figure 1), with small hills appearing along the way. A coin is inside one of three hills, while the other two hills have nothing. The game ends after the participant encounters 20 sets of hills, and the approximate duration of this video game is about three minutes. The purpose is to get as many coins as possible. In this

experiment, the participants were awarded 1 point for each coin that they found. The participants in this experiment were informed that 1 point was equivalent to 50 Japanese yen (about 50 US cents) and that after the experiment, they could use their points to purchase some stationery supplies (e.g., ballpoint pen or mechanical pencil) of equivalent value. The position of the coin in the three hills was randomly assigned. In each trial, a car navigation system next to the driver's seat on the screen told them in which position it expected the coin to be placed. The navigation told the expected position of the coin using their speech sounds. The participants could freely accept or reject the navigation's suggestions. In each trial, even though the participants selected one hill among three, they did not know whether the selected hill had the coin or not (actually, the selected hill just showed a question mark and closed treasure box, as depicted in the center of Figure 1). The participants were informed of their total game points only after the experiment.

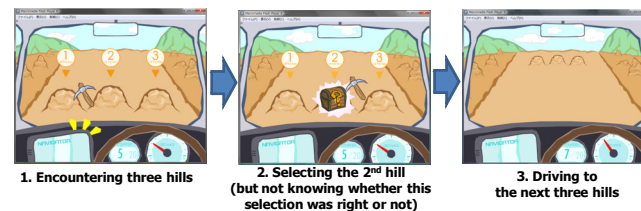


figure 1. Treasure hunting video game

Utilized ASEs

We utilized the audio ASEs in the navigation's speech sounds. In this experiment, the navigation expressed Japanese artificial speech sounds to tell the expected position of the coin; that is, "ichi-ban (no. 1)," "ni-ban

(no. 2)," and "san-ban (no. 3)." These artificial speech sounds were created by the text-to-speech (TTS) function of "Document Talker (Create System Development Company)." Just 0.2 second after these speech sounds, one of the two simple artificial sounds was played as the ASE (Figure 2). These two ASEs were triangle wave sounds 0.5 second in duration, but their pitch contours were different; that is, one was a flat sound (onset F0: 250 Hz and end F0: 250 Hz, called "flat ASE"), and the other was a decreasing one (onset F0: 250 Hz and end F0: 100 Hz, called "decreasing ASE"). These ASE sounds were created by "Cool Edit 2000 (Adobe Corporation)." Actually, these speech sounds and ASEs were the same ones that were utilized in our former study [9]. In that study, we already confirmed that the speech sounds with decreasing ASEs informed users of the robot's lower confidence in the suggestions as the robot's internal state.

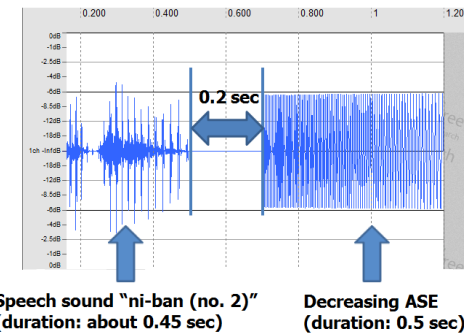


figure 2. Speech sound "ni-ban (No.2)" and ASE

Procedure

Twenty Japanese university students (14 men and 6 women; 21 – 24 years old) participated. The driving treasure hunting video game was projected on a 46-

inch LCD in front of the participants with the distance between them being approximately 100 cm (Figure 3). The navigation's speech sounds were played on the speaker equipped with this LCD, and the sound pressure of these speech sounds at the participants' head level was set at about 50 dB (FAST, A). Before the experiment started, the experimenter told the participant the setting and purpose of the game. However, the experimenter never mentioned or explained the ASEs. Therefore, the participants had no opportunity to acquire prior knowledge about the ASEs. Among the 20 trials, the navigation expressed the flat ASE 10 times and the decreasing ASE 10 times. The order of expression for these two types of ASEs was counterbalanced across participants.



figure 3. Experimental Scene

Here, the experimental stimuli and its procedure were completely same with our former study [9], while the type of artifact to express the speech sounds with ASEs was only different. The purpose of this experiment was to observe the participants' behavior whether they

accepted or rejected the navigation's suggestions in terms of the types of ASEs used. If we could observe the phenomenon that the participants would accept the navigation's suggestion when the flat ASE was added to the speech sounds while they would reject the suggestion when the decreasing ASE was used, we could recognize that the utilized ASEs had succeeded in conveying the navigation's internal states to the participants accurately and intuitively.

Results

ASEs from an on-screen artifact

To investigate the effect of the ASEs from the navigation as on-screen artifact on the participants' behavior, we calculated the rejection rate, indicating how many of the navigation's suggestions the participants rejected for 10 flat ASEs and 10 decreasing ASEs. For all 20 participants, the average rejection rate of the 10 flat ASEs was 1.75 (SD=1.61), while the rejection rate of the 10 decreasing ASEs was 4.35 (SD=2.76, see Figure 4). These rejection rates for the 10 flat ASEs and 10 decreasing ASEs were then analyzed using a one-way analysis of variance (ANOVA) (within-subjects design; independent variable: type of ASE, flat or decreasing, dependent variable: rejection rate). The result of the ANOVA showed a significant difference between the two stimuli ($F(1,19)=8.16$, $p<.05$, (*)); that is, the on-screen artifact's suggestions with the decreasing ASE showed a significantly higher rejection rate compared to the one with the flat ASE. Therefore, we could observe that the ASEs expressed from an on-screen artifact also succeeded in conveying the artifact's internal states to the participants accurately and intuitively.

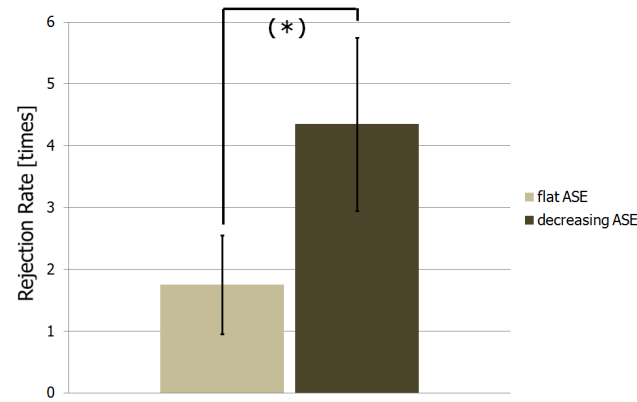


figure 4. Rejection rate for ASEs from on-screen artifact in all 20 participants.

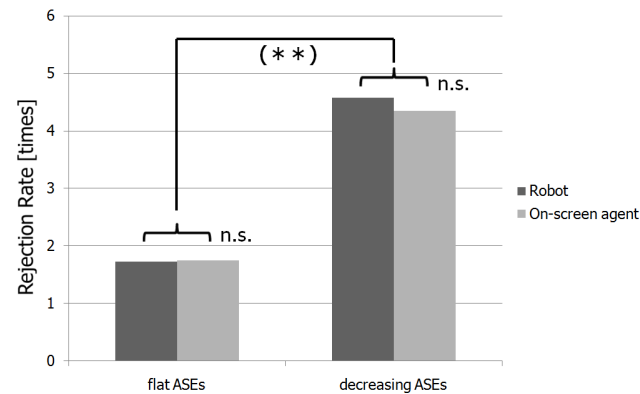


figure 5. Rejection rate for ASEs from an on-screen artifact and a robot.

Comparison of an on-screen artifact with a robot

To investigate how the interpretations of the ASEs from an on-screen artifact were different with the ones from

a robot, we compared the results of this experiment with ones acquired in our former study. Specifically, in the former study, the average rejection rate of the 10 flat ASEs from the robot was 1.73 (SD=1.51), while the rejection rate of the 10 decreasing ASEs was 4.58 (SD=2.43). Because the participants in this experiment did not participate in former study, these rejection rates for the 10 flat ASEs and 10 decreasing ASEs acquired in this experiment and former experiment were then analyzed using a 2 (independent variable in within-subjects factor: types of artifacts, on-screen artifact or robot) x 2 (independent variable in within-subjects factor: type of ASEs, flat or decreasing) mixed ANOVA (dependent variable: rejection rate). The results showed that there were no significant differences in the interaction effects ($F(1,37)=0.04$, n.s.), in the main effects of “types of artifacts” ($F(1,37)=0.08$, n.s.), while there was a significant differences in the main effects of “types of ASEs” ($F(1,37)=20.48$, $p<.01$ (**)) (Figure 5). Thus, we could confirm that the interpretations of ASEs from on-screen artifact were the same with the ones from robot.

Discussions and Conclusions

In this paper, we experimentally investigated whether the ASEs from on-screen artifact could convey the artifact’s internal states to participants in order to confirm whether the ASEs can be interpreted consistently for various types of artifacts. The results of the experiment clearly showed that the ASEs from an on-screen artifact had succeeded in conveying its internal states to the participants accurately and intuitively. And the comparison of the result of this experiment with ones in the former study showed that the interpretations of ASEs from an on-screen artifact were the same with the ones from a robot. These

results succeeded in strongly appealing the robustness and consistency of the ASEs that are simple and low-cost expressions that enable humans to estimate the artifacts' internal states accurately and intuitively. Eventually, we could confirm that the ASEs can be implemented and utilized in various types of artifacts requiring in communicating with users; e.g., spoken dialogue systems such as ATMs or automatic reservation systems. Specifically, we are now focusing on car navigation system for the target application of ASEs; because, current car navigation systems still sometimes give poor driving routes to users. However, if this navigation system's confidence level regarding the route instruction is not very high, the instructions of speech sounds with ASEs could implicitly convey a lower confidence level. If the ASEs are still effective in such situations, they could be utilized in realistic situations in which artifacts have to convey their internal states to users. Therefore, experimental investigations in this paper strongly support this practical application of ASEs.

References

- [1] Kendon, A. Do gestures communicate? A Review. *Research in Language and Social Interaction* 27, 3 (1994), 175-200.
- [2] Cohen, P. R., Morgen, J. and Pollack, M. E. *Intentions in Communication*, The MIT Press, MA, USA, 1990.
- [3] Liu, K. and Picard, W. R. Subtle expressivity in a robotic computer. In *Proc. CHI2003 Workshop on Subtle Expressivity for Characters and Robots* (2003), 1-5.
- [4] Ward, N. On the Expressive Competencies Needed for Responsive Systems, In *Proc. CHI2003 Workshop on Subtle Expressivity for Characters and Robots* (2003), 33-34.
- [5] Kipp, M. and Gebhard, P. IGaze: Studying reactive gaze behavior in semi-immersive human-avatar interactions, In *Proc. IVA2008* (2008), 191-199.
- [6] Sugiyama, O., Kanda, T., Imai, M., Ishiguro, H., Hagita, N. and Anzai, Y. Humanlike conversation with gestures and verbal cues based on a three-layer attention-drawing model. *Connection Science* 18, 4 (2006), 379-402.
- [7] Komatsu, T. and Yamada, S. How do robotic agents' appearances affect people's interpretation of the agents' attitudes? Ext. Abstracts CHI2007, ACM Press (2007), 2519-2524.
- [8] Funakoshi, K., Kobayashi, K., Nakano, M., Yamada, S., Kitamura, Y. and Tsujino H. Smoothing human-robot speech interactions by using blinking-light as subtle expression. In *Proc. ICMI 2008*, ACM Press (2008), 293-296.
- [9] Komatsu, T., Yamada, S., Kobayashi, T., Funakoshi, K. and Nakano, M. Artificial Subtle Expressions: Intuitive Notification Methodology of Artifacts, In *Proc. CHI2010*, ACM Press (2010), 1941-1944.
- [10] Funakoshi, K., Kobayashi, K., Nakano, M., Komatsu, T. and Yamada, S. Non- humanlike Spoken Dialogue: a Design Perspective, In *Proc. SIGDIAL2010* (2010), 176-184.
- [11] Wainer, J., Feil-Seifer, D. J., Shell, D. A., and Mataric, M. J. Embodiment and Human-Robot Interaction: A Task-Based Perspective, In *Proc. IEEE ROMAN2007* (2007), 872-877.
- [12] Shinozawa, K. Naya, F. Yamato, J. and Kogure, K. Differences in effects of robot and screen agent recommendations on human decision-making, *International Journal of Human-Computer Studies*, 62, (2004), 267-279.