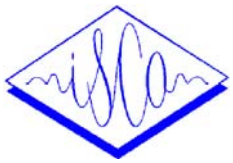

6th European Conference on
Speech Communication and Technology
(EUROSPEECH'99)
Budapest, Hungary, September 5-9, 1999

Evaluation of Affiliation in Interaction with Autonomous Creatures

Noriko Suzuki, Yugo Takeuchi, Kazuo Ishii and Michio Okada

ATR Media Integration & Communications Research Labs.
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto JAPAN 619-0288
e-mail: {noriko, yugo, ishii, okada}@mic.atr.co.jp

ABSTRACT

The paper presents an evaluation method to assess impressions toward a computer-generated creature through human-computer interaction. The aim of our study is to construct a mechanism on a computer to realize empathic interaction between humans and creatures. In this paper, we examine the effect on personality cognition of a virtual creature when the creature replies with different timing to a human utterances. The creature has an abstract appearance and a non-linguistic voice. We control the following three conditions for timing of the creature's response: a) alternation, b) overlap, and c) neutral condition. As a result, we found that the evaluation value of the alternation condition shows more positive behavior than that of the overlap and neutral conditions when the creature has the above features.

10.21437/Eurospeech.1999-68

1 INTRODUCTION

We have tried to realize an empathic interaction between a human and a computer. This is one of the necessary works to construct a computer not as simply a tool but also as an individual entity that can interact with a human on an equal level. The main reason for this work is that our everyday interaction plays a role of not only conveying an explicit message to others but also of establishing an empathic relationship with them. However, it is extremely difficult to implement emotion and sensitivity on a computer by using a symbolic quantization [1, 2].

The aim of our study is to construct a computer with a mechanism that expresses behaviors so that a person could naturally perceive them as if coming from a human companion. Research on robotics has been focused on such a goal. Some robots have been constructed to create an empathic interaction with a human through their physical motion [3, 4, 5]. On the other hand, we have tried to construct a virtual creature that can create an empathic interaction with a human through voice and motion[6, 7].

In this paper, we report an evaluation method to assess impressions toward a creature through interaction with a human. We conducted a psychological experiment to examine the effect on personality cognition of a virtual creature when the creature replies with different timing toward a human utterance. The creature has an abstract appearance as an eyeball and a non-linguistic voice like a bubbling or baby talk. We controlled the following three conditions for timing of the creature's response: the creature replies to a human a) after a human finishes talking (alternation condition), b) as soon as a human speaks (overlap condition), and c) randomly after or as soon as a human speaks (neutral condition). This paper discusses how the timing of a creature's response contributes to affiliation in the interaction between a human and a creature under these conditions.

2 EXPERIMENT

2.1 Environment

For this experiment, subjects interacted with a creature displayed on a 17 inch computer screen via speech (Fig. 1). The creature is shaped like an eyeball and was generated by 3-D computer graphics. It is an effective device for human-computer interaction because it does not cause any special bias in subjects. The creature also has eight kinds of non-linguistic voices such as bubbling or baby talk, which are output randomly. Subjects can start to interact with it in a neutral state because it has the above features.

2.2 Method

Subjects: Total: 24 university students (female: 12 people, male: 12 people, from 18 to 23 years old)

Conditions: The conditions for timing are decided according to the analysis of human-to-human interaction [8] as follows:

Figure 1: Sample snapshot of interaction between human and creature

Alternation: A creature returns its response alternately i) after a subject finishes talking: 80% for all voices (Fig. 2 left), ii) on random timing: 20% for all voices (Fig. 2 right).

Overlap: A creature returns its response i) overlapped subject's talk: 80% for all voices (Fig. 2 middle), ii) on random timing: 20% for all voices.

Neutral: A creature returns its response on random timing: 100% for all voices.

Hypothesis: In general, we can observe much overlapped speech in the interaction among humans when they are familiar. Therefore, we assume that subjects receive a positive impression from a creature that returns its response while overlapping their speech.

Procedure: The sequence of the experiment is as follows:

1. The experimenter ran a video tape for the subject showing a recorded animation where a penguin boy behaves mischievously for about one minute.
2. The experimenter had the subject do a test interaction for about one minute. During this test interaction, the experimenter gave instructions to the subject.
3. After giving the instructions, the experimenter gave a start signal and left the room.
4. After one minute of interaction, the experimenter gave an end signal to stop experiment.
5. After the experiment, the subject answered a post-experimental questionnaire as an evaluation of personality cognition in the creature.

Instruction: This system implements the everyday behavior of children from three to five years old. Please scold the creature shaped like an eyeball for one minute instead of the penguin boy who appears in the last video tape. Please deal leniently with him when he pouts, cries, or resists your speech. You should interact with the three children who have different characters. Please answer questionnaire after each interaction.

Evaluation Items: Table 1 shows the evaluation items in the experiment. The subject evaluates them using seven grades.

2.3 Results and Discussion

Figure. 3 shows average evaluation values for all evaluation items within sociality, familiarity, stability and activity under the alternation, neutral and overlap conditions.

These results of the experiment for every evaluation items are summarized as follows:

1. Sociality: There is no significant difference in the average evaluation value between the alternation and neutral conditions [$t(23) = 0.744$, $p > .05$] as well as between the overlap and neutral conditions [$t(23) = 0.868$, $p > .05$]. Therefore, we assume that the change in timing of the creature's response does not have any effect on the social property of the creature.

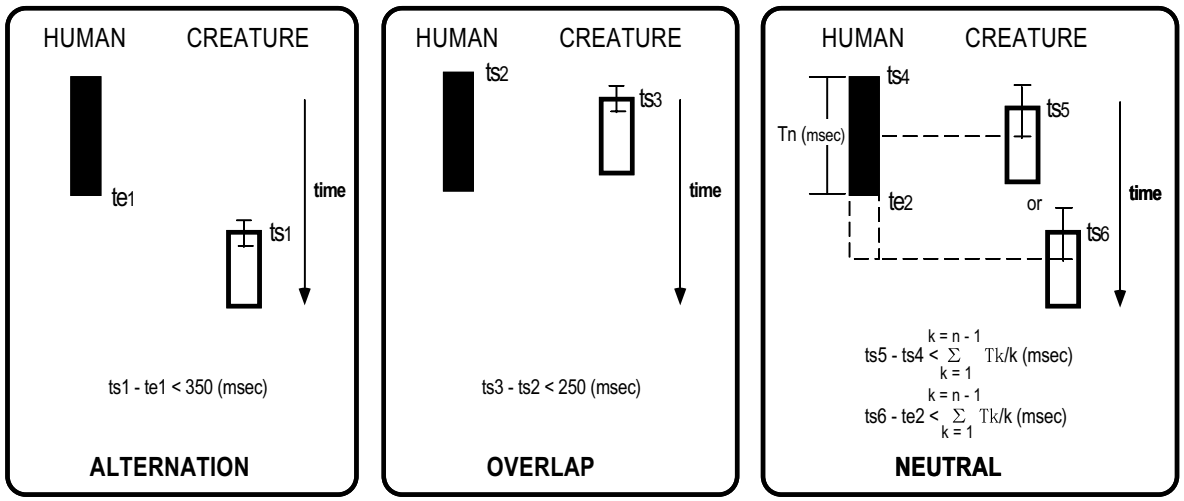

Figure 2: Conditions for timing in experiment

Table 1: Evaluation items concerned with personality cognition

Sociality	Familiarity	Stability	Activity
insincere ↔ serious	unkind ↔ kind	weak ↔ strong	introverted ↔ diplomatic
untidy ↔ tidy	taciturn ↔ talkative	cowardly ↔ brave	passive ↔ aggressive
sentimental ↔ rational	egotistic ↔ thoughtful	unfaithful ↔ faithful	internal ↔ external
insensitive ↔ sensitive	calm ↔ lively	dirty ↔ clean	nerveless ↔ enthusiastic

(*) These are evaluated on seven grades and marked from one (left) to seven (right).

2. Familiarity: There is a significant difference in the average evaluation values between the alternation and the neutral conditions [$t(23) = 2.066$, $p < .05$] as well as between the overlap and neutral conditions [$t(23) = 2.225$, $p < .05$]. Accordingly, we assume that the subjects tend to pay more attention when the creature is under the alternation condition than under the overlap condition. In the alternation condition, the creature moves and talks only after the subjects finish talking. In the overlap condition, it stops its motion in a moment because it starts to move and talk as soon as the subject starts. The main reason is that the subject favorably interprets favorably the behavior of the creature starting in the alternation condition only after s/he finishes talking. On the other hand, s/he unfavorably interprets the behavior on the creature as it moves and talks in the overlap condition while the subject is still talking.

3. Stability: There is no significant difference in the average evaluation value between the alternation and neutral condition [$t(23) = 0.668$, $p > .05$] as wall as between the overlap and neutral condition [$t(23) = 1.182$, $p > .05$]. Therefore, we assume that the change in timing of the creature's response does not have any effect on the stability of the creature.

4. Activity: There is a significant difference in the average evaluation value between the alternation and the neutral conditions [$t(23) = 1.760$, $p < .05$], though there is no significant difference in the average evaluation value between the overlap and the neutral conditions [$t(23) = 0.000$, $p > .05$]. We assume the the subject can feel the behavior of the creature more actively under the alternation condition than under the overlap condition. The main reason is that s/he can better see and hear the creature's motion and voice when it moves and talks after s/he talks than when it does as soon as s/he does.

From these above results, we found that our hypothesis (see 2.2) was not supported in the experimental environment. This suggested that a human's affiliation toward a creature is increased not only by the overlap of voice in interaction but also by interaction with the other modality, i.e. the creature's image.

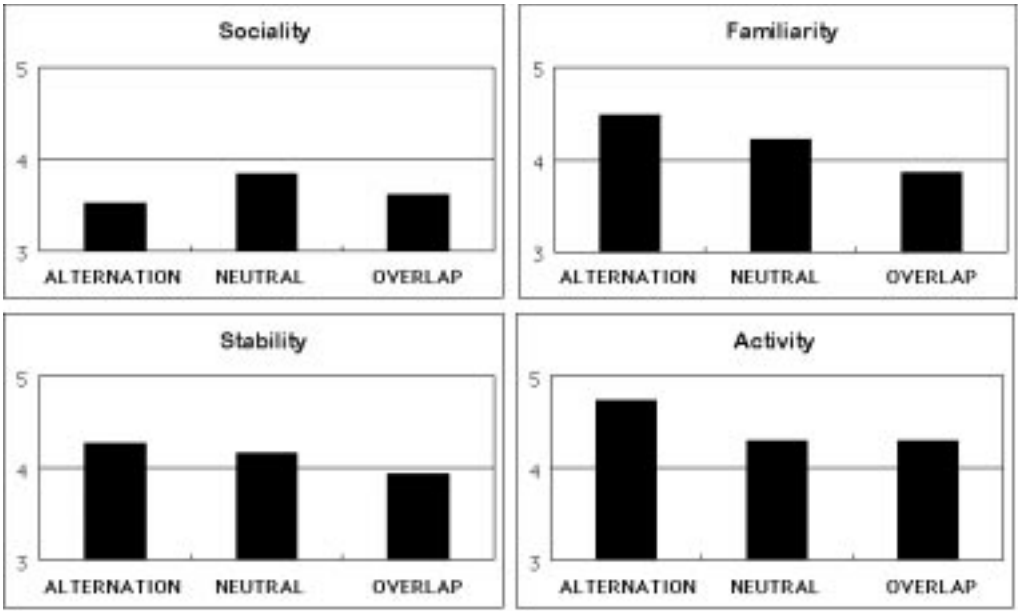

Figure 3: Results of average evaluation value

3 CONCLUSION

We examined whether the change in timing of a creature's response has some effect on a human's impression toward the creature when the creature has the an abstract appearance and a non-linguistic voice.

From the results of an experiment, we found a significant difference for two evaluation items, familiarity and activity. These results suggested the following. The subject can feel a more positive impression about familiarity and activity in the creature's behavior when it starts to move and talk after s/he finishes talking. The main reason for this is that the subject can more favorably interpret the behavior of the creature while it is still since s/he can regard it turning its attention to the subjects' talk. On the other hand, s/he more unfavorably interpreted the creature's behavior under the overlap condition because it starts to move and talk while the subject is talking.

In this experiment, it is difficult to decide which kind of modality, image or voice, contributes to the positive impression of familiarity and activity. In future work, we plan to examine human affiliation toward a creature for a) the effect of the image by using other computer graphics, and b) the effect of voice quality by using more meaningful and linguistic voices.

References

[1] H. L. Dreyfus. *WHAT COMPUTERS CAN'T DO — The Limits of Artificial Intelligence*. Harper & Row Publishers, 1979.

[2] Terry Winograd and Fernando Flores. *Understanding Computers and Cognition — A New Foundation for Design*. Ablex Publishing, 1986.

[3] Toru Nakata, Tomomasa Sato, Taketoshi Mori, and Hiroshi Mizoguchi. Expression of emotion and intention by robot body movement. In *Intelligent autonomous systems 5 (IAS-5)*, pages 352–359, 1998.

[4] Masahiro Fujita and Koji Kageyama. An open architecture for robot entertainment. In *The 1st International Conference on Autonomous Agents*, pages 435–442, 1997.

[5] Hirohide Ushida, Yuji Hirayama, and Hiroshi Nakajima. Emotional model for life-like agent and its evaluation. In *American Association for Artificial Intelligence*, pages 62–69, 1998.

[6] Michio Okada, Noriko Suzuki, and Jacques Terken. Situated dialogue coordination for spoken dialogue systems. In *ICSLP '98*, pages 1227–1230, 1998.

[7] Noriko Suzuki, Kasuo Ishii, and Michio Okada. Organizing self-motivated dialogue with autonomous creatures. In *ICSLP '98*, pages 499–502, 1998.

[8] Hanae Koiso. *A corpus-based study on conversational interaction in Japanese: discourse structures, turn-taking, and backchannels*. Doctor's thesis, 1998.