*Affective and Pleasurable Design,* Vol. 164, 2025, 156–162
https://doi.org/10.54941/ahfe1006071

# Design of a Pupil Response Robot That Listens With Empathy

**Yoshihiro Sejima**

Faculty of Informatics, Kansai University, Osaka, 5691095, Japan

## ABSTRACT

Many proverbs using nonverbal behaviors in human communication are related to human eyes, such as "the eyes speak as much as the mouth" and "the eyes are the mirror of the soul." This is because human eyes are an interface that easily reflects own emotions and are recognized as a social clue to convey emotions and feelings in human communication. It is considered that people infer emotions of others in communication by integrating not only global cues such as "gaze" but also local cues such as "pupils." Therefore, it will be possible to clarify human emotional behaviors and design human-like emotional behaviors by recreating the artificial eyes that integrates gaze and pupils. In particular, one of the human-like behaviors is listening attitude by the listener. Listening attitude is an essential behavior for empathizing with the speaker's emotions or cheering his/her up and an important factor in building a trusting relationship with the speaker. This listening attitude has not been realized even in current AI and is an important challenge in solving human-likeness. In this study, a pupil response robot that listens with empathy was designed. This robot imitates human-like eyeball by projecting CG images of the iris and pupil onto hemispherical displays. In addition, it generates nodding movements, eye contact, and pupil dilation as an active listening attitude. Through experience events, it was confirmed that the developed robot is effective for the listening with empathy.

**Keywords:** Nonverbal communication, Human robot interaction, Social robotics, Human interface, Affective interaction design

## INTRODUCTION

Many proverbs using nonverbal behaviors in human communication are related to human eyes, such as "the eyes speak as much as the mouth" and "the eyes are the mirror of the soul." This is because human eyes are an interface that easily reflects own emotions and are recognized as a social clue to convey emotions and feelings in human communication (Argyle and Dean, 1965). For instance, humans have the characteristic of being easily aware of eye movement called "gaze." Using this characteristic, people can perceive what others are paying attention to (Kendon, 1967). Furthermore, by using the "pupils" controlled by the autonomic nervous system, people can perceive the level of arousal of others such as their interests (Hess, 1965). In other words, it is considered that people infer the emotions of others by integrating not only a global clue of "gaze" but also a local clue of "pupils" (Sejima, 2025).

 

On the other hand, it has been reported that facial expressions and body movements are important role in emotional expressions (Ekman, 1971; McNeill, 1992). However, they are unreliable and have been studied as targets for deception, because these are capable of voluntary movements based on the intention. Therefore, it is expected to clarify the perceptual principle of trust by removing the components that generate voluntary movements and realizing communication with only minimum elements.

In this study, focusing on a counselling where a relationship of trust is essential, a pupil response robot that listens to the speaker was designed. Based on a minimal design, the robot is composed of the iris and pupil in human-like eyeball. Furthermore, it generates nodding movements, eye contact, and pupil dilations as an active listening attitude. Through experience events, it was confirmed that the developed robot is effective for the listening with empathy.

## RELATED WORKS

Generally, a counselling begins with speaking to another person. The listener observes nonverbal behaviors from the speaker and synchronizes with their body movements and emotions. This synchronization creates a sense of unity in communication and emotional empathy (Watanabe, 2020).

According to Rogers, C. R., the listener's synchronous, empathetic and accepting attitude is important in active listening (Rogers, 1946). In this synchronization of nonverbal behaviors, processes of "mirroring" and "pacing" have been confirmed. Mirroring is a process of imitating the visual actions and movements such as facial expressions, gestures, and body movements (O'Connor, and Seymour, 1990). Pacing is a process of sharing time with others such as speaking speed, intervals, and timing (Rogers, 1946). Both mirroring and pacing are important in sharing time with each other and essential in building rapport. Therefore, previous research has investigated the effect of mirroring on forming impressions of talkers (Sharpley et al., 2001), the development of a system that mirrors body movements (Nagai, 2016), the pacing effect in robot dialogue (Nishimura et al., 2019), and the development of a communication system that responds to speaking speed (Yokota et al., 2023). However, no research has been conducted focusing on the eyes which play an important role in building trust. If a sense of acceptance and empathy can be formed through robot's eyes, it could provide an important clue to understanding trust.

## Pupiloid: PUPIL RESPONSE ROBOT

Figure 1 shows a pupil response robot called "Pupiloid" as a minimal design of human eyes (Sejima et al., 2019). "Pupiloid" is a newly created word formed by combining "Pupil" and the suffix "-oid," which means "resembling." Pupiloid mimics the "appearance" of the eyeball using a hemispherical display with a diameter of 250mm, which is approximately 10 times the size of an adult human eyeball, which is about 24mm (Bekerman et al., 2014). The iris and pupil are displayed in proportions similar to those of a human eye within 10 times. To control robot's gaze and pupil response in real-time, CG models were introduced into the virtual space (in Fig. 2). The CG models consist of the iris part and the pupil part. By making the iris part blue, it enhances the contrast of pupil. Additionally, it seems the sclera (white part of the eye) by setting the entire virtual space to white.

**Figure 1:** Pupiloid: pupil response robot (reference: Sejima, 2025).

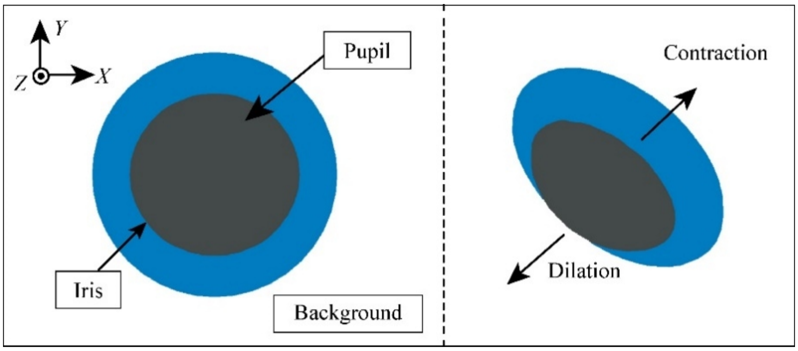


**Figure 2:** CG models of iris part and pupil part (reference: Sejima, 2025).

## DESIGN OF LISTENING ATTITUDE

The concept of the pupil response robot that listens with empathy is shown in Fig. 3. When a person desires to talk to someone about enjoyable events or activities that he/she is interested in or motivated to engage in, the designed robot generates synchronized nodding movements and backchanneling as well as mirrored pupil responses as an active listening attitude. In particular, based on the findings of previous research, it is expected that emotional empathy will be created through mutual synchronization.

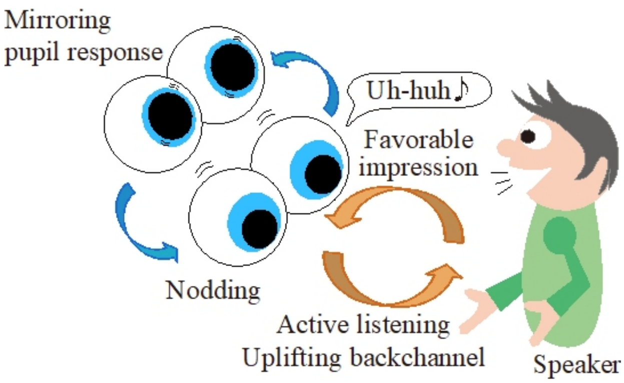


**Figure 3:** Concept of pupil response robot that listens with empathy.

Our previous research indicated that pupils dilate in the synchronization with the speaker's speech as a characteristic of communication (Sejima et al., 2017). This phenomenon makes it possible to generate a pupil dilation response based on speech input. In this study, the speaker's pupil response generated from the speech is mirrored and expressed as a virtual pupil response of the listener. As a model for generating the speaker's pupil response, an estimation model is introduced. This model can estimate the talker's activation level by relating real-space communication to a virtual temperature space (Sejima et al., 2016). Here, the estimated talker's activation level is associated with the pupil size and mirrored as the listener's pupil response (Hashimoto et al., 2024). Specifically, the standard value of pupillary area ($P_0$) is set to the minimum value as default value, and current pupillary area calculated using estimated activation level by equation (1).

$$P = P_0 + \alpha \cdot \Delta T \tag{1}$$

$P$: Current pupillary area
$P_0$: Standard pupillary area
$\alpha$: Coefficient
$\Delta T$: Estimated talker's activation level.

Above-mentioned mirroring method is visual behavior, in which the estimated activation level is reflected directly as the size of the pupils. In contrast, pacing is a process of sharing time, both nodding responses and backchanneling that are synchronized with the speech were incorporated into. Specifically, a nodding response model that predicts the timing of nodding based on the speaker's speech input was introduced (Watanabe et al., 2004). Based on this nodding response model, the robot generated a nodding movement and two types backchanneling. When the estimated virtual temperature is less than or equal to a preset threshold, a general backchanneling is generated. If the threshold was exceeded, an excited backchanneling was generated. Here, general backchanneling was set to "yeah," and an excited backchanneling was set to "wow."

## EFFECTS OF LISTENING ATTITUDE IN COMMUNICATION

In order to confirm the effect of designed listening attitude, an experiential experiment was conducted for children. The participants were 132 children (64 boys, 60 girls, 8 others) aged 11 to 12 years old who belong to the same elementary school. This experiment was conducted after obtaining approval from the Okayama Prefectural University Ethics Committee and the consent of the children.

The experimental procedure was as follows: First, the robot was explained to the children. Then, they engaged in conversation with the robot. Although conversational topics were not specified, almost topics were favorite foods, anime, and music groups. After the experience, the children were asked to evaluate the robot on a 7-point scale (neutral 0) for (a) ease of conversation, (b) familiarity, (c) active listening, (d) (robot) curiosity, and (e) (robot) gaze.

The questionnaire results are shown in Fig.4. The survey targeted 80 people (44 boys, 35 girls, and 1 other) who responded that they had actually engaged in conversation with the robot. As Fig.4 shows, 68% of children gave a positive evaluation of "(a) ease of conversation." 78% of children gave a high evaluation of "(b) familiarity," and 81% of children gave a high evaluation of "(c) active listening." It confirmed that interaction with the robot has a high active listening effect that increases closeness for children. In addition, 81% of children gave a positive evaluation of "(d) curiosity," and 86% of children gave a positive evaluation of "(e) gaze," confirming that a good impression was formed of the robot. It is considered that the robot stimulated children's curiosity by realizing eye-contact and dilating pupils in response to their speech input. These results demonstrated the effectiveness of the robot's active listening attitude.

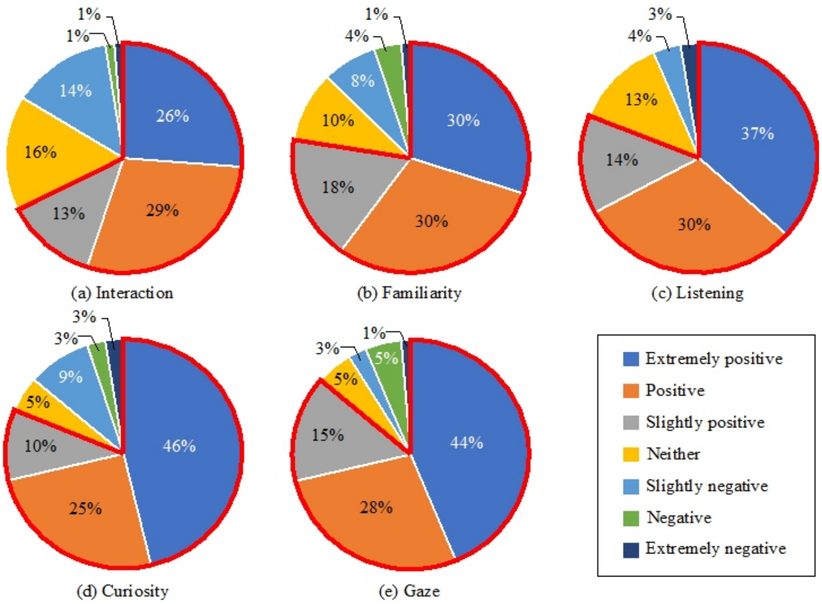


**Figure 4:** Result of the designed robot for children.

## CONCLUSION

In this study, a pupil response robot that expresses the listening attitude using only the eyeballs was designed. Based on the speaker's speech input, the robot generates a mirroring motion by dilating the pupils as well as a pacing behavior by generating nodding movements and backchanneling. Furthermore, sensory evaluations by children demonstrated that the designed listening attitude gave children a sense of familiarity and active listening.

In the future, an algorithm that can change the listening attitude by adjusting for individual characteristics will be designed and developed.

## ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 22H04871 and the Kansai University Organization for the Promotion of Advanced Science and Technology.

## REFERENCES

Argyle, M., and Dean, J. (1965). "Eye contact, distance and affiliation," Sociometry, Volume 28, pp. 289–304.

Bekerman, I., Gottlieb, P. and Vaiman, M., (2014). "Variations in eyeball diameters of the healthy adults," Journal of Ophthalmology, Volume 2014, Article ID 503645, pp. 1–5.

Ekman, P. (1971). "Universals and cultural differences in facial expressions of emotion," in Nebraska Symposium on Motivation, ed. J. K. Cole (Lincoln, NE: University of Nebraska Press), pp. 207–283.

Hashimoto, S., Sejima, Y. and Watanabe, T. (2024). "A speech-driven embodied listening system with mirroring of speaker's pupil response," Paper of Human Interface Society, Volume 26, No. 3, pp. 305–314.

Hess, E. H. (1965). "Attitude and pupil size," Scientific American, Volume 212, No. 4, pp. 46–55.

Kendon, A. (1967). "Some functions of gaze-direction in social interaction," Acta Psychologica, Volume 26, No. 22, p. 63.

McNeill, D. (1992). "Hand and Mind: What Gestures Reveal about Thought," Chicago: The University of Chicago Press.

Nagai, Y. (2016). "Mechanism for Cognitive Development," Cognitive Neuroscience Robotics: A: Synthetic Approaches to Human Understanding, Springer, pp. 51–72.

Nishimura, S., Nakamura, T., Kanbara, M. Sato. W., and Hagita, N. (2019). "Evaluation of Pacing for Dialog Robots to Build Trust Relationships with Human Users," Proceedings of 7th International Conference on Human-Agent Interaction, pp. 300–302.

O'Connor, J. and Seymour, J. (1990). "Introduction neuro-linguistic programming," London, UK: Harper Collins.

Rogers, C. R., and Wallen, J. L. (1946). "Counseling with returned servicemen", McGraw-Hill Book Company.

Sejima, Y., Watanabe, T., and Jindai, M. (2016). "Estimation Model of Interaction-activated Communication based on the Heat Conduction Equation," Journal of Advanced Mechanical Design, Systems, and Manufacturing, Volume 10, No. 9, p. JAMDSM0103.

Sejima, Y., Egawa, S., Maeda, R., Sato, Y. and Watanabe, T. (2017). "A speech-driven pupil response robot synchronized with burst-pause of utterance," Proceedings of 26th IEEE International Symposium on Robot and Human Interactive Communication, pp. 437–442.

Sejima, Y., Egawa, S., Sato, Y., and Watanabe, T. (2019). "A pupil response system using hemispherical displays for enhancing affective conveyance," Journal of Advanced Mechanical Design, Systems, and Manufacturing, Volume 13, No. 2, p. JAMDSM0032.

Sejima, Y. (2025). "Eye Appeal Communication," Springer, Emotional Engineering, Vol. 10, Chapter 4, pp. 39–53.

Sharpley, C. F., Halat, J., Rabinowicz, T., Weiland, B. and Stafford, J. (2001). "Standard posture, postural mirroring and client-perceived rapport," Counselling Psychology Quarterly, Volume 14, pp. 267–280.

Watanabe, T., Okubo, M., Nakashige, M., and Danbara, R. (2004). "InterActor: Speech-Driven Embodied Interactive Actor", International Journal of Human-Computer Interaction, Volume 17, No. 1, pp. 43–60.

Watanabe, T. (2020). "Human-entrained embodied interaction and communication technology for human-connected IoT design", Journal of Advanced Mechanical Design, Systems, and Manufacturing, Volume 14, No. 2, pp. jamdsm0025.

Yokota, M. Hisatomi, A. Ishii, Y. and Watanabe, T. (2023). "A voice-driven embodied interaction system with response actions according to speech speed by the mora," Transactions of the JSME (in Japanese), Volume 89, No. 919, pp. 22–00228.