A New Good Listener, the Digital Human: A comparative research analysis of conversational virtual agents and robots*

Boeun Kwak, Jeongyun Heo** and Eunsoon You**

Kookmin University, Seoul, Republic of Korea
kwakboeun@kookmin.ac.kr,
yuniheo@kookmin.ac.kr, Jiwony71@gmail.com

Abstract. This paper aims to discover the potential of the digital human to develop as a listener and the ability to generate appropriate non-verbal feedback. We look at what aspects of the current digital human are easier to interact with compared with older robots or traditional virtual agents. We examine comparative studies of conversational virtual agents and robots in various contexts and review previous studies investigating non-verbal expressions and characteristics. Based on the research results, four major listener response functions of digital humans are proposed.

Keywords: Digital Human, Conversation, Robot, virtual agent, Listener Feedback, Non-verbal.

1 Introduction

Since the advent of computers, the scope of the conversation partner in the human–artificial agent dialogue system has been developed in various ways. The external appearance of a digital human has reached a level high enough to be recognized as a real human, and we often encounter them on the Internet, kiosks, and TV screens. In addition to message information, human dialogue interactions include tone, pitch, and nonverbal dialogue cues that constitute the context of speech intent and contain emotional expressions.[1] Therefore, human–digital human interaction should be as natural and intuitive as actual human interaction to avoid miscommunication. If a virtual agent's appearance is unnatural, it may offend the user's feelings (e.g., the Uncanny Valley Effect). In this sense, a digital human should resemble a real human being and produce natural sounding/nonverbal responses to the human user. Nonverbal communication

 In: N. D. Vo, O.-J. Lee, K.-H. N. Bui, H. G. Lim, H.-J. Jeon, P.-M. Nguyen, B. Q. Tuyen, J.-T. Kim, J. J. Jung, T. A. Vo (eds.): Proceedings of the 2nd International Conference on Human-centered Artificial Intelligence (Computing4Human 2021), Da Nang, Viet Nam, 28-October-2021, published at http://ceur-ws.org

** Corresponding Authors.

makes up a large proportion of human-to-human communication, comprising about two-thirds of human-to-human contact Nonverbal expressions are made up of the gaze, facial expression, and gestures.[2] In addition to verbal expressions, nonverbal expressions provide important communication functions such as providing information ahead of spoken language in face-to-face interactions, controlling interactions, and expressing intimacy. Therefore, the perception and social effects of digital human behavior should be treated as important.[3, 4]

Nonverbal expressions can sometimes be misleading in meaning transfer because they are represented by symbolic symbols reflecting the culture of each country. People also express their feelings unconsciously and instinctively. Because nonverbal expressions appear as symbolic symbols reflecting the culture of each country, it can sometimes cause misunderstandings in conveying meaning. Sometimes, the latent content of communication can play a more decisive role through these unrecognized nonverbal expressions. Information that is not conveyed through language gives the impression of more than the cognitive activity required for language generation. It can help improve the reliability of digital humans by giving an impression of the mind.[5] In addition, related studies have found that nonverbal communication can improve the likeability of, interest in, and satisfaction with virtual agents.[6, 7] Natural nonverbal communication serves as a key function of relationship formation and is a means by which communication can embody information beyond messages. In particular, since digital humans have a higher degree of freedom of expression of emotions than other types of conversational agents and have the characteristics of manipulation in digital space, it is expected that the nonverbal expressions of digital humans will have a significant impact on design.

The purpose of this paper is to find out why digital humans as virtual agents have greater potential as future conversation partners than robots through a case study comparing virtual agents and robots. This paper consists of the following: Section 2. A comparative study case analysis of virtual agents and robots. Section 3. Deriving the strengths as a good listener of the digital human based on the results of research case analysis. Section 4. Proposed the response four functions of a digital human as a listener (good listener). Section 5. Discussion, conclusions, and future work.

2 Related Work

2.1 Concept and Application of Digital Human and Robot

A digital human (or, virtual human) is an artificial agent with both a human-like body (expression or natural body movement) and intelligent, cognitively-driven behavior.[8] A set of joints is added to the 3D face for expression and movement. The 3D face has features such as eyes, teeth, tongue, and skin. Current research on virtual agents is largely conducted in four areas: environmental design, training, culture and education, and medical care.[9] Recently, digital humans have helped humans by taking various roles such as advertising models (e.g., Lil Miquela, Oh Rozy, Imma), virtual idols, teachers, counselors, coaches, and bankers. Digital human production companies aim

to replace most corporate chatbot services with a digital human. Digital human production companies aim to replace most corporate chatbot services with digital humans. Digital humans cannot perform tasks beyond the environment outside the interface. The representative virtual agent Greta can display hand gestures, but her lower body is motionless, i.e., the activity space is limited to the virtual world. If digital humans can be used to perform tasks or collaborate in real environments for humans, it is expected that the scope of their contribution will be expanded further than now. Robots can have a variety of capabilities, mimicking human emotional states, intention and behavior recognition, interpretation of contextual information, communication, and contextual behavior.[10] Current applications of robots include a variety of areas such as counseling,[11] education and training,[12] security and rescue operations,[13] social services and business,[14] entertainment,[15] and industrial assistance,[16] Robots are becoming more and more advanced in human interaction with humans and compared to the virtual agent, the biggest advantage is that they can exhibit a physical presence. They are now designed to be supported in a personal environment, such as at home. However, the application of robots is still limited in cooperating with humans or in carrying out social tasks for human welfare. This is because robots lack both physical dexterity and expressive ability to imitate simple expressions. In general, robots must meet space and cost requirements and have less hand degree of freedom,[17] The expressions displayed through most postures and facial expressions are also limited, and only a few systems can respond to visual response demands. Physical limitations include the robot's angle, joint speed and torque limitations, awkward arm composition or trajectory, and excessively fast movement. Performing tasks using these action systems can make humans uncertain and anxious. Even state-of-the-art humanoid robots are still unnatural, unhuman, and expensive.[18]

2.2 The concept and function of the listener's nonverbal response

Listener feedback may be defined as a response by the listener to the content of the speaker's utterance. The listener can switch to the speaker's position at any time and can express what they think or feel. During the speaker's turn, they can provide feedback without interfering with the utterance.[19] The expression of the listener's reaction is collectively referred to by various terms such as listener feedback, listener response, backchannel, and nonverbal communication strategy (NVCS). It is used as feedback on receiving the communicative behavior of the interlocutor, and through language and gestures, the listener can indicate the level of participation in the conversation. For example, the speaker may stop the conversation or restructure the sentence if the listener is not interested.[20] Among many feedback types, listener feedback is an important feature in face-to-face interactions because it represents the willingness to continue to hear or invites speakers to continue with the conversation.[21] It can also be used to express evaluations such as surprise, interest, and sympathy. If there is no feedback, the speaker may feel anxious about whether the communication is going well and the listener may feel as if they are talking to a hard "machine," so it should be handled with interest in the communication process. The listener can switch to the speaker's position at any time, express what he or she thinks or feels, and can provide

feedback without interfering with the utterance during the presenter's turn.[19] Fig. 1 reconstructed a Shannon-Weaver-based model to supplement our current understanding of some basic concepts. The listener creates the meaning as code and then transmits it through the message. The speaker receives the code and understands the meaning. The speaker sends a code back to the listener's area in response.[22] In this way, the listener and the speaker are mutually cyclical, suggesting that if you become a good listener, you can become a good speaker at the same time. This paper focuses on the nonverbal listener feedback as the ultimate goal of digital human development as a dialogue listener and the ability to generate appropriate nonverbal expressions.

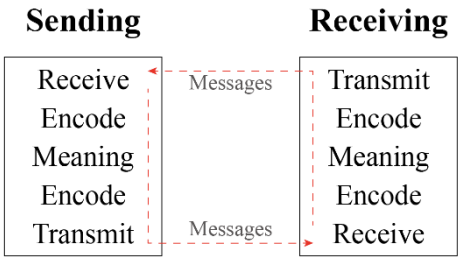

Figure 1 Communication models (Sabah Al-Fedaghi, 2012)

3 Method

3.1 Cases of Comparative Studies of the Virtual Agents and Robots

In human–robot interaction (HRI) studies, there have often been comparative studies of physical and virtual implementations; therefore, we assume that the more human-related social characteristics digital humans present, the more likely they will lead to natural communication, so we would like to look at comparative studies of existing virtual agents and robots. In the following, we list and explain the results of various previous studies focusing on communication between robots and virtual agents.

[23] that when giving recommendations to users in a color selection experiment, robots were less convincing than virtual characters on the screen. Participants had to choose one of four colored square names displayed on a computer monitor. Before making a decision, a robot or virtual character recommended one option to the user, noting that it was the option chosen by other users. As a result, participants followed virtual character recommendations more than robots. In the post-questionnaire response, the subjective "familiarity" factor was found to be different from the subject's behavior. The familiarity factor of the robot group was much stronger than that of the agent group, and the subject accepted more recommendations from the agent.

[24] compared people's responses to robots, projected robots, and agents in health interviews to help them understand differences in people's social interactions with agents and robots. The researchers hypothesized that robots would have more social

impact than agents simply because of their physical proximity. The results showed that the robot had more social impact, but the participants who interacted with the agent remembered more key information in the recall test than the participants who interacted with the robot. studied whether humanoid robots in real life could elicit stronger anthropomorphic interactions than software agents and whether physical presence modulates this effect. The researchers predicted that subjects would anthropomorphize with more anthropomorphic humanoid robots than less anthropomorphic agents. As a result, the participant interacted more with the robot as a person than with the agent, and the more anthropomorphic, the more subjects treated artificial agents as a person.

[25] demonstrated the importance of non-functional aspects that can enhance the level of enjoyment and social presence of older people. It was hypothesized that the more natural and human the conversation with the conversational agent, the higher the perceived pleasure and acceptance. The virtual agent used in this study is "Steffie," and Steffie is a virtual 3D agent that can use various facial expressions, hand/arm gestures, lip-syncing, and voice repetition in the form of a woman. The robot used Philips Electronics' iCat. iCat can make a variety of facial expressions using lips, eyes, eyelids, and eyebrows, has a female voice and is in the shape of a cat. Statistics show a stronger relationship between intention and use of virtual agents than robots. M Heerink revealed that the two agents could not explain why the virtual agent had a stronger influence on intention and use than the robot because of the fundamental difference in the appearance and action system of the two agents.

[26] studied how the physical presence of a robot affects human judgments about a robot as a social partner. Subjects participated in a simple book-moving task with either a physically present robot or a humanoid robot displayed via live video. The Nico robot, which was used in the experiment, was a humanoid robot in the upper body, wearing children's sportswear and a baseball cap. Nico's head has a total of seven degrees of freedom and six degrees of freedom (two on the shoulder, elbow and wrist) on each arm. In the experiment, subjects easily approached Nico in video and augmented conditions, while avoiding face-to-face encounters with the physically present Nico. The researcher identified two causes for these results. First, physical robots can be perceived as more expensive than monitors used in video display conditions, so robots may be reluctant to come closer. Second, the granting of personal space between the robot and the subject can be interpreted as a sign of respect. However, space was also created between Nico in video and augmented conditions, indicating that the first case would be more appropriate.

[13] invited Brazilian subjects to interact with two types of receptionists with different appearances (agent vs mechanical robot) and voices (human vs mechanical) to investigate factors related to designing a receptionist robot for deployment in Brazil. In the interaction experiment, participants interacted with two receptionists with different characteristics (a conversational virtual agent and a humanoid robot) and voice (human-like vs robot). Two receptionists directed the participants to a specific room where the assessment was conducted via questionnaire. The researchers found that when comparing Ana and Kobiana through all categories of questions, they preferred Ana in both groups of participants and that the main reason was its human appearance.

3.2 Cases of Comparative Study of Virtual Agents and Robots

Through the review of related prior studies, we derived the following.
1. Although it was shown that people had stronger behavioral and attitudinal responses to physically existing agents, when both physically implemented agents and virtually implemented agents were presented, each had different results depending on the purpose of the study (Table 1). Depending on the appearance of the virtual agent and its degree of freedom to express, it is assumed that each resulted from a different result.
2. The nonverbal expression of the virtual agent usually has a positive effect on users in the experiment, but unnatural expression or repetition may give a feeling of dis- traction or discomfort. Therefore, it is necessary to provide natural and appropriate feedback.
3. Users expect a natural conversation response from these anthropomorphic agents. Therefore, the more similar to a person the agent is, the more likely the user will be to treat the agent as a person.
4. Finally, since the difference in the social reality given by the implementation environment is greater than the difference in appearance and function, it is necessary to consider how this sense of presence can be realized in digital humans.

Table 1. Table of comparative study cases.

Author	Robot	Virtual Agent	Research Result
[23]	Rabbit Robot	3D Modeling Rabbit Robot	Robots prevailed in familiarity, but agents were higher in acceptance.
[24]	Nursebot Pearl	3D Modeling Nursebot Pearl	Robots had more social impact, but agents remembered more key information.
[25]	Nursebot Pearl	3D Modeling Nursebot Pearl	Depending on the degree of personification, treat as a real person.
[26]	Philips iCat	Steffie	Interactions with agents show a stronger relationship between intent and use than with robots.
[27]	Humanoid Nico	3D Modeling Nico	Overall, participants preferred the robot but easily approached the virtual Nico while avoiding face-to-face interaction with the robot Nico.
[13]	Humanoid Kobiana	Ana	Preferring virtual agents that resemble humans to robots that do not resemble humans.

Based on the above four points, we felt that for a digital human to become a good conversational partner, a new design unique to a digital human is needed that is different from the existing robot design. The current nonverbal representations of virtual

agents are not as natural as the real appearance of the digital human. The above study results are experiments that exclude realistic human-like virtual agents (digital humans), and since they did not focus on subjective evaluation criteria or use representative evaluation scales, there is a possibility that participants' responses may be different. What virtual agents and robots have in common is that they have a body. Whether virtual or physical, due to differences in the physical specifications of the agent, the nonverbal expression and implementation method of the two are different, and various nonverbal expressions can be generated due to this transformation.

3.3 Cases of Comparative Study of Virtual Agents and Robots

Efforts to communicate with humans naturally are continuing as in previous studies and several improvement methods have also been proposed.[28, 29] Allwood proposed four feedback functions: contact, perception, understanding, and attitudinal reactions. In this paper, we reconstruct the overlapping functions of the four feedback functional elements of the preceding studies as "attention," "understanding," and "opinion."[30, 31] Here, we would like to examine the case of a listener's reaction studies by additionally using the element of "timing" (Fig. 2).

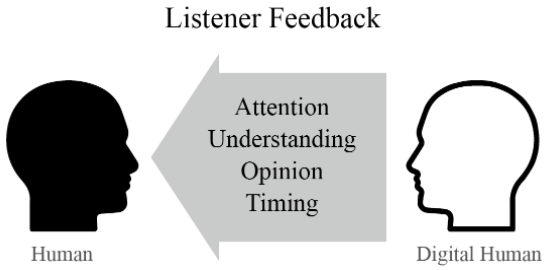

Figure 2 Listener Feedback Model

Attention is an expression of the listener's willingness and ability to recognize a message.[30] Attention can help users interact with agents to provide even more listener feedback.[32] A representative expression of attention is staring at the speaker. Gaze can signal usually that the speaker's continuing encouragement of utterance and that communication channels is open. Yoichi found in healthcare studies that patients pay more attention to agents when returning listener feedback while the patient is speaking.[33] Oh studied the degree of attention conveyed by nodding, audio and audio-visual feedback.[34] The robot was evaluated more positively when it displayed hand and arm gestures with words and asked participants to pay attention to the robot during the interaction.[35]. Allwood and Cerrato found that nodding the head conveys that the listener is paying attention and further triggers a sympathetic reaction.[36]

The listener understands the speaker's intentions through the language information, and the speaker monitors the listener to see if what the speaker wants to convey has been achieved.[32] The listener can express understanding by nodding or staring.[31, 37] Nakano et al. found that nonverbal cues perceived as positive evidence of comprehen-

sion were context-dependent. They also found that staring at the speaker was interpreted as evidence of incomprehension that provoked further explanation from the speaker.[31]
The speaker checks how the partner receives the message. Listeners can express their opinions (acceptance, consent, preference, etc.) to make communication livelier. The expression of opinion may take an expression form similar to the above understanding element. Understanding, however, is simply focused on the listener's understanding of the information, and expression of opinion is an implicit confirmation of the understanding. Nodding proved to be very important because all participants responded "agree" when displayed alone. Smiling, nodding and raising eyebrows also received high marks as signs of consent.[31] When virtual agent Billie requests confirmation, nodding is considered an acceptance and shaking the head is considered an expression of rejection. On the other hand, if the user nodded while agent "Billie" presented the information, nodding was interpreted as evidence of understanding.[38]
Timing needs to be considered a digital human listener feedback element because it can provide an unnatural feeling and a sense that we are indeed talking. For human-like communication, proper timing of the response to feedback is important.[39, 40] Even a good expression of consent can cause misunderstanding in the process of conveying meaning if it appears when it is not appropriate. Also, timing can be a signal of turn-taking and can contribute to creating a natural and realistic digital human.[21]
[41] scrutinized when and how the listener inserted responses in line with the speaker's context. They suggested that the speaker's gaze mediates this cooperation. [42] The "Rapport Agent" creates rapport by providing feedback to the person speaking about the comics they have seen before. The camera analyzes the speaker and determines the appropriate moment to provide feedback with head nods, head shakes, head rolls, and gaze.[39] Previous investigation of nonverbal feedback from avatars or robots revealed that cues or reactions, such as head turns, are effective when they occur at meaningful times rather than at random times

4 Discussion

Existing studies have not explored the potential mediated effects of digital human appearance as no work has been done using digital human beings to any degree. Therefore, we felt the need for a new guide to digital humans in line with the rapid commercialization trend of digital humans. For natural communication, we showed that people had stronger behavioral and attitudinal responses to physically present agents as opposed to a virtual presence, but when both the physically implemented agent and the virtually implemented agent were presented, the potential for development as a good listener compared to the robot was found respectively. It is presumed that different results were derived depending on the appearance of the virtual agent and the degree of freedom to express it. Also, depending on the appearance and degree of freedom of the virtual agent, the response can be linked to the agent's overall satisfaction. Unnatural expressions or excessive repetition may cause discomfort, so it is necessary to consider providing natural and appropriate feedback. Based on this, we defined four feedback

functional factors for interactive agents to become good listeners and proposed a wide range of concepts that can be applied. The four functional feedback elements presented were summarized into three elements (attention, understanding, and expression) that overlap or have the same meaning in previous studies, then redefined a total of four elements by adding the "timing" elements that stand out in other nonverbal communication studies.

First, an expression that pays attention to the speaker is required. Second, whether the understanding of the content of the ignition is successful or not. Third, it should be possible to express the listener's opinion about the content. Finally, the expression of the listener's attention, understanding, and opinion should be expressed in a timely manner. Digital humans, which are currently commercialized, may be suitable as low-cost personal assistants because they can be less expensive than robots and less constrained by their environment of use. The digital human can be generally better than a robot in that it can represent behavior, emotions, gestures, and expressions like humans. All told, it suggests that the digital human has sufficient potential to be utilized as a virtual listener. From the robots and agents used in this study alone, it is not clear to what extent the results will be the same for each evaluation in different implementations with different agent types. In future research, it will be necessary to check whether there is an empirical effect as a good listener through the feedback presented.

5 Conclusion

This study is just the first to examine how listener feedback from digital humans can affect human conversations. As the use of various agents increases, studies focusing on specific contexts and the need for nonverbal representation design studies in digital human conversation are shown. For several reasons, it was not possible to go deep into the functional analysis of listener feedback as originally intended in this study. We have just begun exploring listener feedback in digital humans, which should be combined with a larger number of studies in the future. In the next study, we intend to verify the four feedback factors proposed in this study through experiments to create a digital human for conversation. Thus, our ultimate goal is to build a digital human that users will want to talk to.

References

1. Kim, J., Kim, W., Nam, J., Song, H.: "I Can Feel Your Empathic Voice": Effects of Nonverbal Vocal Cues in Voice User Interface. Ext. Abstr. 2020 CHI Conf. Hum. Factors Comput. Syst. 1–8 (2020). https://doi.org/10.1145/3334480.3383075.
2. Gobron, S., Ahn, J., Garcia, D., Silvestre, Q., Thalmann, D., Boulic, R.: An Event-Based Ar- chitecture to Manage Virtual Human Non-Verbal Communication in 3D Chatting Environment. In: Perales, F.J., Fisher, R.B., and Moeslund, T.B. (eds.) Articulated Motion and Deformable Objects. pp. 58–68. Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642- 31567-1_6.

3. Judee K, B., Laura K, L.K.G., Valerie Manusov: Nonverbal signals. SAGE Handb. Interpers. Commun. 239–280 (2011).
4. Ekman, P., Friesen, W.V.: Nonverbal Leakage and Clues to Deception. Psychiatry 321. 88–106 (1969).
5. Cassell, J.: Nudge Nudge Wink Wink: Elements of Face-to-Face Conversation for Embodied Conversational Agents. (2000).
6. Bergmann, K., Kopp, S., Eyssel, F.: Individualized Gesturing Outperforms Average Gesturing
7. –Evaluating Gesture Production in Virtual Humans. Intell. Virtual Agents. 104–117 (2010). https://doi.org/10.1007/978-3-642-15892-6_11.
8. Davis, R.O., Wan, L.L., Vincent, J., Lee, Y.J.: The effects of virtual human gesture frequency and reduced video speed on satisfaction and learning outcomes. Educ. Technol. Res. Dev. (2021). https://doi.org/10.1007/s11423-021-10010-x.
9. Traum, D.: Models of Culture for Virtual Human Conversation. Univers. Access Hum.-Com- put. Interact. Appl. Serv. 5616, 434–440 (2009). https://doi.org/10.1007/978-3-642-02713-0_46. 9. Carrozzino, M.A., Galdieri, R., Machidon, O.M., Bergamasco, M.: Do Virtual Humans Dream of Digital Sheep? IEEE Comput. Graph. Appl. 40, 71–83 (2020). https://doi.org/10.1109/MCG.2020.2993345.
10. Leite, I., Martinho, C., Paiva, A.: Social Robots for Long-Term Interaction: A Survey. Int. J. Soc. Robot. 5, 291–308 (2013). https://doi.org/10.1007/s12369-013-0178-y.
11. David, D., Matu, S., David, O.: Robot-Based Psychotherapy: Concepts Development, State of the Art, and New Directions. Int. J. Cogn. Ther. 7, 192–210 (2014). https://doi.org/10.1521/ijct.2014.7.2.192.
12. Forbrig, P., Bundea, A.-N.: Modelling the Collaboration of a Patient and an Assisting Human- oid Robot During Training Tasks. Hum.-Comput. Interact. Multimodal Nat. Interact. 592–602 (2020). https://doi.org/10.1007/978-3-030-49062-1_40.
13. Trovato, G., Lopez, A., Paredes Venero, R., Cuellar, F.: Security and guidance: Two roles for a humanoid robot in an interaction experiment. (2017). https://doi.org/10.1109/ROMAN.2017.8172307.
14. Nakanishi, J.: Can a Humanoid Robot Engage in Heartwarming Interaction Service at a Hotel? Proc. 6th Int. Conf. Hum.-Agent Interact. (2018).
15. Johnson, D.O., Cuijpers, R.H., Kathrin, P., J, van de V.A.A.: Exploring the Entertainment Value of Playing Games with a Humanoid Robot. Int. J. Soc. Robot. 8, 247–269 (2016). http://dx.doi.org/10.1007/s12369-015-0331-x.
16. Hasunuma, H., Kobayashi, M., Moriyama, H., Itoko, T., Yanagihara, Y., Ueno, T., Ohya, K., Yokoil, K.: A tele-operated humanoid robot drives a lift truck. Proc. 2002 IEEE Int. Conf. Robot. Autom. Cat No02CH37292. 3, 2246–2252 vol.3 (2002). https://doi.org/10.1109/ROBOT.2002.1013566.
17. Kose, H., Yorganci, R., Algan, E.H., Syrdal, D.S.: Evaluation of the Robot Assisted Sign Language Tutoring Using Video-Based Studies. Int. J. Soc. Robot. 4, 273–283 (2012). https://doi.org/10.1007/s12369-012-0142-2.
18. Rahman, S.M.M.: Generating human-like social motion in a human-looking humanoid robot: The biomimetic approach. 2013 IEEE Int. Conf. Robot. Biomim. ROBIO. 1377–1383 (2013). https://doi.org/10.1109/ROBIO.2013.6739657.
19. Schroder, M., Heylen, D., Poggi, I.: Perception of Non-Verbal Emotional Listener Feedback. 4 (2006).
20. Bevacqua, E., Heylen, D., Pelachaud, C., Tellier, M.: Facial Feedback Signals for ECAs. AISB. 328–334 (2007).

21. Huang, L., Morency, L.-P., Gratch, J.: Learning Backchannel Prediction Model from Parasocial Consensus Sampling: A Subjective Evaluation. 6356, 172 (2010). https://doi.org/10.1007/978-3-642-15892-6_17.
22. Al-Fedaghi, S.: A Conceptual Foundation for the Shannon-Weaver Model of Communication. Int. J. Soft Comput. 7, 12–19 (2012). https://doi.org/10.3923/ijscomp.2012.12.19.
23. Yamato, J., Shinozawa, K., Naya, F., Kogure, K.: Effects of Conversational Agent and Robot on User Decision. (2000).
24. Powers, A., Kiesler, S., Fussell, S., Torrey, C.: Comparing a computer agent with a humanoid robot. Proceeding ACMIEEE Int. Conf. Hum.-Robot Interact. - HRI 07. 145 (2007). https://doi.org/10.1145/1228716.1228736.
25. Kiesler, S., Powers, A., Fussell, S.R., Torrey, C.: Anthropomorphic Interactions with a Robot and Robot-Like Agent. Soc. Cogn. 26, 169–181 (2008). http://dx.doi.org/10.1521/soco.2008.26.2.169.
26. Heerink, M., Kröse, B., Evers, V., Wielinga, B.: Influence of Social Presence on Acceptance of an Assistive Social Robot and Screen Agent by Elderly Users. Adv. Robot. 23, 1909–1923 (2009). https://doi.org/10.1163/016918609X12518783330289.
27. Bainbridge, W.A., Hart, J.W., Kim, E.S., Scassellati, B.: The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. Int. J. Soc. Robot. 3, 41–52 (2011). https://doi.org/10.1007/s12369-010-0082-7.
28. Jung, J., Kanda, T., Kim, M.-S.: Guidelines for Contextual Motion Design of a Humanoid Robot. Int. J. Soc. Robot. 5, 153–169 (2013). https://doi.org/10.1007/s12369-012-0175-6.
29. Wang, I., Ruiz, J.: Examining the Use of Nonverbal Communication in Virtual Agents. Int. J. Human–Computer Interact. 0, 1–26 (2021). https://doi.org/10.1080/10447318.2021.1898851. 30.Allwood, J.: Feedback in Second Language Acquisition. Adult Lang. Acquis. Cross Linguist. Perspect. II Results. 196–235 (1993).
30. Heylen, D., Bevacqua, E., Pelachaud, C., Poggi, I., Gratch, J., Schröder, M.: Generating Listening Behaviour. In: Cowie, R., Pelachaud, C., and Petta, P. (eds.) Emotion-Oriented Systems. pp. 321–347. Springer Berlin Heidelberg, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-15184-2_17.
31. Buschmeier, H., Kopp, S.: Communicative Listener Feedback in Human–Agent Interaction: Artificial Speakers Need to Be Attentive and Adaptive. 9 (2018).
32. Sakai, Y., Nonaka, Y., Yasuda, K., Nakano, Y.I.: Listener agent for elderly people with dementia. In: Proceedings of the seventh annual ACM/IEEE international conference on Human-
33. Robot Interaction. pp. 199–200. Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2157689.2157754.
34. Oh, C.S., Bailenson, J.N., Welch, G.F.: A Systematic Review of Social Presence: Definition, Antecedents, and Implications. Front. Robot. AI. 0, (2018). https://doi.org/10.3389/frobt.2018.00114.
35. Salem, M., Eyssel, F., Rohlfing, K., Kopp, S., Joublin, F.: Effects of Gesture on the Perception of Psychological Anthropomorphism: A Case Study with a Humanoid Robot. In: Mutlu, B., Bart- neck, C., Ham, J., Evers, V., and Kanda, T. (eds.) Social Robotics. pp. 31–41. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25504-5_4.
36. Allwood, J., Cerrato, L.: A study of gestural feedback expressions. 13.
37. Nakano, Y., Reinstein, G., Stocky, T., Cassell, J.: Towards a Model of Face-to-Face Grounding. Proc. 41st Annu. Meet. Assoc. Comput. Linguist. 553–561 (2003). https://doi.org/10.3115/1075096.1075166.

38. Buschmeier, H., Kopp, S.: Towards Conversational Agents That Attend to and Adapt to Com- municative User Feedback. Intell. Virtual Agents. 169–182 (2011). https://doi.org/10.1007/978- 3-642-23974-8_19.
39. Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M., Kuzuoka, H.: Precision timing in human-robot interaction: coordination of head movement and utterance. Proc. SIGCHI Conf. Hum. Factors Comput. Syst. 131–140 (2008). https://doi.org/10.1145/1357054.1357077.
40. Poppe, R., Truong, K.P., Reidsma, D., Heylen, D.: Backchannel Strategies for Artificial Listeners. In: Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., and Safonova, A. (eds.) Intelli- gent Virtual Agents. pp. 146–158. Springer, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15892-6_16.
41. Bavelas, J.B., Coates, L., Johnson, T.: Listener Responses as a Collaborative Process: The Role of Gaze. J. Commun. 52, 566–580 (2002). https://doi.org/10.1111/j.1460-2466.2002.tb02562.x.
42. Gratch, J., Okhmatovskaia, A., Lamothe, F., Marsella, S., Morales, M., van der Werf, R.J., Morency, L.-P.: Virtual Rapport. Intell. Virtual Agents. 14–27 (2006). https://doi.org/10.1007/11821830_2.