



Communication in Human-Robot Interaction

Andrea Bonarini¹

Published online: 27 August 2020
© The Author(s) 2020

Abstract

Purpose of Review To present the multi-faceted aspects of communication between robot and humans (HRI), putting in evidence that it is not limited to language-based interaction, but it includes all aspects that are relevant in communication among physical beings, exploiting all the available sensor channels.

Recent Findings For specific purposes, machine learning algorithms could be exploited when data sets and appropriate algorithms are available.

Summary Together with linguistic aspects, physical aspects play an important role in HRI and make the difference with respect to the more limited human-computer interaction (HCI). A review of the recent literature about the exploitation of different interaction channels is presented. The interpretation of signals and the production of appropriate communication actions require to consider psychological, sociological, and practical aspects, which may affect the performance. Communication is just one of the functionalities of an interactive robot and, as all the others, will need to be benchmarked to support the possibility for social robots to reach a real market.

Keywords Human-robot interaction · Communication · Robot design

Introduction

Robots that are involved in communication are growing exponentially [1], and they will increase even faster as the communication abilities will open new applications. Social robots [2, 3•] used in public spaces (for instance, hotels [4, 5], malls [6, 7], airports [8, 9], hospitals), education [10•], assistance [11, 12], and personal care [13–15], co-bots [16, 17] used in production plants, but also smart toys [18] and autonomous cars [19, 20], all need to interact effectively with people.

Interaction may be based on communicative acts [21, 22], performed intentionally to produce some effect in the interacting agent(s), but also obtained by unintentional acts, since it is impossible not to communicate when a channel is open between two agents [23].

Human-robot interaction requires from both sides the exploitation of different sense channels [24], typically hearing, sight, and touch, respectively, presented in sections “[The Hearing Channel: Sounds and Speech](#),” “[The Sight Channel: Light and More](#),” and “[The Touch Channel](#).” The signals should be not only produced and detected along these channels but need to be elaborated to produce information compatible with the decision function of both the human and robot agents, so that appropriate actions can be selected as a consequence of the communicative act. Interpretation could be either programmed, or learned, as discussed in “[Learning to Interact](#).”

In communication, so also in HRI, it is important that all the signals involved on the different channels be coherent in order to obtain effective message exchange and to establish a good relationship between the interacting agents. From the point of view of robot expression, this could be obtained by considering all the limitations imposed by sensors, computational power, mechanical implementation, and the role to be played, to keep the robot coherently placed on the Mori’s curve [25], possibly refraining from trying to achieve impossible performances and making a robot able to play a role compatible with its physical, mechanical, and computational features.

This article belongs to the Topical Collection on *Service and Interactive Robotics*

✉ Andrea Bonarini
andrea.bonarini@polimi.it

¹ Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milan, Italy

The quality of communication between robot and humans should be evaluated to support the effectiveness of the robots to perform tasks where communication is needed, in particular for social robots. “**Benchmarking: the “Real” Performance**” section mentions some of the efforts that are being performed in this direction. They are part of a much larger set of activities aimed at providing some kind of certification of the robot’s abilities, needed to make them reach the real market, issue still open in most potential situations [26].

The Hearing Channel: Sounds and Speech

Sounds can be used by robots to express emotional content by explicit production, as we have been used with R2D2 and BB8 robots of the Star Wars saga [27, 28], or by exploiting the “natural” sounds that motors are producing [29].

The production of content of a speech can be done directly by text-to-speech systems [30], which may require to define some structure of the dialogue a priori and to integrate the speech production in a framework for the interaction. Recently, this activity can be learned, typically by deep learning, to produce dialogue models [31]. A rich speech production may come from the use of dialogue management systems [32, 33] that can produce or adapt the structure of the dialogue online by using rules, statistical models, or machine learning [34]. These systems are often computationally expensive and require computation that cannot be done on board but needs a connection to large systems, like Watson [35]. This solution may lead to lag issues, due to the time needed to generate the proper text to speech and to the unpredictable delays due to the network supporting the connection. Although these lags may produce non-fluent interaction, some tricks may be used to limit this issue, such as simulating thinking, either gesturally, or with a short text, or with generic interjections, like: “Ehm...”.

The speech production can include signals to convey emotion, typically through prosody [36, 37], which make the interaction more natural and linked to the context defined by the content and the situation.

Speech interpretation by a machine is knowing a great improvement in the last years, also due to the development of home assistants [38], which, again, need a connection to a provider that could interpret the speech, with the consequent possibility of lags. While question-answering dialogues have achieved a good, market-grade quality, several challenges to obtain fluent general dialogue are still open [39]. Most of the solutions are not using the traditional natural understanding procedures, based on signal analysis, phoneme detection, grammars, and text structure interpretation, but exploit deep learning models trained off-line and able to interpret many of the possible sentences, possibly integrated in dialogue management systems.

From prosody and analysis of text, it is possible to capture some emotional content of the human speech [40–42], which can be used to adapt the dialogue and the expression of the robot.

Short commands, useful in directive interaction, can also be successfully interpreted by systems on-chip [43], which can be hosted on-board without requiring network connections to providers. Recent developments on both deep learning algorithms and hardware technology are bringing on-chip also more powerful speech understanding systems, reaching with FPGA technology more than real-time performances [44, 45].

The Sight Channel: Light and More

Visual Interaction: Robot to Human

Robots may exploit the vision channel of people to produce messages through light, images, or motion.

Light can be emitted by LED or other light sources that by exploiting color, intensity, and rhythm may convey emotional states or messages with a content [46]. These can also be organized in matrices, which can depict eyes and mouth, providing an immediate expression of affective and attention signals.

Interaction through a screen, often a touch screen, allows to convey quite a lot of information, either textual or visual. This is a solution taken by many social robots to overcome possible issues affecting other channels. In many cases, it may seem not natural, being in front of a humanoid, for instance, to have to read from the screen on its chest and to push virtual buttons on it, but this is becoming common, giving to the robot the minor role of “screen bearer” and diminishing the feeling of rich interaction with an autonomous agent. In most cases, this is a way to circumvent the current limits of a full speech interaction but also an effective way to produce complex interactions, such as in the case of asking to select an element in a relatively long list of alternatives.

Images are most often reproduced on screens, but they can also be projected on the floor or on objects, e.g., to convey specific information about the intention of the robot [47]. Other surfaces can also be used for projection; an interesting example is the Furhat head [48], where facial expressions are projected from inside an opaline face giving it the possibility to show rich and highly believable facial and emotional movements at a cost and complexity relatively lower than physical and mechanical faces.

A last, interesting possibility for screens is to use them to show either an animated face or parts of it sometimes integrated in physical faces (e.g., [49]), with a much less natural effect than the mentioned Furhat, or even actual faces of people interacting with the user through the robot, in a telepresence experience. These robots are used to have a remote

audiovisual contact with people that cannot be reached in a specific moment, as we have seen in the recent COVID-19 pandemics, and most of them do not even pretend to be more than a bare screen bearer.

Another way to exploit the visual channel of the human for the robot is to move its body or parts of it. This is very important also to accompany signals presented on other channels. It has been said that body language conveys most of the communication content in humans [23], and it is indeed important also for robots, to support their claims of animacy [50] and improve naturalness of the interaction. Most of the movements are pretending to mime as much as possible the analogous movements in humans [51], with possible limitations given by the mechanical implementation of joints, usually different from the biological ones, which bring to inconsistency, mostly in movements (uncanny valley [25]). It is possible to exploit bioinspired movements also in non-bioinspired robot bodies [52–54], obtaining recognizable expressions [55], as it was studied for cartoon animation [56–58].

Visual Interaction: Human to Robot

The robot has to understand the explicit and implicit communicative cues people produce with their body, mostly the affective expressions. Although many systems have been proposed to detect emotion from facial expressions [59, 60], they often have relatively low accuracy, leading to too quick shift in interpretation and quite demanding requirements both about face resolution and the expression of facial cues [61], often required to be unnaturally expanded to be recognized. Work has still to be done on body expressions [62] and on subtle cues, whose detection is also limited by sensor resolution, the learning models which cannot come with too high complexity, and the situations where interaction actually occurs, with subjects moving fast in front of the robot, and reaching positions out of the camera range.

Explicit, ample gestures can be easily detected by cameras, mostly to be interpreted as commands. More natural human activities can be recognized, at least under constrained situations not so common in social robot applications, since most of the models have been developed for surveillance or other purposes related to entertainment and require settings not all common for social robots, such as depth cameras, fixed in the environment, presence of a single user, models developed in controlled environments, and subjects distant from the camera. Moreover, only limited sets of actions have been considered in the data sets used to learn the available models, mostly by deep learning [63], and many actions interesting for HRI are not included in those sets. Reliable identification of common gestures, in a wild environment, from a mobile camera mounted on a social robot is still an open research issue.

A simpler, although less informative, interaction channel related to vision is based on low-cost range sensors (sonar,

infrared), which provide the distance from an unidentified object, possibly a moving person. The analysis of the dynamics of these signals can establish interesting interaction in low-cost, low computational power applications, such as robotic toys used in games.

The Touch Channel

Some robots rely on touch for at least some of the possible interactions.

Simple touch detectors, such as buttons, are integrated in many robotic toys, while resistive or capacitive sensors are common in many other robots, often used to detect affective gestures such as hugs, caresses, and punches (e.g., [52, 64]). More complex, distributed, and expensive sensors can be used to detect also punctual interactions, similarly to an artificial skin [65].

Manipulation of small robots can also be interpreted from accelerometer and gyroscope sensors, widely used to detect all sort of activity in smart phones and watches, thus enabling interaction through, usually implicit, communication acts. For instance, in [66], these data are used to interpret the manipulation of a plush robot intended to be used by autistic children, in order to objectively evaluate their activity and to have the robot react to undesired actions, such as being thrown to others. Accelerometers can also be used to interpret gestures or human activities. For instance, in [67], a human player involved in a robogame wears an accelerometer that gives implicitly to the opponent robot information about the type and the quality of activity that the person is performing, thus allowing the robot to adapt its playing style to the human player.

There are also situations, e.g., in robotic rehabilitation, in robogames, or in cobots, where the way the robot enters in contact with the human body should convey messages about safety and confidence, which require this channel to be effectively expressed.

Learning to Interact

Machine learning plays a relevant role in supporting different aspects of the communication between robots and humans, in particular in supporting the basic interpretation of signals: image content from camera images, speech content from audio signal, manipulation characteristics from contact sensors, and accelerometers. Another important aspect concerns the generation of behaviors, including planning and execution.

We have already mentioned many possible aspects of HRI communication that exploit machine learning, and we only present in this section some general considerations.

Interpreting the Signals

In most cases, signal interpretation is based on classification, today performed in many cases by using quite complex architectures belonging to the wide category of *deep learning*. In many cases, models are available to classify objects and people's actions: they have been developed spending a large amount of time and effort, and they are regarded as general models, good for many purposes. Being all layered models, it is also possible to retain the first, more data intensive levels, and train models on specific situations, for instance, for the recognition of activities or objects not included in the original data set. Usually, the last layers start from higher level features, requiring less time and effort to be learned.

Key points for the application of deep learning in HRI are the need of extensive data sets covering all the aspects interesting for the specific application and the need of massive computational power to learn the models. While this second issue can be addressed by off-line learning activity and with the support of cloud computing and data centers, the first one is critical and leads in many situations to the use of tools and models that some member of the community kindly made available, with few possibilities of tuning or obtaining what would really be needed for the specific application. In other situations, the collection of proper data sets is simply not possible, so the need for a different learning approach is emerging, but still not addressed.

Learning and Adaptation of Interactive Behaviors

Behaviors can be learned by following different approaches.

Imitation learning requires that a task be performed by some other agent (either human or another robot), and descriptions of both situation and taken actions are considered to build the new model. This approach is generalized by *supervised learning* where the proper behaviors for given situations are directly provided in their representational form.

Another promising approach is *reinforcement learning*, where some evaluation of the performance of the robot is used to promote correct actions and discard the worst ones.

To implement effective social robots, learning in the actual situations where they will operate is crucial. In particular, it is important to model not only the multimodal features describing the person the robot is interacting with, but the whole situation, including other possible persons, objects, surroundings, and possibly anticipating what will happen next [68]. Replicating the multimodal interactive actions could also be learned, e.g., modeling synchronous speech and gestures for a humanoid robot [69].

The complexity of interaction requires, mostly in long-term activities [70], to understand what are the general attitudes, personality [71], preferences, and possible limitations of the interlocutor(s) and try to match them [72]. The identification

of these aspects, and consequent adaptation of the robot behavior to the specific situation, including the selection of the appropriate multimodal communicative actions, can be done by learning models of general attitudes, trying to classify the interlocutor from the interaction, to identify the type of the most appropriate modality, and to apply the learned model to generate the proper interaction. It has been observed both for verbal interaction (e.g., [73]) and for non-verbal interaction (e.g., in robotic games [74, 75]) that robots matching the characteristics of the interacting people can have better performance.

Given safety and performance issues in real environments, and the usually long learning time, needed to identify complex models general enough to be used in the desired range of situations, learning is often done in simulation, with possible issues about how much realistic could be the learning experience in a simulated world that cannot include real people, whose behavior is difficult to simulate. The same issues hold for *adversarial learning*, where two learning systems learn from each other, again in a simulated environment, the only one where it is possible to perform the needed, very large number of iterations. Also for behaviors, different learning approaches may be beneficial, at least for some aspects and for complex applications.

Benchmarking: the “Real” Performance

“Robot benchmarking can be defined as an objective performance evaluation of a robot system/subsystem under controlled, reproducible conditions. [...] A benchmark includes a set of metrics together with a proper interpretation, allowing the evaluation of the performance of the system/subsystem under test according to well-specified objective criteria. In particular, a benchmark can be used to certify properties and functionalities, and therefore takes a key role in demonstrating the worth of specific solutions to prospective adopters, be they companies contemplating the realization of new products, or their clients interested in the purchase of such products” [76].

Benchmarking is becoming relevant also in HRI when an objective evaluation of the performance of the robot would be required to match the market requirements. HRI benchmarking activities are still in their infancy [77–79], but their importance will increase as certification processes will be defined for interacting robots, and the real market will require guarantees for value, performance, and safety.

Conclusion

We have presented a concise list of issues concerning communication in HRI. It is evident that communication includes transmission of signals among the communicating agents as

explicit, implicit, and involuntary interaction stimuli, in an integrated flow, that has to be considered as a whole.

We have left apart all considerations about physical aspects of the robots—such as shape, dimension, skin material, and weight—which are also important elements that affect communication, but would have required as much space, dedicated to robot design; also, this is knowing a new season, since social robots will face a market of people used at high-quality design.

Despite the great efforts required to make acceptable the communication with a complex device as a robot, considering the different aspects and the inherent limits involved in this activity, the scientific community facing these challenges is growing exponentially, striving to contribute to the definition of objects that could be really considered as companions in our activities.

Funding Information Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement.

Compliance with Ethical Standards

Conflict of Interest The author declares that he has no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance

1. World Service Robots 2019. International Federation of Robotics, 2019.
2. Leite I, Martinho C, Paiva A. Social robots for long-term interaction: a survey. *Int J Soc Robot.* 2013;5(2):291–308.
3. Breazeal C, Dautenhahn K, Kanda T. Social robotics. In: Springer handbook of robotics; Springer; 2016. p. 1935–72. **This paper surveys some of the principal research trends in Social Robotics and its application to human–robot interaction (HRI).**
4. Pinillos R, Marcos S, Feliz R, Zalama E, Gómez-García-Bermejo J. Long-term assessment of a service robot in a hotel environment. *Robot Auton Syst.* 2016;79:40–57.
5. Yu C-E. Humanlike robots as employees in the hotel industry: thematic content analysis of online reviews. *J Hosp Mark Manag.* 2020;29(1):22–38.
6. Sabelli AM, Kanda T. Robovie as a mascot: a qualitative study for long-term presence of robots in a shopping mall. *Int J Soc Robot.* 2016;8(2):211–21.
7. Niemelä M, Heikkilä P, Lammi H, Oksman V. Shopping mall robots—opportunities and constraints from the retailer and manager perspective. In: *International Conference on Social Robotics*; Springer; 2017. p. 485–94.
8. Nielsen S, Bonnerup E, Hansen AK, Nilsson J, Nellesmann LJ, Hansen KD, Hammsrshoi D. Subjective experience of interacting with a social robot at a Danish airport1. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*; 2018. p. 1163–1170.
9. Tonkin M, Vitale J, Herse S, Williams M-A, Judge W, Wang X. Design methodology for the UX of HRI: a field study of a commercial social robot at an airport. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*; 2018. p. 407–415.
10. Belpaeme T, Kennedy J, Ramachandran A, Scassellati B, Tanaka F. Social robots for education: a review. *Science Robotics.* 2018;3(21). <https://doi.org/10.1126/scirobotics.aat5954>. **This paper is a review of the possible application of social robots in education, some of the relative technical challenges, and puts in evidence how the robot's aspect and behavior may affect learning outcomes.**
11. Cabibihan J-J, Javed H, Ang M, Aljunied SM. Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. *Int J Soc Robot.* 2013;5(4):593–618.
12. Pennisi P, Tonacci A, Tartarisco G, Billeci L, Ruta L, Gangemi S, et al. Autism and social robotics: a systematic review. *Autism Res.* 2016;9(2):165–83.
13. Alonso SG, Hamrioui S, de la Torre Díez I, Cruz EM, López-Coronado M, Franco M. Social robots for people with aging and dementia: a systematic review of literature. *Telemed J E Health.* 2019;25(7):533–40.
14. Kachouie R, Sedighadeli S, Khosla R, Chu M-T. Socially assistive robots in elderly care: a mixed-method systematic literature review. *Int J Hum Comput Interact.* 2014;30(5):369–93.
15. Broadbent E, Stafford R, MacDonald B. Acceptance of healthcare robots for the older population: review and future directions. *Int J Soc Robot.* 2009;1(4):319.
16. Sherwani F, Asad MM, Ibrahim BSKK. Collaborative robots and industrial revolution 4.0 (ir 4.0). In *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*. IEEE; 2020. p. 1–5.
17. Terzioğlu Y, Mutlu B, Sahin E. Designing social cues for collaborative robots: the role of gaze and breathing in human-robot collaboration. In: *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, HRI 20*. New York: Association for Computing Machinery; 2020. p. 343357.
18. Hobby Products Report 2018 - Toys and games. Statista. 2019.
19. Basu C, Singhal M. Trust dynamics in human autonomous vehicle interaction: a review of trust models. In *2016 AAAI Spring Symposium Series*; 2016.
20. Kun AL, et al. Human-machine interaction for vehicles: review and outlook. *Foundations and Trends R in Human–Computer Interaction.* 2018;11(4):201–293.
21. Searle JR, Searle JR. *Speech acts: an essay in the philosophy of language*, vol. 626: Cambridge University Press; 1969.
22. Newcomb TM. An approach to the study of communicative acts. *Psychol Rev.* 1953;60(6):393.

23. Watzlawick P, Bavelas JB, Jackson DD. Pragmatics of human communication: a study of interactional patterns, pathologies and paradoxes: WW Norton & Company; 1967.
24. Yan H, Ang MH, Poo AN. A survey on perception methods for human–robot interaction in social robots. *Int J Soc Robot.* 2014;6(1):85–119.
25. Mori M, et al. The uncanny valley. *Energy.* 1970;7(4):33–5.
26. Hoffman G. Anki, Jibo, and Kuri: What we can learn from social robots that didn't make it. *IEEE Spectrum*, <https://spectrum.ieee.org/automaton/robotics/home-robots/anki-jibo-and-kuri-what-we-can-learn-from-social-robotics-failures?> 2019.
27. Schwenk M, Arras KO. R2-d2 reloaded: a flexible sound synthesis system for sonic human-robot interaction design. In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*; IEEE; 2014. p. 161–7.
28. Jee E-S, Jeong Y-J, Kim CH, Kobayashi H. Sound design for emotion and intention expression of socially interactive robots. *Intell Serv Robot.* 2010;3(3):199–206. **This paper presents an original analysis of the potential of sound to express robot's intention and emotion.**
29. Tennent H, Moore D, Jung M, Ju W. Good vibrations: How consequential sounds affect perception of robotic arms. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*; 2017. p. 928–35.
30. Taylor P. *Text-to-speech synthesis*: Cambridge University Press; 2009.
31. Zhang Q, Guo B, Wang H, Liang Y, Hao S, Zhiwen Y. AI powered text generation for harmonious human-machine interaction: current state and future directions. *arXiv preprint arXiv.* 2019:1905.01984.
32. Schlangen D, Skantze G. A general, abstract model of incremental dialogue processing. *Dialogue & Discourse.* 2011;2(1):83–111.
33. Chen H, Liu X, Yin D, Tang J. A survey on dialogue systems: recent advances and new frontiers. *ACM SIGKDD Explor Newsl.* 2017;19(2):25–35. **This paper presents recent results about the application of Deep Learning for natural language processing.**
34. Su P-H, Gasic M, Mrksic N, Rojas-Barahona L, Ultes S, VanDyke D, et al. Continuously learning neural dialogue management. *arXiv preprint arXiv.* 2016:1606.02689.
35. High R. *The era of cognitive systems: an inside look at IBM Watson and how it works.* IBM Corporation, Redbooks; 2012. p. 1–16.
36. Crumpton J, Bethel CL. A survey of using vocal prosody to convey emotion in robot speech. *Int J Soc Robot.* 2016;8(2):271–85.
37. Li Y, Ishi CT, Inoue K, Nakamura S, Kawahara T. Expressing reactive emotion based on multimodal emotion recognition for natural conversation in human–robot interaction. *Adv Robot.* 2019;33(20):1030–41.
38. Kepuska V, Bohouta G. Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In: *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*. IEEE; 2018. p. 99–103.
39. Ward NG, DeVault D. Challenges in building highly-interactive dialog systems. *AI Mag.* 2016;37(4):7–18.
40. James J, Watson CI, Stoakes H. Influence of prosodic features and semantics on secondary emotion production and perception. In: *ICPhS2019 International Congress of Phonetic Sciences*; 2019.
41. Li Y, Ishi CT, Ward N, Inoue K, Nakamura S, Takanashi K, Kawahara T. Emotion recognition by combining prosody and sentiment analysis for expressing reactive emotion by humanoid robot. In: *2017 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE; 2017. p. 1356–1359.
42. Sailunaz K, Dhaliwal M, Rokne J, Alhaji R. Emotion detection from text and speech: a survey. *Soc Netw Anal Min.* 2018;8(1):28.
43. Stemmer G, Georges M, Hofer J, Rozen P, Bauer JG, Nowicki J, et al. Speech recognition and understanding on hardware-accelerated dsp. In: *Interspeech*; 2017. p. 2036–7.
44. Lee M, Hwang K, Park J, Choi S, Shin S, Sung W. FPGA-based low-power speech recognition with recurrent neural networks. In: *2016 IEEE International Workshop on Signal Processing Systems (SiPS)*; IEEE; 2016. p. 230–5.
45. Gao C, Braun S, Kiselev I, Anumula J, Delbruck T, Liu SC. Real-time speech recognition for IoT purpose using a delta recurrent neural network accelerator. In: *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*; IEEE; 2019. p. 1–5.
46. Cha E, Kim Y, Fong T, Mataric MJ, et al. A survey of non-verbal signaling methods for non-humanoid robots. *Foundations and Trends R in Robotics.* 2018;6(4):211–323.
47. Andersen RS, Madsen O, Moeslund TB, Amor HB. Projecting robot intentions into human environments. In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*; IEEE; 2016. p. 294–301.
48. Al Moubayed S, Beskow J, Skantze G, Granström B. Furhat: a back-projected human-like robot head for multiparty human-machine interaction. In: *Cognitive behavioural systems*: Springer; 2012. p. 114–30.
49. Chen H, Park HW, Breazeal C. Teaching and learning with children: impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Comput Educ.* 2020;150:103836.
50. Kim M, Yi S, Lee D. Between living and nonliving: young children's animacy judgments and reasoning about humanoid robots. *PLoS One.* 2019;14(6).
51. Ishiguro H, Libera FD. *Geminoid Studies: Science and Technologies for Humanlike Teleoperated Androids*: Springer; 2018. **This paper describes the concepts and technology of Geminoids, the most human-like robots.**
52. Bonarini A, Garzotto F, Gelsomini M, Romero M, Clasadonte F, Yilmaz ANÇ. A huggable, mobile robot for developmental disorder interventions in a multi-modal interaction space. In: *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2016)*. New York: IEEE Computer Press; 2016. p. 823–30.
53. Kozima H, Michalowski MP, Nakagawa C. Keepon. *Int J Soc Robot.* 2009;1(1):3–18.
54. Anderson-Bashan L, Megidish B, Erel H, Wald I, Hoffman G, Zuckerman O, et al. The greeting machine: an abstract robotic object for opening encounters. In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*; IEEE; 2018. p. 595–602.
55. Bonarini A. Can my robotic home cleaner be happy? Issues about emotional expression in non-bio-inspired robots. *Adapt Behav.* 2016;24(5):335–49.
56. Thomas F, Johnston O, Thomas F. *The illusion of life: Disney animation*. New York: Hyperion; 1995.
57. Ribeiro T, Paiva A. Nutty-based robot animation—principles and practices. *arXiv preprint arXiv.* 2019:1904.02898.
58. Schulz T, Soma R. The role of animacy for communicating behavior in robots. In: *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*; 2018. p. 676–80.
59. Marechal C, Ajewski DM, Tyburek K, Prokopowicz P, Bougueroua L, Ancourt C, et al. Survey on ai-based multimodal methods for emotion detection. In: *High-Performance Modelling and Simulation for Big Data Applications*: Springer; 2019. p. 307–24.
60. Mehta D, Siddiqui MFH, Javaid AY. Facial emotion recognition: a survey and real-world user experiences in mixed reality. *Sensors.* 2018;18(2):416.

61. Deshmukh S, Patwardhan M, Mahajan A. Survey on real-time facial expression recognition techniques. *Iet Biometrics*. 2016;5(3): 155–63.
62. Noroozi F, Kaminska D, Corneanu C, Sapinski T, Escalera S, Anbarjafari G. Survey on emotional body gesture recognition. *IEEE Trans Affect Comp*. 2018. <https://doi.org/10.1109/TAFFC.2018.2874986>.
63. Wang J, Chen Y, Hao S, Peng X, Hu L. Deep learning for sensor-based activity recognition: a survey. *Pattern Recogn Lett*. 2019;119:3–11.
64. Shibata T, Mitsui T, Wada K, Touda A, Kumasaka T, Tagami K, et al. Mental commit robot and its application to therapy of children. In: 2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Proceedings (Cat. No. 01TH8556), vol. 2: IEEE; 2001. p. 1053–8.
65. Cheng G, Dean-Leon E, Bergner F, Olvera JRG, Leboutet Q, Mittendorfer P. A comprehensive realization of robot skin: sensors, sensing, control, and applications. *Proc IEEE*. 2019;107(10):2034–51.
66. Alhaddad AY, Cabibihan J-J, Bonarini A. Influence of reaction time in the emotional response of a companion robot to a child's aggressive interaction. *Int J Soc Robot*. 2020:1–13.
67. Oliveira E, Orrù D, Nascimento T, Bonarini A. Modeling player activity in a physical interactive robot game scenario. In: Proceedings of the 5th International Conference on Human Agent Interaction; 2017. p. 411–4.
68. Tapus A, Bandera A, Vazquez-Martin R, Calderita LV. Perceiving the person and their interactions with the others for social robotics—a review. *Pattern Recogn Lett*. 2019;118:3–13.
69. Aly A, Tapus A. Speech to head gesture mapping in multimodal human robot interaction. In: *Service Orientation in Holonic and Multi-Agent Manufacturing Control*: Springer; 2012. p. 183–96.
70. Tapus A, Tapus C, Mataric M. Long term learning and online robot behavior adaptation for individuals with physical and cognitive impairments. In: *Field and service robotics*: Springer; 2010. p. 389–98.
71. Paunonen SV, Ashton MC. Big five factors and facets and the prediction of behavior. *J Pers Soc Psychol*. 2001;81(3):524.
72. Bandler R, Grinder J. *The structure of magic*. Palo Alto: Science and Behavior Books; 1975.
73. Cruz-Maya A, Agrigoroaie R, Tapus A. Improving user's performance by motivation: matching robot interaction strategy with users regulatory state. In: *International Conference on Social Robotics*: Springer; 2017. p. 464–73.
74. de Oliveira E, Donadoni L, Boriero S, Bonarini A. Deceptive actions to improve the attribution of rationality to playing robotic agents. *Int J Soc Robot*. 2020:1–15.
75. Bonarini A, Boriero S, de Oliveira E. Robot player adaptation to human opponents in physical, competitive robogames. In: *Proceedings of the 29th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2020)*. New York: IEEE Computer Press; 2020. p. in press.
76. Amigoni F, Bonarini A, Fontana G, Matteucci M, Schiaffonati V. Benchmarking through competitions. In: *European Robotics Forum—Workshop on Robot Competitions: Benchmarking, Technology Transfer, and Education*, vol. 604; 2013.
77. Feil-Seifer D, Skinner K, Mataric MJ. Benchmarks for evaluating socially assistive robotics. *Interact Stud*. 2007;8(3):423–39.
78. Amigoni F, Bastianelli E, Berghofer J, Bonarini A, Fontana G, Hochgeschwender N, et al. Competitions for benchmarking: task and functionality scoring complete performance assessment. *IEEE Robot Autom Mag*. 2015;22(3):53–61.
79. Chrysostomou D, Barattini P, Kildal J, Wang Y, Fo J, Dautenhahn K, et al. Rehri'17-towards reproducible HRI experiments: scientific endeavors, benchmarking and standardization. In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*; 2017. p. 421–2.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.