



Pepper as a Storyteller: Exploring the Effect of Human vs. Robot Voice on Children's Emotional Experience

Berardina de Carolis, Francesca D'errico, Veronica Rossano

► To cite this version:

Berardina de Carolis, Francesca D'errico, Veronica Rossano. Pepper as a Storyteller: Exploring the Effect of Human vs. Robot Voice on Children's Emotional Experience. 18th IFIP Conference on Human-Computer Interaction (INTERACT), Aug 2021, Bari, Italy. pp.471-480, 10.1007/978-3-030-85616-8_27 . hal-04196844

HAL Id: hal-04196844

<https://inria.hal.science/hal-04196844>

Submitted on 5 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Pepper as a Storyteller: Exploring the Effect of Human vs. Robot Voice on Children's Emotional Experience

Berardina De Carolis¹, Francesca D'Errico² and Veronica Rossano¹

¹ Department of Computer Science, University of Bari Aldo Moro, Bari, Italy

² For.Psi.Com Department, University of Bari Aldo Moro, Bari, Italy
berardina.decarolis@uniba.it, francesca.derrico@uniba.it,
veronica.rossano@uniba.it

Abstract. Social robots are autonomous entities able to engage humans at the emotional and social level. They are being used in several domains, especially in those where kids are the primary users (i.e., education, games, rehabilitation). The paper presents an experience in which the social robot Pepper is used as a storyteller. A storyteller robot should engage humans by combining its verbal and non-verbal behaviors and 'immerse' the user into the story. Therefore, to design an engaging and effective storytelling experience we started to address a first design issue: does a human voice have an advantage over a synthesized voice of the robot in this context? To this aim, two versions of the same story for kids from 8 to 9 y.o. have been developed. The social robot Pepper was used to tell the story in two modalities. In the first modality, Pepper storyteller was designed as a kind of audiobook in which the robot had just the role of a device, but the story was narrated by a human voice; in the second modality, Pepper was designed to tell the story using its own voice combined with non-verbal behaviors. The system has been tested in a real context and results show that Pepper's voice affected more positively the children's emotional experience, also by giving the children the perception that they learn more easily.

Keywords: Storytelling, Social Robots, Educational Technology.

1 Introduction

Since childhood storytelling is one of the most appreciated activities by kids. Storytelling is a participatory and immersive experience that allows children to acquire language skills in a dynamic and entertaining way. In recent years, the educational power of storytelling has also been newly discovered. Storytelling was found to be a valuable tool for motivating students to listen, engaging them during lessons, improving reading skills, and as a springboard for beginning units and skill development [1].

Recently, several digital solutions have been applied to the domain of storytelling [2,3] and each solution aims at enhancing the value of storytelling using the features of the digital medium. Among these solutions, the most similar to the original experience of the human storytelling activity is the audiobook [4]. An audiobook allows hearing a human voice reading a story, and sometimes there are sound effects or background music to allow the users to have an emotional experience. The audiobooks are widely

accepted and appreciated, but they are not an actual innovation, but they represent a traditional approach in which communication happens through a new device.

To contribute to this context, the research herein aims at investigating whether social robots can engage children in listening stories. Social robots are embodied, autonomous agents that communicate and interact with humans on a social and emotional level. They represent an emerging field of research focused on developing a social intelligence that aims to maintain the illusion of dealing with a human being [5]. Thanks to their ability to interact with humans naturally and familiarly, social robots are spreading more and more often into human life not only for entertainment purposes, but also to support users in their daily activities, or in teaching and educational settings [6, 7].

The paper presents a case study in which the social robot Pepper acts as a storyteller. The main aim of the study was to investigate the effect of the social interaction of this innovative device on children's emotional experience. The research questions will be: (RQ1) is there a difference in user's emotions comparing human voice and synthesized voice of the robot and (RQ2) is there a difference in user's experience in the two conditions?

To this aim, the story was implemented using two different approaches. In the first one, Pepper was used as a device to convey the story using a pre-recorded human voice, in this context Pepper acts as a sort of audiobook. In the second approach, Pepper tells the story using its own voice, slightly modified with appropriate pitch and speed variations. In both cases, the storytelling was enriched by typical Pepper's non-verbal behaviors. The paper presents a between-subjects study involving 34 children, 17 of them interact with Pepper which uses the human pre-recorded voice (Real), the others interact with Pepper which uses its own voice (Robot). Results show that children feel positive emotions and user-experience mainly in the Robot condition, pointing how voice signal can affect the perception of robot authenticity.

2 Related Works

Storytelling is the most ancient form to transfer knowledge from across the centuries. Telling stories has been successfully used to narrate experiences, teach social norms, past traditions, and so on. Stories provide a realistic and authentic opportunity to capture children's attention and help them listen and learn more actively than other forms of instruction [8]. Stories are not just for literary narratives but can be used to illustrate even the most complex and abstract concepts or subjects (e.g., math and science). According to Bruner [9], stories engage our thinking, emotions, and imagination all at once. As listeners, we participate in the story with both mind and body as we enter the narrative world and react to it. Stories help teachers reach novices in ways they cannot with other strategies.

When storytelling becomes digital, through new media and innovative technologies, it can be enriched with other communicative elements to enhance the user experience. Digital storytelling has been shown to be a valuable tool to acquire content [10,11], to enhance motivation and reflection for deep learning [12, 13].

In [14] digital storytelling has been applied to develop digital literacy skills in a higher educational context. Students were asked to design and develop a short self-

introduction video using technological tools including cameras, microphones, and video editing software. The findings suggest that students acquired or improved their digital skills and communication skills, since thanks to digital storytelling they were able to communicate and express their ideas effectively using digital media. Similar findings are reported in [15] where digital storytelling helps secondary school students to enhance their writing skills.

The effectiveness of digital storytelling has been investigated also in the context of health professions education. The investigation of the literature in this context underlines that the creation and use of health professionals' digital stories can positively enhance learning [16]. Digital storytelling has been used also as a tool to foster reflection, in [17] some undergraduate students were asked to develop digital stories to reflect on their study abroad experiences. Findings show that the process of creating digital stories can promote deep reflection.

Robots as storytellers have a long tradition and their evolution has been in line with technological support. One of the first talking robots was launched on the market in 1959 by Mattel®, Chatty Cathy, a doll able to tell a story when a string was pulled. Recently, smart speakers, such as Alexa, are becoming popular digital storytellers.

In the research context, the trend is to imply social robots as a storyteller in order to investigate if beyond learning and motivational improvements there are more deep benefits in terms of emotional or evaluative processes and how they can affect their attitude.

Leite et al, for example, use socially assistive robots to help children build their emotional intelligence skills through interactive storytelling activities [18].

The results of the study on the influence of stories narrated either by a humanoid robot or by a human teacher suggest a positive effect of the expressive behavior in robot storytelling, whose efficacy is comparable to the human expressing the same behavior [19]. Social robotic storytellers have been used also to teach new vocabulary to children showing increased vocabulary learning [20].

In [21] a comparison between recorded voice of an actor with text-to-speech synthesis and the effect on the participants was investigated. The emotional reaction of the participant towards the story told was measured. The results were twofold, they found out that the physically embodied robot attracts the listener as compared to a virtual embodiment and that a human voice is preferable to the text-to-speech.

All these studies have in common the fact that social robots are a good tool for digital storytelling, and they have a positive impact on children's engagement and learning. The new issue herein investigated is that the robot, in this case, is used in this study to convey the human voice, like an audiobook, and to tell a story. The differences in children's emotions during the storytelling will be measured.

3 User Study

In order to design an engaging and effective storytelling experience we started to address a first design issue: does a human voice (Real) have an advantage over a synthesized voice of the Pepper robot (Robot) in this context? This section describes the user study designed for this purpose. In particular, the research questions the study was addressing are:

- RQ1: are the user's emotions (concentration, enjoyment, anger and feeling of learning) different in the two conditions: Real and Robot?
- RQ2: is the user's experience (pleasantness, clarity and easiness during the interaction) different in the two conditions: Real and Robot?

To answer these RQs different measures were collected to evaluate the general engagement in listening to the story, the experience with the robot, and the affective and evaluative elements that came into play during the story listening. A between-subject design was used to collect data, since two different groups of people tested each condition, so that each subject was only exposed to a single use of Pepper.

3.1 Equipment

The Robot. The robot used in this study is the Pepper, a semi-humanoid robot developed by SoftBank Robotics (formerly known as Aldebaran Robotics). It is an omnidirectional wheeled humanoid robot 1.21 m tall, with 17 joints and 20 degrees of freedom. The robot is equipped with several LEDs to signal and support communication. These are software-controlled to change colors and intensity. It is equipped with four directional microphones in its head that allow it to detect the origin of entries and thus to turn its face to whoever is talking. The microphones can eventually be used to analyze the voice tone and therefore interpret the emotional state of the interlocutor. A 3D camera and two HD cameras allow Pepper to recognize not only faces but also images and objects. The robot is equipped with 20 motors that allow it to move its head, back and arms. It has other accessories, such as six laser sensors, two ultrasonic sensors and three obstacle detectors.

The Story. The story used in this first experience is “Tu Sei la Regina delle Oasi Incantate” (in English “You are the Queen of Charmed Oases”) [22] that in 1994 was published as a story game since it has different points in which the reader can choose what the heroine, the princess, can act in the story. Choices influence the story and its ending. Only in some cases the princess finds and rescues her prince, in other endings she fails the rescue mission. The interesting part of the story is that some choices have to be made simply by choosing to trust the characters you meet (should I trust the shadow I encountered in the garden?).



Fig. 1. Pepper robot during a storytelling session

The questionnaires. Two questionnaires were used to collect data. The pre-test one, filled in before the activity, was aimed at collecting users' demographic data. It was composed of 6 questions: 2 measuring the confidence with a robot, and 4 measuring the confidence with stories. The post-test was composed of 14 questions, and it was aimed at measuring:

- Children's emotions: concentration, easiness, enjoyment, feeling able, feeling of learning.
- Children User Experience: pleasant to listen to , pleasant way of telling story, easy to play, easy to understand, Clear image.

			
NIENTE	POCO	ABBASTANZA	MOLTO

Fig. 2. An example of the emoticons used as answers (labels: Never, Little, Enough, High)

In both questionnaires, children have to express a score between 1 (very low) to 4 (very high). To help children the questions were very simple, and the answers were represented using emoticons (Fig. 2).

3.2 Participants and Procedure

The study involved 34 children, (balanced for gender, 52% males), aged from 8 to 9 y.o. attending the third grade of a primary school in (anonymized). Children were divided into two groups: in the first group Pepper uses a human pre-recorded voice (Real), in the second group Pepper uses its own voice (Robot). The voice of Pepper was manipulated by modifying the pitch and the speed parameters in order to follow the same intonation of the human voice during story narration. Before the activity, the pupils underwent a pre-test to collect demographic information and to measure trust and potential enjoyment in playing with the robots and fun in storytelling activities and the user's perceived enjoyment of listening to the stories.

Then the activities, using two different robots, Peppers, were carried out at the same time in two different rooms: in one room the group Real listened to Pepper using a pre-recorded human voice, and in the second room the group Robot listened to Pepper as storyteller. During the activity, only one pupil at a time was asked to interact with Pepper and make their own choices to reach the end. In this way, all pupils could interact with Pepper.

At the end of the game session, the participants were asked to evaluate their experience with Pepper as a storyteller by answering the post-test.

3.3 Results and Discussions

As mentioned above, our participants before the interaction with Pepper answered the questions aimed at measuring their general level of confidence in playing with robots and their enjoyment in listening to stories. The analysis of data reports that the involved children did not show significant gender differences but a general low confidence in playing with robots and a medium level of enjoyment in playing with robots and in listening to stories. The analysis reveals a significant difference between the means of playing with robots and the remaining ones [$F(1, 24) = 7.49$; $p < 0.05$].

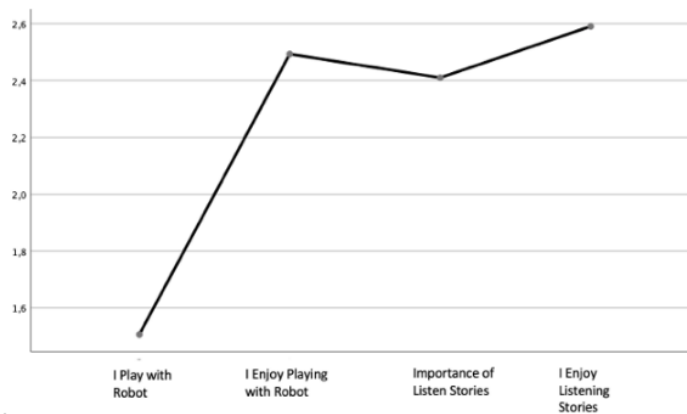


Fig. 3. Descriptive on level of confidence in robots and enjoyment in listening stories.

As to the experimental conditions (human voice vs robotic voice) an Anova repeated measures was performed, which revealed a within effect among the items of the children's emotions [$F(1, 24) = 12.75$; $p < 0.000$] especially considering the feeling of a positive sense of mastery - feeling able and feeling of having learned after the interaction with Pepper - and negative emotions ('anger') that is very low. Furthermore, including the experimental variable of voice type, children felt more positive emotions when interacted with a robotic voice than the human one [$F(1, 24) = 3.98$; $p < 0.05$], by feeling able, concentrated, at-ease, enjoyed and they also perceived to have learned more, maybe due to a kind of satisfied expectations. In the *Real* condition, with the human voice, children evaluated less positively their emotions since they feel a kind of violation of hearing a robotic voice, more coherently with Pepper (Figure 4).

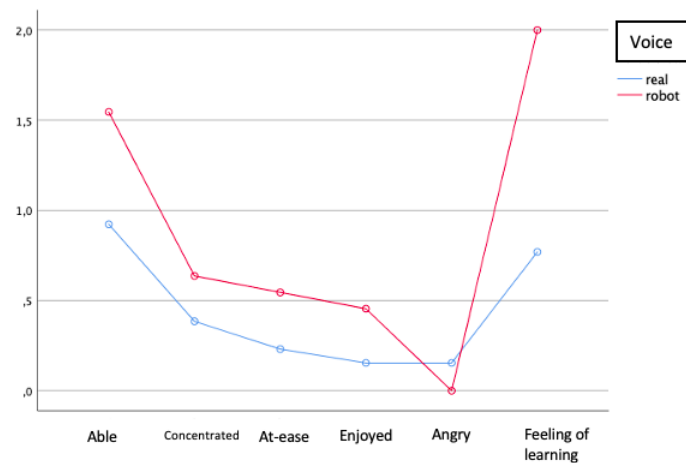


Fig. 4. Emotional experience of the children*type of voice (real or robot)

As regarding the user experience, emerged how children evaluate the interaction with Pepper pleasant to listen, easy to play and understand, pleasant also the way of telling stories, and image clarity, mainly when the voice of Pepper is robotic [$F(1, 24) = 3.97$; $p < 0.05$]; they evaluate a sense of pleasantness in listening to Pepper and this positive evaluate is also high in the story understanding (Figure 5). This result is also in line with the previous one on the children's emotionality, and it shows a tendency to feel positive emotions and evaluations toward Pepper with 'robotic' features.

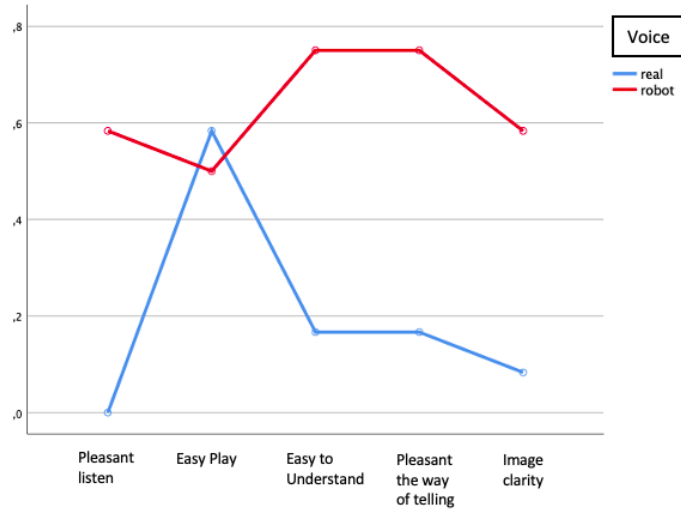


Fig. 5. User Experience of the children*type of voice.

4 Conclusions and future works

As the first step of our research for assessing the effectiveness of using a social robot, Pepper in this case, as a storyteller the issue of investigating whether there was an advantage in using a human voice instead of the robot's voice was addressed. To this aim, the study carried out shows that robot interaction is best evaluated by children when the social signals [23] expressed by the robot are consistent with the expectations of the users, in this case, the use of the robotic voice goes in this direction. When Pepper uses a coherent voice there is a lower feeling of disorientation instead present when they listen to the human voice. On the contrary of the results reported in [21], in our sample, this effect is quite clear, even if the children involved do not have many opportunities to interact with the robots, as stated before the interaction with Pepper.

This might be due to the fact that the robot used in that experiment has a human-like face, while Pepper has a cute but still robotic appearance.

Our results are quite promising from a learning point of view, since the children's emotional state was positive and their level of engagement and mastery on the task (easy to play and easy to understand), mainly in the robotic voice condition.

In the future, more stories will also be implemented to measure the effects of Pepper storytellers on knowledge acquisition or persuasive processes. Moreover, future studies will also measure the long-term effect of using Pepper as a storyteller on the affective and cognitive aspects of the learning process [24] by having more and regular sessions with children, who will interact with a robot having differentiated stance - humble vs dominant [25] - toward the young users.

References

1. Groce, R. D.: An experiential study of elementary teachers with the storytelling process: Interdisciplinary benefits associated with teacher training and classroom integration. *Reading Improvement*, 41(2), 122-129 (2004).
2. Alexander, B.: *The New Digital Storytelling: Creating Narratives with New Media--Revised and Updated Edition*. Abc-clio (2017).
3. Syafryadin, H., Salniwati, A. R. A. P.: Digital storytelling implementation for enhancing students' speaking ability in various text genres. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(4), 3147-3151 (2019).
4. Van Maas, S.: Opening the audiobook. *Comparative Literature*, 70(3), 337-356 (2018).
5. Malerba, D., Appice, A., Buono, P., Castellano, G., De Carolis, B., de Gemmis, M., Polignano, M., Rossano, V., Rudd, L. M.: Advanced programming of intelligent social robots. *Journal of e-Learning and Knowledge Society*, 15(2) (2019).
6. Westlund, J. M. K., Dickens, L., Jeong, S., Harris, P. L., DeSteno, D., Breazeal, C. L.: Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction*, 13, 1-9 (2017).
7. Castellano, G., De Carolis, B., D'Errico, F., Macchiarulo, N., & Rossano, V.: *Pepperecycle: Improving Children's Attitude Toward Recycling by Playing with a Social Robot*. *International Journal of Social Robotics*, 13(1), 97-111 (2021).
8. Isbell, R., Sobol, J., Lindauer, L., Lowrance, A.: The Effects of Storytelling and Story Reading on the Oral Language Complexity and Story Comprehension of Young Children. *Early Childhood Education Journal*, 32, 157-163 (2004).
9. Bruner, J.: *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press, (1986).
10. Robin, B. R.: Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory into practice*, 47(3), 220-228, (2008).
11. Rossano, V., and Roselli, T.: Game-based learning as effective learning method: An application of digital storytelling. In 2018 22nd International Conference Information Visualisation (IV), IEEE, pp. 542-546 (2018).
12. Barrett, H.: Researching and evaluating digital storytelling as a deep learning tool. In *Prof. Of Society for information technology & teacher education international conference*, pp. 647-654, AACE (2006).
13. Smeda, N., Dakich, E., Sharda, N.: The effectiveness of digital storytelling in the classrooms: a comprehensive study. *Smart Learning Environments*, 1(1), 1-21 (2014).
14. Chan, B. S. K., Churchill, D., Chiu, T. K. F.: Digital Literacy Learning In Higher Education Through Digital Storytelling Approach. *Journal of International Education Research (JIER)*, 13(1), 1-16 (2017).
15. Rong, L. P., Noor, N. M.: Digital Storytelling as a Creative Teaching Method in Promoting Secondary School Students' Writing Skills. *International Journal of Interactive Mobile Technologies*, 13(7), 10-11 (2019).
16. Moreau, K. A., Eady, K., Sikora, L., Horsley, T.: Digital storytelling in health professions education: a systematic review. *BMC medical education*, 18(1), 1-9 (2018).
17. Hamilton, A., Rubin, D., Tarrant, M., Gleason, M.: Digital storytelling as a tool for fostering reflection. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 31(1), 59-73 (2019).
18. Leite, I., McCoy, M., Lohani, M., Ullman, D., Salomons, N., Stokes, C., ... Scassellati, B.: Emotional storytelling in the classroom: Individual versus group interaction between children and robots. In *Proc. of 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 75-82, IEEE (2015).

19. Conti, D., Di Nuovo, A., Cirasa, C., Di Nuovo, S.: A comparison of kindergarten storytelling by human and humanoid robot with different social behavior. In Proc. of the Companion of the 2017 ACM/IEEE International Conference on Human-robot Interaction, pp. 97-98, ACM (2017).
20. Westlund, J.K., Breazeal, C.: The interplay of robot language level with children's language learning during storytelling. In Proc. of the 10th ACM/IEEE International Conference on Human-robot Interaction - Extended Abstracts ACM, pp. 65-66, ACM (2015).
21. Costa, S., Brunete, A., Bae, B.-C., and Mavridis, N.: Emotional storytelling using virtual and robotic agents, *International Journal of Humanoid Robot*, 15(3), (2018).
22. Fabri, S., Pasini, M.: *Tu sei la Regina delle Oasi Incantate*, ed. Giunti (1994).
23. Poggi, I., D'Errico, F.: Social signals: a framework in terms of goals and beliefs. *Cognitive Processing*, 13(2), 427-445 (2012).
24. D'Errico, F., Paciello, M., De Carolis, B., Vattani, A., Palestra, G., Anzivino, G.: Cognitive emotions in e-learning processes and their potential relationship with students' academic adjustment. *Journal of educational psychology*. 10 (1), 89-111 (2018).
25. D'Errico, F., & Poggi, I.: Tracking a leader's humility and its emotions from body, face and voice. In *Web Intelligence*, 17 (1), pp. 63-74 (2019). IOS Press.