International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.23, August 2019 – www.ijais.org

Storytelling Robot for Activities in the Alfadown Project

Adam Everson S. Araújo[1], Ariane Ruben C. Di Menezes[2], Kárita Cecília Rodrigues Silva[3], Edson Douglas F. De M. Santos[4], Juliana Santos de Souza Hannum[5], Talles Marcelo G. de A. Barbosa[6]

Pontifical Catholic University of Goiás, PUC-Goiás

ABSTRACT
The project "Computer technologies as a facilitator during the literacy process of people with down syndrome" known as Alfadown, is an extension project and developed by the Program of Reference in Social Inclusion (PRIS) of the Pontifical Catholic University of Goiás. The project aims to assist in the development of the written and verbal language of people with Down Syndrome (DS), through the use of computational technologies. The storytelling robot is a specific case of a sociable robot offering mechanisms that enables a natural and intuitive communication, while also aiming to replace the human during a narrative. The storytelling prototype used some elements from various technologies that includes, Emolab, Curious Learning, Application Programming Interfaces (API), Text to Speech from IBM Watson, and Unity. In order to evaluate volunteer learners of the Alfadown project, a questionnaire and stipulated activities were formulated for the implementation of test cases. To be able to measure the level of understanding of the story told and to attach data analysis from Emolab. This made it possible to analyze the following variables obtained from emolab: attention, engagement, valence, anger, contempt, disgust, fear, joy, sadness, and surprise. The conclusions obtained in this work, are in related to complexity, originality, and future works.

Keywords
Storyteller; Sociable Robot; Down Syndrome

1. INTRODUCTION
The storytelling robot is a specific case of a sociable robot, offering mechanisms that allows a more natural and intuitive communication [8]. These robots can be utilized in various applications, such as in teaching-learning, pediatric treatment, and motor training for people with special necessities. The same has the objective of replacing the human, during a narrative.

Playtime Computing is an environment that utilizes a combined reality, having physical objects interacting with virtual objects [4]. The main objective of this environment is to enable learning through playing. For this, playful objects are offered, while searching at the same time to explore the greatest number of interaction possibilities between the system components and the users, in general children. New activities are generated from the insertion of components into the environment, which can be virtual (projection) or real (electronic devices). TinkRBook is a system that offers interfaces for shared reading, serving to tell stories [9]. It seeks to create a new experience for the story reader with interactivity, by allowing the use of voice commands and touch, to modify a story. The authors of this project perceived that children learning is greater with toys than static texts.

Curious learning is a project whose objective is to offer computational tools that may be utilized in comprehending learning behaviors for early literacy [7][10]. One of its tools is an application with the same name that utilizes a virtual sociable robot, to help in early literacy by means of lettering activities, involving sounds and images, testing the ability to memorize. The P2PSTORY is a project that captured a dataset of children as storytellers and listeners in interactions peer-to-peer [6]. Its objective being to better understand the socio-emotional behaviors of children in storytelling, a grand challenge in the implementation of a sociable robot, tablets, and virtual agents. Natural interactions between child-to-child were carried out, assuming a future interaction of child-to-technology. For analysis, video and audio of each interaction was captured, detecting behavioral characteristics, prosodic features (pitch, speed, and pauses), personal characteristics, and child perceptions through a questionnaire.

The Pediatric Companion is a project that aims to minimize stress, anxiety and the pain of pediatric patients with a robot companion [5]. Helping to fill the gaps that exist in the socio-emotional support already provided to pediatric patients. Tests were made with the huggable robot (sociable robot), including a virtual character, a screen, and a plush teddy bear detecting the effects it caused. Initially it was detected that patients would be more eager to emotionally connect with the physical robot than a virtual character. This sociable robot is capable of reproducing the voice of an external person, while also performing movements (like simulating a hug).

The project "Computer technologies as a facilitator during the literacy process of people with down syndrome" known as Alfadown, is an extension project and developed by the Program of Reference in Social Inclusion (PRIS) of the Pontifical Catholic University of Goiás. The project aims to assist in the development of the written and verbal language of people with Down Syndrome (DS), through the use of computational technologies.

The purpose of this study is the development of a sociable robot (storyteller) for activities in the Alfadown Project. The object of study is justified by the increase in interactivity (due to the use of multimedia resources), and the increase in availability (from the use of computational resources).

Section two covers material and methods, subdivided into technologies used, the developed prototype, pilot testing off field and field testing (in the Project Alfadown). Followed by the results and discussions obtained from the analyses of the questionnaires and technologies (software). Finalizing the conclusion with emphasis on complexity, originality, and future works. Lastly, the credits with the acknowledgements and the references.

2. MATERIALS AND METHODS
A. Technologies
The Affdex project provides libraries with specific functions, that enables face analysis [2]. Through the Affdex Project, the Emolab application was developed by D. R. Szturm [11][12]. Emolab is a technology (software) that is capable of

extracting expressions and emotions of present faces in recorded videos, camera captures and photos.

As shown in Fig. 1, when it detects a face it analyses the following variables: attention, engagement, valence (a measurement for how good or bad the experience was), anger, contempt, disgust, fear, joy, sadness, and surprise generating a respective value (within a range) for each variable. In this work, only previously recorded videos are used for analysis. Fig. 1 illustrates a screen prototype developed for the Emolab application, using libraries from the Affdex Project.

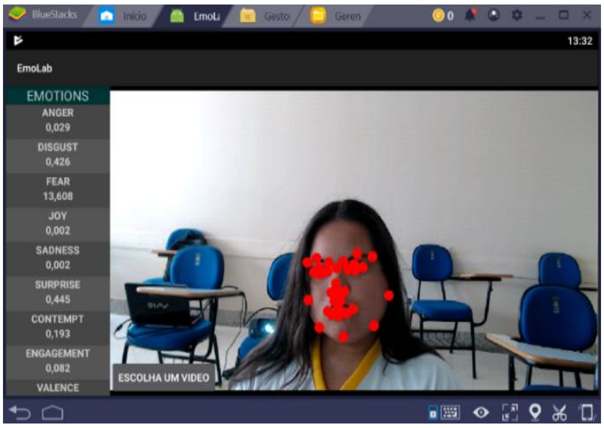


Fig. 1 An image showing the analysis of expressions and emotions of a participant by using the Emolab application

Curious Learning is an application used for early literacy [3]. The code is open source for developers and available in a GitHub [13] repository. Its graphic elements were used as the base application of the prototype.

The Application Programming Interface (API) Text to Speech (IBM Watson) is an API that converts text into speech, its package is available for use in a GitHub repository [14]. This work uses the Lite Plan that offers a free monthly conversion of 10,000 characters. There is also the Standard Plan that charges $0.02 USD for every 1,000 characters used. The Speech Synthesis Markup Language (SSML) is used to transmit text for speech synthesis, and to accept punctual alterations in the form of audio return, adapting to certain text peculiarities such as pitch, speed, amplitude, pauses, and classification of specific cases [1].

Unity is a useful application in maintaining and integrating miscellaneous packages. Allowing the manipulation of open source code projects [13], and executing them in the Android platform together with packages from the Text to Speech API [14]. In addition, for programming it uses the Integrated Development Environment (IDE) Microsoft Visual Studio.

B. Prototype

In this work, the base application uses graphic elements from the Curious Learning Project. Using its interfaces from the virtual sociable robot as background expanding possibilities by bringing a more playful environment. Fig. 2 demonstrates the scenario, where the monster represents a virtual sociable robot.

To utilize the functionalities of the API text to Speech, a button and textbox interface was used as seen in Fig. 2, including the API configuration for the selected Portuguese language. Since multiple devices were used, an account was created for each device with the available API and with their own credentials. In Fig. 2, the credentials are covered by a black stripe to avoid compromising the applications integrity.

The Unity platform was used to integrate all the packages [13][14], graphics editing, enabling electronic keys from the API, and to extend the Microsoft Visual Studio IDE used in coding for the C Sharp (C#) language.

C. Tests

After obtaining the parents or legal representative authorization, through the signed Free and Informed Consent Term (TCLE), testing began for the volunteer learners with SD. The tests were performed individually and in reserved rooms.

In initial testing, the software narrated the story "The Little Red Hen", with the support of an illustrative video. Afterwards, monitors evaluated the level of understanding of the story from the volunteer learners with SD, by means of a request for the individual to say what happened in the addressed story. In addition, questions about the presented story from the software would be asked in order to evaluate the volunteer comprehension.

In the second moment, the illustrated video is captured again. However, the narrator of the story "The Little Red Hen" is a human being. This way the monitors can reevaluate the level of understanding of the story, with a human as the narrator while following the same criteria of evaluation as previously.

In the third and last moment, videos were analyzed using Emolab, extracting data and filtering the most accurate ones for a better analysis. Thus, generating several varied results searching to explain the phenomena that occurred.

It is worth noting that to perform the tests, a pilot testing was first conducted off field, without the physical sociable robot, to estimate the amount of time and of possible unforeseen events during testing. Images of the scenario from the perspective of both the monitor and participant is demonstrated in Fig 3.

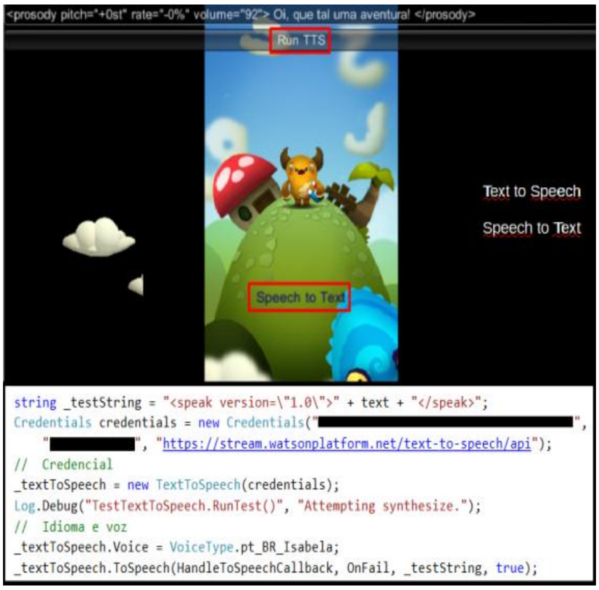


Fig. 2 Storytelling Software and a snip of code responsible for the functionality of the API Text to Speech

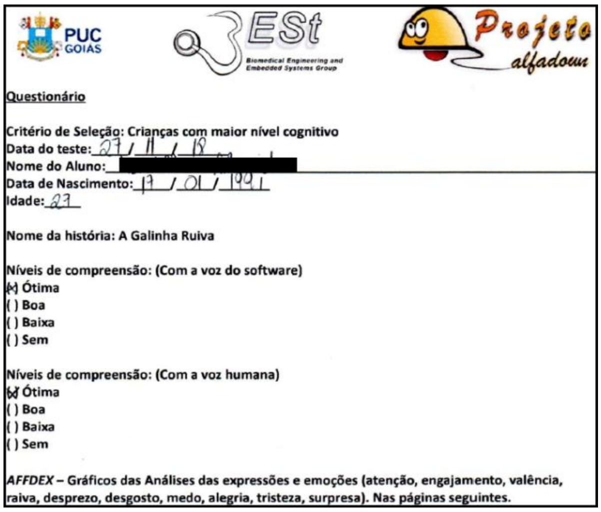
Fig. 3 Pilot testing scenario, tests carried outside of the Project AlfadowAfter concluding all the preliminary tests, the field testing in the Project Alfadown was carried out and Fig. 4 illustrates the testing environment

The equipment used for testing included two computers, one containing the storytelling software, and the other with an illustrative presentation of the story through slides shown by a projector. A speaker was used to enable a higher volume for the story teller, and a webcam for video capture. Lastly, an illustrative table that brings a playful environment and a plush chicken on top of the speakers, simulating a physical sociable robot.

Fig 4 Testing environment in the Project Alfadown

Fig 4 shows a participant watching a story, first the story is told by the software. Afterwards, the story is told by a human. The black stripe on the participants face is to help maintain their integrity. The computer next to the illustrated table, contains the storyteller software.

The questionnaire shown in Fig. 5 is a document that serves to measure the level of understanding of the story told and to attach data from the Emolab (Affdex) analysis. The objective of the questionnaire is to evaluate the participants, that are scored by the monitors Ariane and Kárita in two separate stages. After the software tells a story, and after the human retells the story. The scoring levels are shown in Fig. 5, the black stripe over the student's name is to help maintain their integrity.



Questionário

Critério de Seleção: Crianças com maior nível cognitivo
Data do teste: 27 / 11 / 18
Nome do Aluno: [redacted]
Data de Nascimento: 17 / 01 / 1991
Idade: 27

Nome da história: A Galinha Ruiva

Níveis de compreensão: (Com a voz do software)
(x) Ótima
( ) Boa
( ) Baixa
( ) Sem

Níveis de compreensão: (Com a voz humana)
(x) Ótima
( ) Boa
( ) Baixa
( ) Sem

AFFDEX – Gráficos das Análises das expressões e emoções (atenção, engajamento, valência, raiva, desprezo, desgosto, medo, alegria, tristeza, surpresa). Nas páginas seguintes.

Fig 5 A participants Questionnaire

3. RESULTS

The results for the level of understanding of the participants are presented in Fig. 6, ranging from 0 to 3 (None-Great). In this way, comparative results were made from the understanding results of the learner with SD, through the software storyteller and storytelling by a human. The best results came from participants 1 and 8, with the worst result being from participant 10 that did not obtain understanding of the content with the software or the human. It can be observed in Fig. 6 that there are cases where the understanding with a human telling the story was superior to when the software was telling the story.

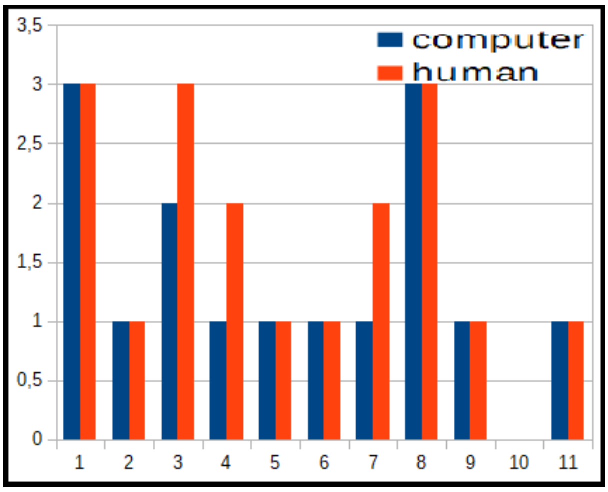


Fig 6 Bar graph displaying the levels of understanding of the participants in regards to the story

Through Emolab, it was possible to analyze the following variables: attention, engagement, valence, anger, contempt, disgust, fear, joy, sadness, and surprise. Each value compares to the individual in regards to their level of understanding, in order to assess which variables could be causing their level of understanding. Fig. 7 displays graphical results from three selected candidates which had good analyses, for a demonstration of the possible comparisons based on their respective averages. The other remaining participants wore glasses which proved to be a hindrance in performing the analysis.

Fig. 7 shows the average of attention, the remaining variables could also be generated, and the analyses of the graphs can be made together since one or more graphs could imply in other graphs. For participant 3 it cannot be said that the attention variable was important regarding the level of understanding, since the participant had a higher attention with a computer yet a better understanding with the human. Now for participant 7 it can be said that attention was an important variable for the level of understanding since the levels of attention reflect the understanding. For participant 6, it can be said that the levels of attention do not portray the level of understanding with exactness.

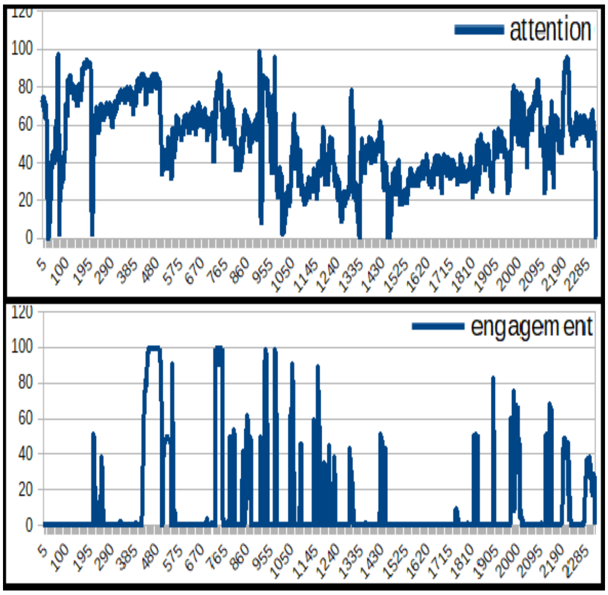


Fig. 7 Bar graph comparing the results of the average attention variable for 3 participants

Fig. 8 demonstrates the attention and engagement of participant 6 varying during the story told by the software and Fig. 9 shows the same variables but with a human telling the story. The attention graph for participant 6, presents which areas of the story had a higher relevance, generating new forms of individual interpretations, which can generate new conclusions, even about the type of story being told. Thus, by verifying all the variables: attention, engagement, valence, anger, contempt, disgust, fear, joy, sadness, and surprise, not only the two variables presented, it is possible to assess which activities would be the best for each individual participant. In addition, to estimate which variables are most important to improve for future works, for example, the use of a sociable robot that could minimize the level of sadness during an activity.

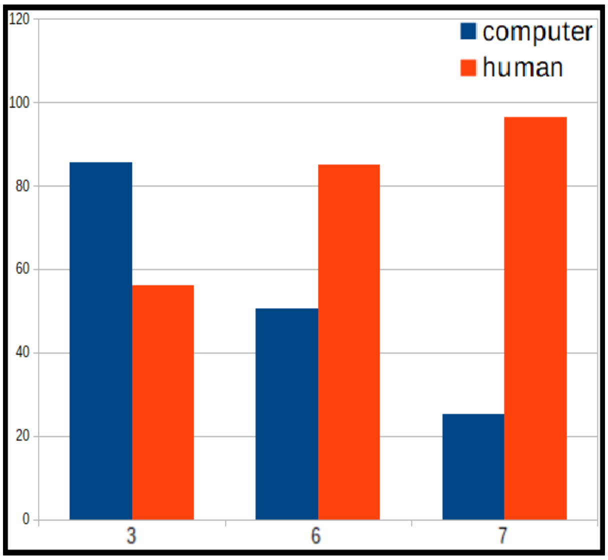


Fig 8 Analysis of participant 6 with the software telling the story, displaying attention and engagement variables

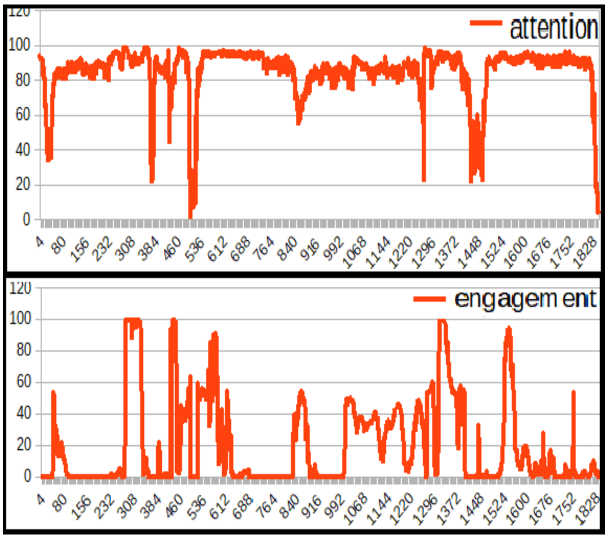


Fig 9 Analysis of participant 6 with a human telling the story, displaying attention and engagement variables

4. CONCLUSION

The conclusion obtained in this work, in relation to complexity, was about the implementation of the prototype, tests, and the analyses of the results. While in relation to originality, it was about the analysis of the interaction between the storyteller sociable robot with people that have Down Syndrome. For Future works it is in the open, an implementation of a moving sociable robot, and the integration of converting speech to functional text, for a better interactivity.

5. ACKNOWLEDGMENTS

This study was supported by FAPEG *(in portuguese, Fundação de Amparo à Pesquisa de Goiás)* at Pontifical Catholic University of Goiás. Special thanks, to Israel Brito, Jéssica Borges, Sammy Azar, Kaliston de Sousa, Neilto Ribeiro, Matheus Rodrigues, Jesuel Alves and Anelita Neila.

6. REFERENCES

[1] (2018) SSML elements. IBM Clould. [Online]. Available: https://console.bluemix.net/docs/services/text-to-speech/SSML-elements.html#elements

[2] (2018) Android.:) Affectiva. [Online]. Available: https://knowledge.affectiva.com/v3.2/docs/getting-started-with-the-emotion-sdk-for-unity

[3] (2018) Curious Learning. [Online]. Available: https://www.curiouslearning.org/

[4] C. Breazeal. (2018) Playtime Computing. MIT Media Lab. [Online]. Available: https://robotic.media.mit.edu/portfolio/playtime-computing/

[5] C. Breazeal. (2018) Pediatric Companion - Personal Robots Group. [Online]. Available: https://robotic.media.mit.edu/portfolio/pediatric-companion/

[6] C. Breazeal. (2018) Overview ‹ P2PSTORY: Dataset of Children as Storytellers and Listeners in Peer-to-Peer Interactions — MIT Media Lab. [Online]. Available: https://www.media.mit.edu/projects/p2pstory/overview/

[7] C. Breazeal. (2018) Curious Learning: Understanding Learning Behaviors for Early Literacy. MIT Media Lab. [Online]. Available: https://www.media.mit.edu/projects/curious-learning/overview/

[8] C. Breazeal. Sociable Machines: Expressive Social Exchange Between Humans and Robots. 2000. 264 f. Tese (Doutorado) - Curso de Electrical Engineering and Computer Science, Massachusetts Institute Of Technolociy, Santa Barbara, 2000.

[9] C. Breazeal. et al. ThinkRBook: Shared reading interface for storytelling. Global Village. Cambridge, MA. June 20–23, 2011.

[10] C. Breazeal. et al. Mobile Devices for Early Literacy Intervention and Research with Global Reach. Global Village. Edinburgh, UK. April 25–26, 2016.

[11] D. A. A. Santos. et al. Wearable device for literacy activities with people with down syndrome. Cambridge, MA, USA. November 3–5, 2017.

[12] D. R. Szturm. et al. Emotion Analysis Tool to Support People with Down Syndrome. Cambridge, MA, USA. October 5–7, 2018.

[13] (2018) Norad-Eduapp4syria/Feed the Monster at master · curiouslearning/Norad-Eduapp4syria · GitHub. [Online]. Available: https://github.com/curiouslearning/Norad-Eduapp4syria/tree/master/Feed%20the%20Monster

[14] (2018) GitHub - watson-developer-cloud/unity-sdk: Unity SDK to use the IBM Watson services. [Online]. Available: https://github.com/watson-developer-cloud/unity-sdk