Improving language skills and encouraging reading habits in primary education: A Pilot Study using NAO Robot

Dipesh Pandey[1], Anushka Subedi[2], and Deepti Mishra[3], *Senior Member, IEEE*

***Abstract*— NAO is a popular social robot that has been used in educational settings to assist with teaching as well as to increase engagement and motivation of students. While features like speech recognition, object detection, sensors in robots may be close to making them human-like, they are still a long way from being able to fully mimic human behavior. In this study, we present a pilot study that involves a quiz based interaction with primary level students for an English language learning course where these features are thoroughly tested and analysed. Based on the results of the study, it was found that the voice module of NAO requires refinement for smooth dialog-based communication, while the object detection engine performed fairly well in a quiz based scenario with primary level students.**

I. INTRODUCTION

Educational interventionists are continually seeking new ways to improve children's learning experiences. Designers are increasingly considering robots and other social agents as instructional tools, as a result, there has been increasing use of social robots for education and research purposes [4][19][17][14]. NAO is a humanoid robot created by SoftBank Robotics [2] and designed to work as an assistant to educators. Using several project based learning approaches, NAO is often used to develop problem solving and analytical skills as well as language acquisition among the students. It can be programmed using Choregraphe which provides a graphical interface with drag-and-drop features as well as using the Python SDK.

Social robots could provide unique support for young language learners due to their various strengths - their ability to perform actions and gestures in addition to being adaptive i.e., they can detect humans' motivational and educational needs through various sensors and change their behavior accordingly [11]. However, there is a dearth of empirical research and insufficient evidence supporting the unique benefits of robot tutors should not be taken as definitive [11]. Therefore, this study aims to explore and evaluate the use of NAO to aid a human teacher in English language learning in a Primary level education setting.

[1]Dipesh Pandey is a master's student at the Department of Computer Science (IDI), Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology (NTNU), Teknologivegen 22, 2815 Gjøvik, Norway; Email: dipeshp@ntnu.no

[2]Anushka Subedi is a master's student at the Department of Computer Science (IDI), Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology (NTNU), Teknologivegen 22, 2815 Gjøvik, Norway; Email: anushkas@ntnu.no

[3]Deepti Mishra is affiliated with Educational Technology Laboratory at the Department of Computer Science (IDI), Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology (NTNU), Teknologivegen 22, 2815 Gjøvik, Norway; Email: deepti.mishra@ntnu.no

Research team conducted brainstorming sessions with a team of primary school teachers to identify how Nao can be integrated in the current language learning course to improve students' language abilities and motivate them to read a variety of books. Subsequently, various scenarios were designed based on children story books followed by their implementation using Choregraphe to develop a dialog system that interacts with primary level students. These scenarios involve students of grade 1 and grade 2 who would be provided with story books as part of English language class and NAO would then interact with the kids to ask questions from those books. This paper describes two specific scenarios where in one, NAO asks questions related with the story similar to a human teacher aiming to conduct a dialog between them and in the other, NAO asks students to show pictures of objects used in the story so as to aid in visual learning of students.

Further, we evaluate the speech recognition and object detection engine of NAO with the help of a pilot test with a limited number of participants. The pilot test with representatives of the target group is essential in order to test the ability of NAO to hold conversation with kids and to make sure that they understand and react to the question asked to them. The outcomes of this pilot test would serve as an assessment of the whole system's effectiveness prior to the actual workshop with grade 1 and 2 students.

The present paper is structured as follows: In the next section, recent relevant studies are discussed. Section 3 presents the implementation of scenarios using Choregraphe followed by a section describing the procedure and data collection during evaluation. Section 5 discusses lessons learned in light of existing literature and finally the paper concludes with the conclusion, limitation and future work.

II. RELATED WORK

NAO humanoid robot has been used to facilitate the educational process for teachers and to make the class more interactive and entertaining to kids [22] [1] [16]. Tiago et al. [20] used NAO as tutor to play a multi-role serious game with an aim to support students studying geography on a multi-touch table. A study by Alkhalifah et al.[3] proposed a system to support fun learning through different activities and games for students in kindergarten. The activities also included quizzes where NAO would ask kids to touch their specific body part or say the name of the color as NAO changes his eye color.

Quiz can be considered as one of the fun ways to gain knowledge on various topics. In a recent study by Matsuura

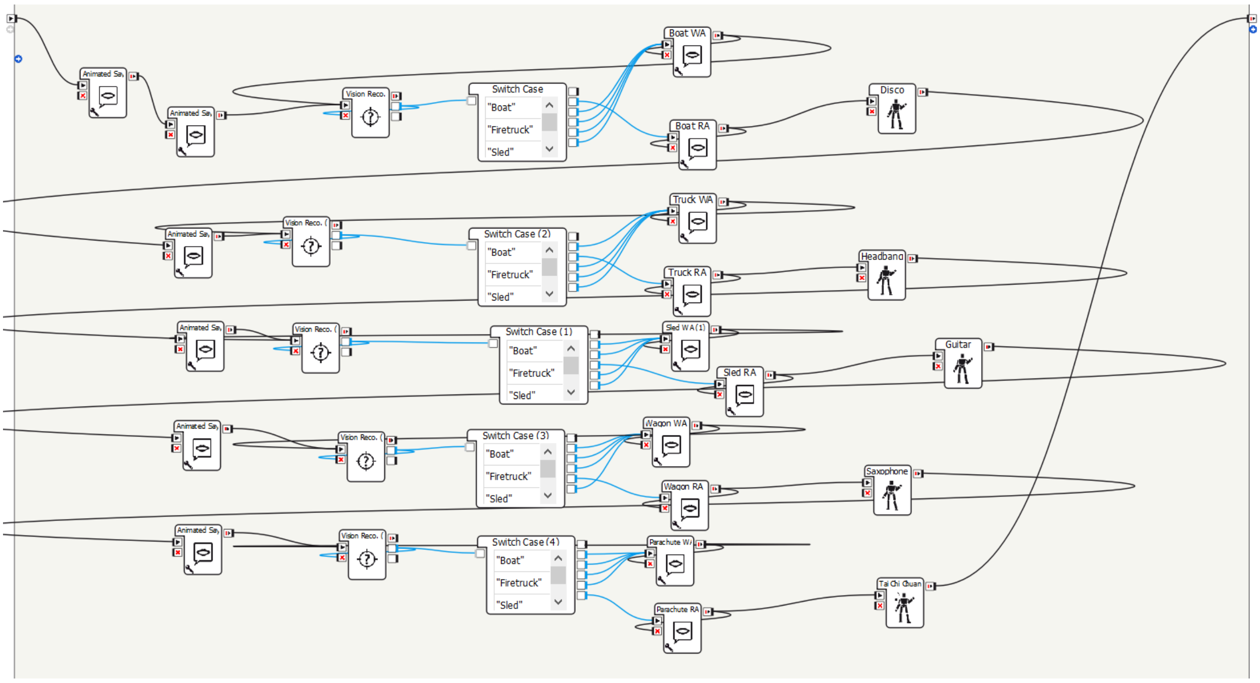

Fig. 1. Vision Recognition implemented in Choregraphe for Story Task 2

et al. [15], NAO Robot was used as a quiz presenter every week for a period of six months. According to the study, the students accepted the robot sympathetically, and the study developed children's interest in the quiz topic. The experiments from Shiomi et al. [21], show that using a social robot in a classroom setting can increase the individual curiosity of the students, although the overall curiosity doesn't seem to be very significant. Similarly, Eguchi and Okada [8] present a pilot study examining the experience of students participating in a robotics competition. They conclude that humanoid robots can be used in schools to provide assistance to teachers as well as other staff members. Another study by Pedro et al. [18] shows that the use of a robotic platform in class helps in the improvement of students' knowledge acquisition and increases their motivation and attention span.

Similarly, research has been done to explore the relation between age of the participants and their learning using robots. An exploratory analysis done by Wit et al. [7] showed that age played a role and older children learned more than the younger children.

Several researchers have looked into using social robots as an aid in learning first or second language [6][13] and majority of children find learning language with social robots engaging [9][23][24]. However, these interactions often hinge on verbal interaction to effectively achieve their goals [12]. The success of these interaction depends not only appropriate speech production by robots but also transcribing and understanding speech from young users as well as the ability of dialogue system to cope relatively well with noisy ,i.e., real world environments [12]. The work presented in this paper builds on the contribution of using social robots in education in addition to exploring the technical capabilities of NAO in regards to speech recognition and object detection.

III. Implementation

The implementation of the project was divided into three main sections: introduction, story task 1, and the story task 2. While both the first two parts were implemented using the dialog script functionality available with the NAO SDK, the last part involved object recognition using NAO camera.

A. *Introduction with NAO*

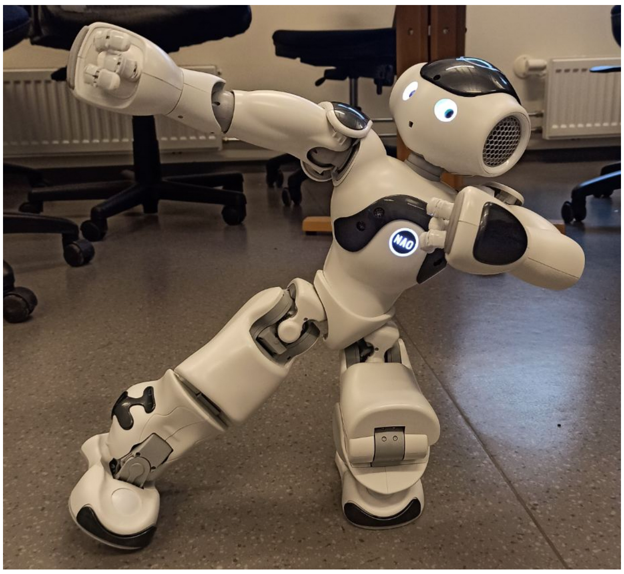

Fig. 2. Nao showing one of its moves during Story Task 2

First impressions are always very important in human-robot-interaction [10]. Keeping this in mind, several interesting interaction items were included in the introduction part. It was important that we make the robot show its capabilities to the children and make them familiar with it. While some primary level school children are already exposed to humanoid robots, for some it was the first time. So, we included greetings, using soft, friendly tone, showing off dance moves, responding to their generic commands like "sit down", "stand up", etc. in this session. Fig. 2 shows NAO performing the Tai Chi martial art step.

The implementation was done using Choregraphe which comes with the ready-to-use movement features like dance moves and body postures. We combined this with a dialog program built using QiChat, a library available with the Python SDK.

B. Story Task 1

In the primary level classes, students are taught the English language using story books by the teachers, and later asked related questions from the books. An attempt to simulate a similar scenario was made using NAO as an teaching assistant. Participants were asked to read a story book, after which some questions were asked from the same book. To implement this, NAO's speech recognition engine was used using QiChat, and a script was built based on the list of possible answers to the formed questions in the quiz. The syntax of QiChat script file allows for programming a dialog between the robot and the user in a flexible way. The first question in the quiz included five sentences while the rest of the question had one or two sentences. Q 2A was an action based question in which participants were instructed to touch the Robot's head. The goal of this question was to observe if the student would be hesitant of approaching or touching the robot.

C. Story Task 2

This story is based on mode of transportation and the objective is to make children understand and differentiate between different mediums of travel. Firstly, participants were provided with five pictures of vehicles that are used in one of the books they read. The idea was to understand how well the robot can help them recognize the pictures of the vehicles. To do this, NAO asks them specific questions like: "Can you show me a picture of something we can use to cross a river?" If the participant shows the picture of a boat in front of the robot, he says, it's right and rewards the participant with a dance move. If, however, the participant shows the wrong picture, say a truck, the robot encourages to try again with another picture, until the picture of a boat is shown in front of the robot. The aim of using this method is to help the children learn visually, by attaching the picture of a vehicle in their mind when they hear the name of the vehicle.

Choregraphe's object recognition module was used for this task which provides an easy-to-use interface to train objects. Additionally, simple modules like speech, conditional logic, gestures, and movements were also used to complement the object recognition. The schematic diagram used for the task is shown in Fig. 1.

IV. Evaluation

A. Participants

For the pilot study, four students aged 5, 9, 10, 11 years were chosen as subjects. Although, the final experiment is targeted to students of grade 1 and grade 2, the kids of different ages were chosen so as to evaluate the capability of the robot to recognise and respond to different speech rate, voices and fluency. Among the four participants, only one student speaks English as a first language while others speak Norwegian or Spanish at home. One participant withdrew from the study because she felt uncomfortable approaching or interacting with the robot.

B. Procedure

Participants were told about the objective of the study, and parents were informed about the data collection procedure, how long data will be kept, their right to withdraw, and whom they can contact in case they wish to withdraw. A sheet of paper was provided to the participants explaining the purpose of the study, along with the list of general queries they can make to the robot. For the quiz part, the questions were also provided in a printed format, so that they can refer to it in case they do not understand what the robot is asking. This was done to make the conversation smooth because we had a concern that there might be a situation when children may not understand what the robot is saying, and the dialog may gets stuck in an awkward silence. Also there was a provision of operator intervention if the robot fails to recognize answer from a participant. In that case, an operator could intervene through the terminal with the answer (provided by the participant) so that the participant can experience a smooth transition to the next question.

C. Data Collection and analysis

During the pilot test, two evaluators were employed to note down the cues in a quantitative format. An audio recording of the session was also made with the consent of the participants and their parents for later reference to confirm evaluator's findings. Additionally, an operator was present who gave instructions to the participants before each session and was ready to intervene for smooth execution of different tasks. While the introduction was ongoing, for each participant, the following markers were noted which are analyzed in the results section later:

- What questions did the participants ask?
- Did the robot respond correctly based on the questions asked?
- Did the robot transition smoothly to the next item in the dialog?

It is represented in a tabular form in Table I.

Table II represents the data that was collected during the story task I. Among the four participants that agreed to be a part of the experiment, only 3 participated in this task because one of the students was still in pre-school and was

	No. of participants who asked the question	No. of participants correctly understood by the root	No of participants who interacted clear and loud
Q1	2 1 repeated thrice	2 1 intervention	2
Q2	1- repeated twice	1 intervention	1
Q3	3	3	3
Q4	2 1 repeated twice	1 2 operator intervention	3
Q5	1 repeated twice	1 repeated	1

TABLE I
DATA COLLECTED FROM THE INTRODUCTION SECTION

	NOP who understood the question	NOP who responded in clear loud voice	NOP correctly recognized by Robot	correct response by Robot
Q.1	2 1 referred to paper+help	3	3 although 1 had to repeat the question	3
Q.2A	1 1 referred to paper 1 referred to paper+help	3	3	3
Q.2B	1 1 referred to paper 1 referred to paper+help	3	3	3
Q.3	1 referred to paper 1 referred to paper+help 1 asked to repeat the question+ referred to paper+help	2 1 responded slowly	1 1 operator intervention 1 incorrectly recognized	2 1 incorrect response
Q.4	1 1 referred to paper 1 referred to paper+help	2 1 responded slowly	2 1 incorrectly recognized	2 1 incorrect response

TABLE II
DATA COLLECTED FROM STORY TASK 1

**NOP*:Number of participants

	Picture 1	Picture 2	Picture 3	Picture 4	Picture 5
Participant 1	✓	✓	✓	✓	✓
Participant 2	✓*	✓	✓	✓**	✓
Participant 3	✓	✓	✓**	✓	✓
Participant 4	✓	✓	✓	✓	✓

TABLE III
DATA FROM STORY TASK 2

✓: Recognized correctly
*: needed help in understanding the question
**: took some time since picture was not positioned correctly

not well versed at reading. During story task 1, the following cues were noted:

- How difficult was it for children to understand the questions?
- Did the transition between questions happen smoothly?
- Was intervention required from the operator to make the quiz smooth?

During story task 2, the following events were noted and presented in Table III:

- Did the participant get the question right?
- In how many attempts did the participant get it right?
- How often did the participant have to look at the written script because he/she didn't understand the robot?

V. Lessons Learned

For students to feel at ease with the robot, the introduction session was beneficial. Before the trial, the students were shy and reluctant to carry out conversation with the robot, but the introduction session turned out as a good exercise and children actively participated as the session progressed by asking many questions to the robot as can be seen from the data shown in Table I. The importance of first impression in child-robot interaction was observed to be crucial as can also be seen in [5] to motivate students in long term interaction. Similar to this, in our experiment, the introductory session prepared students to answer confidently during the main quiz sessions based on story task 1 and 2. However, we can not generalize this to every single participant. This was evident from one of the potential participants who decided not to take part in the experiment.

The data in Table II shows that, during the dialog based quiz, children referred to the paper or needed intervention frequently especially when the recognition was based on similar sounding words and short words. Thus, more work is required to address this problem of wrong recognition in the speech engine of NAO. Based on the data given in Table III, the picture based quiz worked fairly well facilitating a smooth communication, however this possibly was because the objects looked vastly different. In scenarios where objects can be similar, this can be tested again.

It was observed that even though the first question of task 2 was longer than other questions, it was understood by participants whereas although question 3 had just two sentences, all students needed help to understand the question as seen in Table II. The possible reason for this could be the use of a new vocabulary in the third question. However, this shows that as long as the words in the questions are known to the kids, the length of the question may not have an impact in recognition. This also shows that the framing of question with the right vocabulary is important in a quiz based robot interaction.

Students were allowed to refer to the paper having questions asked by the robot, and to ask for help if they do not understand the questions asked by the robot. It was no surprise that in many cases children referred to the written form to understand what NAO was asking (as shown in Table II). Automatic speech recognition still does not work reliably with children, and should not be relied upon for autonomous child-robot interaction [12]. Therefore providing the possible queries in a written form turned out to be useful, both for keeping the conversation going, as well as for having a proper framework for data collection during the interaction.

Given that it was the first time some of them were participating in such workshop, we had anticipated the chances of having a few instances where the robot would not understand properly what the participant is saying requiring the participant to repeat multiple times, or in the worst case scenario, operator requiring to intervene and typing the correct response to keep the conversation going. Overall, the number of repeats required (5) and interventions required (3) can be seen as a bit too high considering the simplicity of the the queries that were required to speak (See Table II). But in cases where the children were confident and loud enough, the robot did a good job taking forward the dialog. Developing a reliable system for recognizing children's speech automatically is a challenge because of factors such as the ungrammaticality of children's utterances and rapid developmental changes in the phonetic characteristics of children's speech [11].

Contrary to the results produced by Jan et al. [7], in our experiment, we observed that the youngest participant was more confident in interacting with the robot compared to the older ones. However, our sample was not enough to derive a conclusion because of limited number of participants. Further analysis with enough number of participants would be required to test the validity of the results.

VI. Conclusions, Limitations and Future Work

The results obtained during the pilot test were encouraging. Most of the students showed keen interest in interacting and conversing with the robot although it was their first time interacting with a humanoid robot. The introduction session was found to be really useful practice prior to actual story tasks. Story task consisting of picture based quiz went smoothly compared to story task involving dialogue based quiz.

As such, it is obviously premature to consider the present results as final due to a limited number of participants based on the data collected from a single interaction between the robot and each participant. Moreover, the participants of this pilot test may not be considered as a representative of target age group. In the future, it would be interesting to involve more number of participants to validate the results of the experiment. Nonetheless the results of this study will form a basis for further improvements and refinement in the implementation of speech and object recognition functionality in order to conduct the final workshop with grade 1 and 2 students.

References

[1] *IADIS International Journal on WWW/Internet*, 13(1):72–86. URL: http://www.iadisportal.org/ijwi/papers/2015131106.pdf.

[2] 2018. URL: https://www.softbankrobotics.com/emea/en/nao.

[3] Atheer Alkhalifah, Bashayer Alsalman, Deema Alnuhait, Ohoud Meldah, Sara Aloud, Hend Suliman Al-Khalifa, and Hind M. Al-Otaibi. Using nao humanoid robot in kindergarten: A proposed system. *2015 IEEE 15th International Conference on Advanced Learning Technologies*, pages 166–167, 2015.

[4] Noraidah Blar, Syahril Anuar Idris, Fairul Azni Jafar, and Mahasan Mat Ali. Robot and human teacher. In *2014 International Conference on Computer, Information and Telecommunication Systems (CITS)*, pages 1–3, 2014. doi:10.1109/CITS.2014.6878953.

[5] Natalia Calvo-Barajas, Giulia Perugia, and Ginevra Castellano. Effects of robot's facial expressions on children's perception of trustworthiness in first encounters. 2019.

[6] Mirjam de Haas, Paul Vogt, and Emiel J. Krahmer. The effects of feedback on children's engagement and learning outcomes in robot-assisted second language learning. *Frontiers in Robotics and AI*, 7, 2020.

[7] Jan de Wit, Arold Brandse, Emiel J. Krahmer, and Paul Vogt. Varied human-like gestures for social robots: Investigating the effects on children's engagement and language learning. *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 359–367, 2020.

[8] Amy Eguchi and Hiroyuki Okada. If you give students a social robot? - world robot summit pilot study. *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018.

[9] Goren Gordon, Samuel Spaulding, Jacqueline Kory Westlund, Jin Joo Lee, Luke Plummer, Marayna Martinez, Madhurima Das, and Cynthia Breazeal. Affective personalization of a social robot tutor for children's second language skills. In *AAAI*, 2016.

[10] Peter A. Hancock, Deborah R. Billings, Kristin E. Schaefer, Jessie Y.C. Chen, Ewart de Visser, and Raja Parasuraman. A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 53:517 – 527, 2011.

[11] Junko Kanero, Vasfiye Geckin, Cansu Oranç, Ezgi Mamus, Aylin Küntay, and Tilbe Goksun. Social robots for early language learning: Current evidence and future directions. *Child Development Perspectives*, 12:146–151, 01 2018. doi:10.1111/cdep.12277.

[12] James Kennedy, Séverin Lemaignan, Caroline Montassier, Pauline Lavalade, Bahar Irfan, Fotios Papadopoulos, Emmanuel Senft, and Tony Belpaeme. Child speech recognition in human-robot interaction: Evaluations and recommendations. *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*, pages 82–90, 2017.

[13] A. Kukulska-Hulme. Intelligent assistants in language learning: friends or foes? 2019.

[14] Edgar Lopez-Caudana, Pedro Ponce, Luis Cervera, Sara Iza, and Nancy Mazon. Robotic platform for teaching maths in junior high school. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 12(4):1349–1360, Nov 2018. doi:10.1007/s12008-017-0405-0.

[15] Shu Matsuura, Satoe Kon, and Sakura Kuwano. Communication robot as a weekly online quiz presenter. In *HCI*, 2021.

[16] Deepti Mishra, Karen Parish, Ricardo Gregorio Lugo, and Hao Wang. A framework for using humanoid robots in the school learning environment. *Electronics*, 10(6), 2021. URL: https://www.mdpi.com/2079-9292/10/6/756, doi:10.3390/electronics10060756.

[17] Alex Polishuk and Igor Verner. An elementary science class with a robot teacher. In Wilfried Lepuschitz, Munir Merdan, Gottfried Koppensteiner, Richard Balogh, and David Obdržálek, editors, *Robotics in Education*, pages 263–273, Cham, 2018. Springer International Publishing.

[18] Pedro Ponce, Arturo Molina, Edgar Omar López Caudana, German Baltazar Reyes, and Nancy Mazon Parra. Improving education in developing countries using robotic platforms. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, pages 1–22, 2019.

[19] Natalia Reich-Stiebert and Friederike Eyssel. Learning with educational companion robots? toward attitudes on education robots, predictors of attitudes, and application potentials for education robots. *International Journal of Social Robotics*, 7(5):875–888, Nov 2015. doi:10.1007/s12369-015-0308-9.

[20] Tânia Ribeiro, André Pereira, Amol Deshmukh, Ruth Aylett, and Ana Paiva. I'm the mayor: A robot tutor in enercities-2. 2:1675–1676, 01 2014.

[21] Masahiro Shiomi, Takayuki Kanda, Iris K. Howley, Kotaro Hayashi, and Norihiro Hagita. Can a social robot stimulate science curiosity in classrooms? *International Journal of Social Robotics*, 7:641–652, 2015.

[22] Eleni Vrochidou, Aouatif Najoua, Christodoulos Lytridis, Michail Salonidis, Vassilios Ferelis, and George A. Papakostas. Social robot nao as a self-regulating didactic mediator: a case study of teaching/learning numeracy. *2018 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 1–5, 2018.

[23] Jacqueline Kory Westlund and Cynthia Breazeal. The interplay of robot language level with children's language learning during storytelling. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, 2015.

[24] Jacqueline Kory Westlund, Leah R Dickens, Sooyeon Jeong, Paul L. Harris, David DeSteno, and Cynthia Breazeal. A comparison of children learning new words from robots , tablets , & people. 2015.