World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:3, 2017

Designing and Evaluating Pedagogic Conversational Agents to Teach Children

Silvia Tamayo-Moreno, Diana Pérez-Marín

Abstract**—**In this paper, the possibility of children studying by using an interactive learning technology called Pedagogic Conversational Agent is presented. The main benefit is that the agent is able to adapt the dialogue to each student and to provide automatic feedback. Moreover, according to Math teachers, in many cases students are unable to solve the problems even knowing the procedure to solve them, because they do not understand what they have to do. The hypothesis is that if students are helped to understand what they have to solve, they will be able to do it. Taken that into account, we have started the development of Dr. Roland, an agent to help students understand Math problems following a User-Centered Design methodology. The use of this methodology is proposed, for the first time, to design pedagogic agents to teach any subject from Secondary down to Pre-Primary education. The reason behind proposing a methodology is that while working on this project, we noticed the lack of literature to design and evaluate agents. To cover this gap, we describe how User-Centered Design can be applied, and which usability techniques can be applied to evaluate the agent.

Keywords**—**Pedagogic conversational agent, human-computer interaction, user-centered design, natural language interface.

I. Introduction

CHILDREN enjoy interacting with the computer and, 71% of the children have stated that they would like to talk to their computers [1]. It is our hypothesis that this is possible by devising Pedagogic Conversational Agents (PCAs) that can be defined as a "lifelike autonomous character that cohabite the learning environment creating a rich face-to-face interface with the student" [2]. Several benefits of using PCAs have already been reported in the literature [3]-[5]. However, regarding how the design of these agents, and to our knowledge, no methodologies have been proposed. Each team of authors seems to create the agent according to their expertise and needs, without a unified procedure.

In particular, given the lack of design patterns and recommendations to develop pedagogic agents, we propose the use of HCI methodologies such as User-Centered Design (UCD) [6], [7], whose main goal is to take into account the opinion of the users during the creation of the computer system to cover their real needs, and to be assisted with the users' guidance and comments. In particular, our case study is to design an agent for school children to teach from Secondary down to Pre-Primary education. Moreover, there is some controversy about the usefulness of evaluation techniques that do not take into account users even in the first steps of software development [8], [9]. All in all, some educational systems use them, mostly through heuristic evaluation [10]. This is the reason why in this work, more indagation and test techniques are proposed for evaluating the agent. In particular, in the case study, several evaluation techniques are used to validate the agent.

This paper is organized as follows: Section Related Work presents the context of this work; Section Methodology focuses on the description of how to design and evaluate a Pedagogic Conversational Agent to teach children from Secondary down to Pre-Primary education; Section Case Study focuses on the application of the methodology to design and evaluate the agent Dr. Roland; and finally, Section Conclusions provides researchers with a set of recommendations to develop and evaluate any Pedagogic Conversational Agent to teach children.

S. Tamayo-Moreno & D. Pérez-Marín are with Universidad Rey Juan Carlos, Spain (e-mail: diana.perez@urjc.es).

Project Number: TIN2015-66731-C2-1-R.

II. Related Work

Human-Computer Interaction (HCI) studies how to design, evaluate and implement interactive computer systems to be used by human beings [11]. Among the frequently discussed usability guidelines for interface design, two can be highlighted: "speak the user's language" and "minimize the user memory load". Both of them can be achieved by using pedagogical agents, that is, static or animated anthropomorphic interfaces that talk to the students on a certain topic of study from University down to Primary education.

The use of agents has shown benefits enhancing teaching and learning processes because students tend to perceive the interaction real and natural. Agents are often given life-like characteristics, such as emotion, reactivity, and speech presented in human form, or portrayed within a fantasy realm as talking animals, cartoons, or bugs. Regarding the agent-student dialogue, it is usually focused on the domain under study with a set of exercises provided by the teachers or sentences previously recorded. Moreover, it has been highlighted that for effective learning students should be allowed to produce explanations and the agent should take that information into account and act accordingly. For instance, using AutoTutor [3] in the University, it has been found out that mixed-dialogue interaction, in which both the agent and the student can change the turn of the conversation, can improve the score of the student up to one point (Fig. 1 for a snapshot of AutoTutor); the adaptive intelligent tutorial dialogue module in the BEETLE II [4] pedagogic agent system provided significant learning gains for Primary students interacting with the system (Fig. 2); and, Mike [5]

behaves differently depending on the Primary students' feelings and some pedagogic choices with good results. Up to our knowledge, no agent has still been provided in the literature for Pre-Primary students.

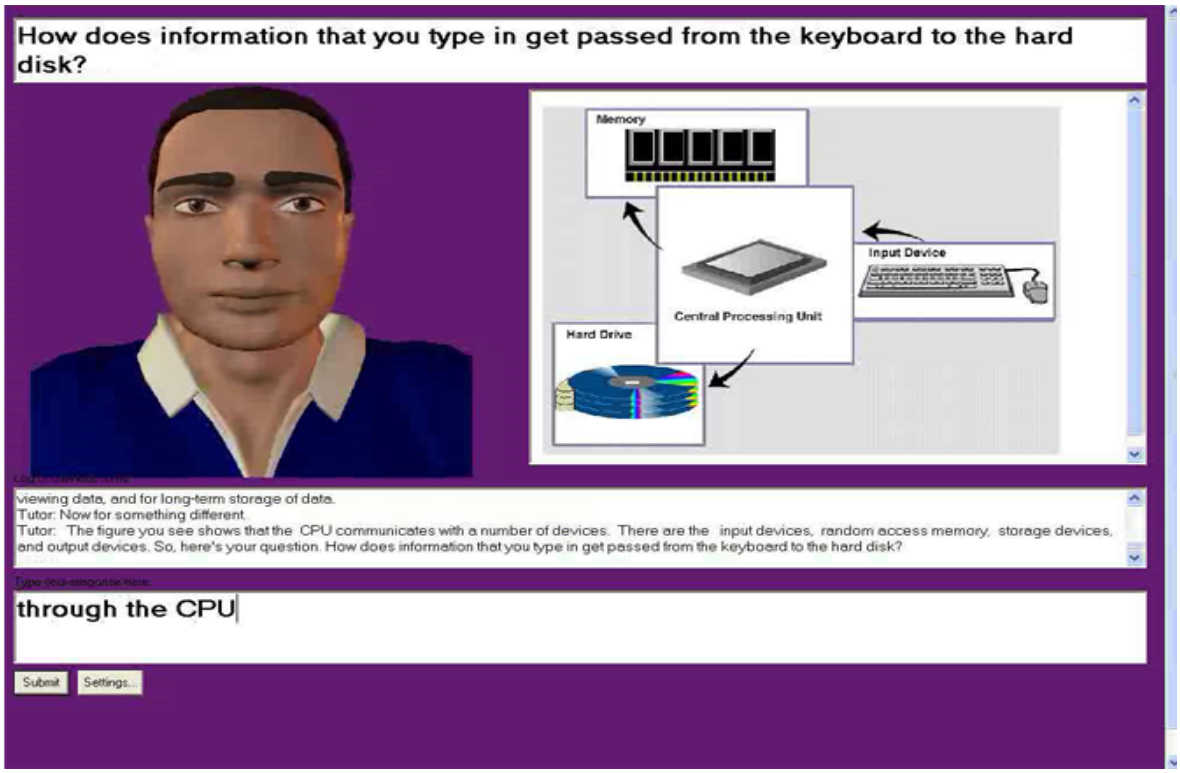

Fig. 1 Snapshot of AutoTutor (sample agent used by University students)

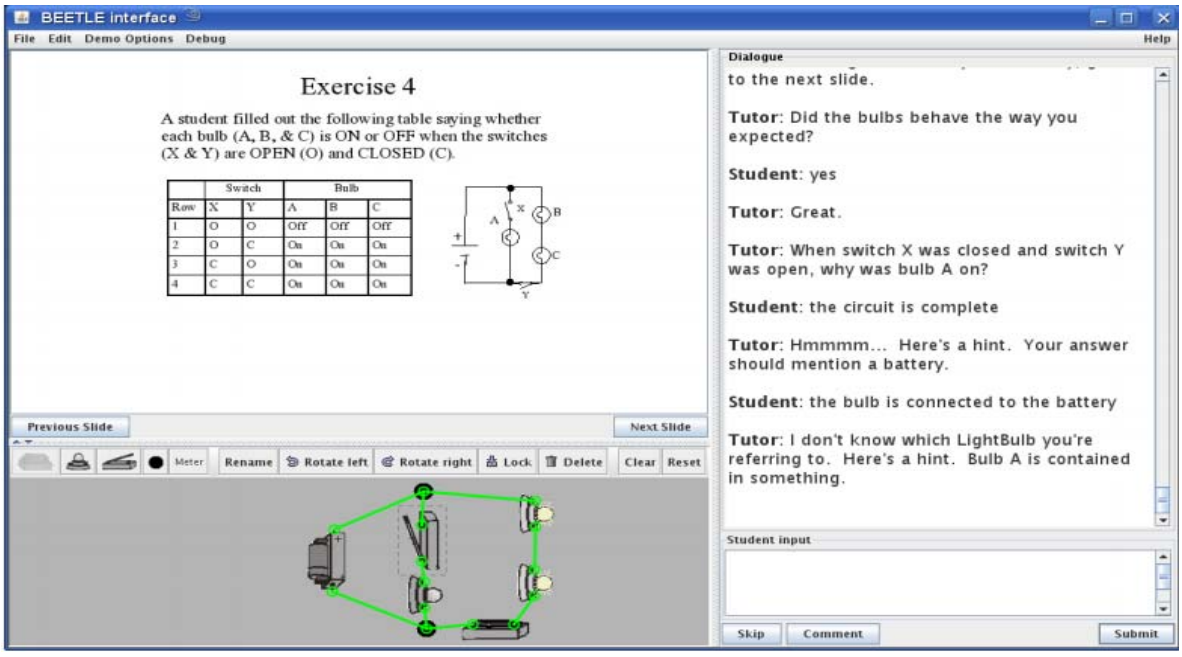

Fig. 2 Snapshot of BEETLE II (sample agent used by Primary students)

III. Methodology

Regarding how to design these agents, and up to our knowledge, no methodologies have been proposed. Each team of authors seems to create the agent according to their expertise and needs, without a unified procedure. Given the lack of design patterns and recommendations to develop pedagogic agents, we propose the use of HCI methodologies such as User-Centered Design [6], [7] whose main goal is to

take into account the opinion of the users during the creation of the computer system to cover their real needs, and to be assisted with the users' guidance and comments.

As can be seen, the first step is to perform the analysis as in any other Computer Science Engineering software development. The difference is that we find out more information about the users, not only requisites but their profiles, needs, roles, etc. Secondly, the design step comes, as in any other software development process, with the difference that in this case users are integrated into the design team. After this step, it is necessary to produce a first prototype so that if there is something wrong, it can be detected before the software is fully implemented (and thus, it is highly expensive to go backwards to make changes).

User-Centered Design is an incremental process. It means that after producing the first prototype it must be evaluated by users, so that their comments can be analyzed again, included in the design and to produce a more complete prototype.

There can be as many analysis-design-prototype-evaluation cycles as users need. Only when the prototype is validated by the users, can it be implemented and tested in a full version of the software. Regarding the evaluation of the systems, one key concept in HCI is usability, which according to the ISO 9241-11 can be defined as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. Usability must be applied during all the software development to improve the quality of the software and reduce costs. Moreover, it is not only necessary to design applying usability principles, but to evaluate the usability of the software according to several techniques.

In the inspection group of techniques, the usability evaluation has to be carried out by experts. It is usually adequate for the first steps of software development. There can be distinguished three main inspection techniques: heuristic evaluation, cognitive walkthrough and standard evaluation. There is some controversy about the usefulness of evaluation techniques that do not take into account users even in the first steps of software development [8]. All in all, some educational systems use them, mostly heuristic evaluation [10]. This is the reason why they are included in this review, and in the context of this work, why albeit they are still used, the results achieved by using only heuristic evaluation is not enough, at least when evaluating the usability of Blended Learning systems [9].

In the indagation group of techniques, the usability evaluation has to be carried out by users that work with the system and they are observed or questioned about it. The users do not need to complete any task required by the expert, but they are working on their own. These techniques are used on the analysis software development phase. There can be distinguished four main inspection techniques: field observation, focus group, interviews/questionnaires, and logs.

In the test group of techniques, the usability evaluation has to be carried out by users that work with the system to complete some tasks requested by the evaluators. These techniques can be used at any software development phase. There can be distinguished by four main inspection techniques: think aloud, constructive interaction, driver test and measurement analysis.

IV. Case Study

From the review of the related work, it seems that using pedagogic agents in learning environments has proven beneficial. However, no methodology to design the agents has still been proposed, and no agents for Pre-Primary students have been proposed. Therefore, we would like to propose the use of UCD to design pedagogic agents to teach children form Secondary down to Pre-Primary education.

We have applied the User-Centered Design (UCD) methodology described in the previous section to develop a Math agent to help students understanding problems and solving them. Moreover, we have evaluated the agent, called Dr. Roland, with 38 Spanish students between 12-13 years old, and three school teachers [12].

We chose to follow a UCD methodology to allow teachers to take part in the development process since the beginning to the end. UCD allowed us to find out the needs of the teachers. There were several meetings to check the requisites, and whether the teachers had identified any new requisite. Moreover, we also wanted teachers to define the features that the agent must have according to their knowledge of the students. Finally, five meetings to design the agent for the Secondary Education students were established: November 7th, 2011; February 10th, 2012; March 22nd, 2012; April 24th, 2012 and May 3rd, 2012. These meetings will be briefly described in the following paragraphs as a sample to follow the four main steps of UCD: analysis, design, prototype and evaluation. As the description is linear, but the methodology is not, it will be indicated the number of the iteration.

A. Analysis

In the first meeting (November 7th, 2011), three school teachers explained to us the need for some kind of help for children unable to solve Math problems. We told them about the Pedagogic Conversational Agents, and started talking about how the agent should address the need detected.

B. Design

Given the complexity of the requirements, it took us some months to design an agent that met the requisites given. During this time, teachers were consulted whenever there were doubts, and teachers worked as members of the design team.

C. Evaluation

First Prototype Evaluation

In the second meeting, teachers were shown the teacher interface prototype. They validate the interface, and the development of the interface for the children started. We talked about the image of the agent, and the teachers told us that it should be something formal like a graduate.

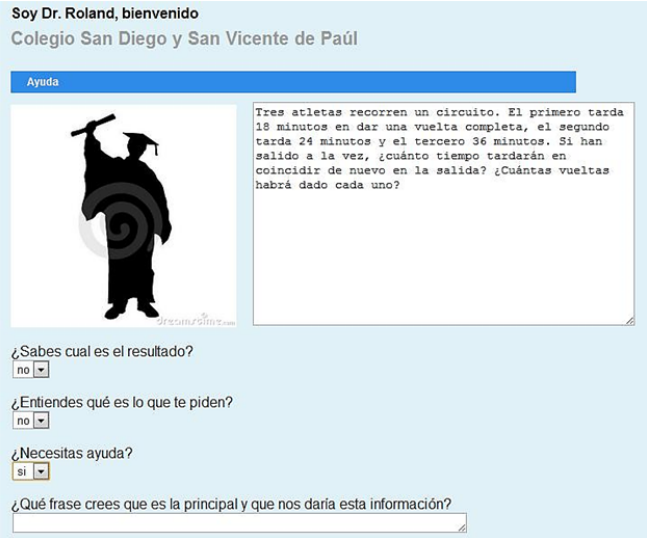

Fig. 3 Interface of Dr. Roland for Secondary Education students

Second Prototype Evaluation

In the third meeting (March 22nd, 2012), the interface for the children was also validated. The teachers only requested the addition of a help button. It took some months to implement the help button and the algorithm. Each problem had been classified according to several types and levels of difficulty. According to the level of difficulty and the amount of help received, the score given to each student varies. The types of exercises are shown in cycles, so that when a cycle of one type is finished, it calculates the score achieved, and that is compared with the score registered in the previous cycle. According to the result of the comparison, the student will be presented with more difficult or easier problems in the next cycle, so that the level of difficulty of the problems is never too hard or too easy.

All interaction data are also registered, so that teachers can check that information later. Teachers can also insert new problems, modify the information of already existing problems, and introduce new levels of help. That way, it was thought that a higher level of flexibility is provided to the interaction between students and Dr. Roland.

Third Prototype Evaluation

In the fourth meeting (April 24th, 2012), 38 students were requested to fill in a pre-test with eight problems of four different levels of difficulty, and a questionnaire, whose results revealed that nearly all the students have a computer at home with Internet connection, and 100% have an e-mail account. All students considered positive the use of the computer to help them in their studies.

Fig. 3 shows the students' interface as it was shown that day to 20 students (test group). As can be seen, the picture of the agent is on the left, with the problem on the right, and below are the questions.

All the students were excited at the idea of having a computer program able to help them solve the Math problems. In 10 minutes, they were able to complete two exercises in average, passing an average of 0.6 problems. Some comments gathered include the following: "I like that the system asks me whenever it is necessary", "I like that Dr. Roland explains the problem to me with detail", "I like to solve the problem step by step", or "I would like that Dr. Roland gives me more samples".

Fourth Prototype Evaluation

Finally, in the fifth meeting (May 3rd, 2012), 19 out of the 20 students of the test group used Dr. Roland (this time for 50 minutes), and 19 students were randomly assigned to a control group (no use of Dr. Roland), keeping the proportion of girls-boys in both groups. Test students completed an average of 6.2 problems with Dr. Roland, passing 3.2 problems (21.6% of increase with respect to the previous use). The control group students were given a paper with the same Math problems.

Finally, students were asked to pass a post-test with the same number of problems (8) and levels of difficulty (4) than the pre-test. No significant difference was found between the control and test groups. We believe that the reason of not finding any significant evidence was the limited exposure to Dr. Roland to produce any significant changes in the test group. Moreover, there were some issues with data input in the agent, and students requested some tools such as a calculator, blackboard, and improvements that are being implemented for allowing the students to use the agent during several lessons together with new features to involve parents to motivate their children to use Dr. Roland after class at home.

All in all, the qualitative study regarding the reaction of the students revealed that despite the short time in which students interacted with Dr. Roland, the levels of motivation and concentration were higher in the test group. When reviewing the state of the art, it was not only detected a lack of recommendations about how to design Pedagogic Conversational Agent for children, but also a lack of metrics to evaluate them. As can be seen from the previously described experience, two metrics were evaluated: the level of motivation (from a qualitative point of view with field observation and questionnaires) and learning gain (with a pre-post test). HCI also provides researchers with several usability evaluation techniques that, in our opinion, should be taken into account when using Pedagogic Conversational Agents. The reason why we consider that it is necessary to evaluate the usability of agents is that using the agent in class it not enough to increase the learning gain, as seen in the previous section, and in order to use the agent at home it should be friendly and easy to use by both children and parents.

The use of inspection techniques was discarded, because researchers and developers may find difficult to think as children given the age difference. On the other hand, constructive interaction is considered essential as one of the first check of the agent with the teachers, students and parents, to avoid input errors, incoherence, and motivate a better and more extensive use of the agent both at home and in class.

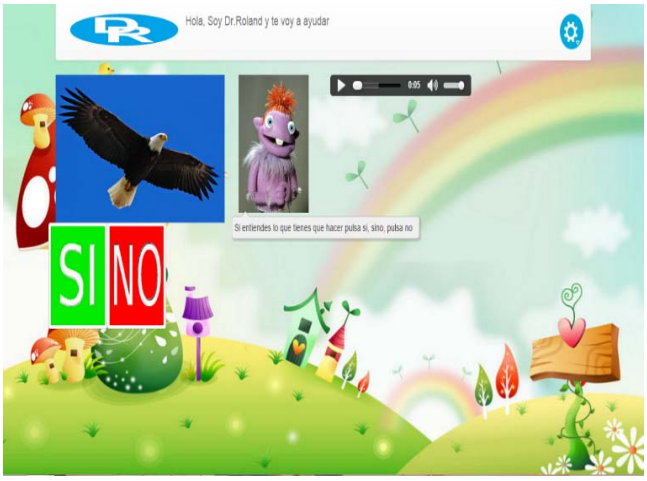

Fig. 4 Interface of Dr. Roland for Pre-Primary Education students

Adapting the Agent to Primary and Pre-Primary Education

In the following courses, Primary and Pre-Primary Education teachers were asked to follow UCD to design an agent for their students too. It took several analysis-design-prototype-evaluation iteration cycles (similar as in the detailed sample provided) until in the 2015/2016 course [12], Dr. Roland was used for the first time by 10 year-old children in Primary Education, and 2-3 years-old children in Pre-Primary Education, as can be seen in Fig. 4.

The evaluation in this case was based on a combination of field observation and interviews. In particular, in Pre-Primary Education, teachers advised us to ask children during the assembly at the beginning of their class when they are more receptive and attentive. The results in both cases were satisfactory, i.e. according to teachers' expectations and allowing a more fun and efficient learning process for children as they were able to remember what they have been taught.

V. Conclusion

User-Centered Design has been key to produce an agent that has successfully been used both by teachers and students from Secondary Education down to Pre-Primary Education.

Some recommendations drawn from this experience, and generalized for any researcher or designer that would like to start developing a Pedagogic Conversational Agent would be the following:

- It is essential a good student-agent algorithm is adapted to the domain of the study. The algorithm should be devised with the help of teachers and parents, and validated before trying the agent with the students.
- The interface should keep the HCI familiarity principle to software liked by the students. For instance, Dr. Roland for Secondary Education students follows the Messenger interface because it was a popular program with those children.
- The use of colors should be personalized to allow students control over the interface design.
- The use of voice or animation is not essential for all domains and teachers. In particular, Secondary Education teachers would prefer not to have animation or voice as they feel it could distract the students. On the other hand, Pre-Primary Education teachers prefer to have lots of colors in the interface.
- The collaboration possibilities should be taken into account, as students tended to work in groups. Moreover, tools such as calculators, shared blackboards and current technology available in educational software should be integrated in the agent interface.
- Given that children are creative when typing their answers, input to the agent should be carefully designed, even knowing the children are highly tolerant and they will try to make the agent understand what they are telling several times.
- Agents increase the motivation of children to study (even Math), but they should be given enough time to practice with the agent to get some kind of learning gain. Help from parents at home is needed to encourage children to study after class.

As a future work, the researchers of this system would like to continue working on the agent to develop the next prototype to include the lessons learned.

References

[1] Narayanan, S., & Potamianos, A. (2002). Creating Conversational Interfaces for Children, IEEE Transactions On Speech And Audio Processing. Vol. 10, no 2, 65-78.
[2] Johnson, W., Rickel, J., and Lester, J. (2000). Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments, Journal of Artificial Intelligence in Education 11, 47-78.
[3] Graesser, A., D'Mello, S., Craig, S., Witherspoon, A., Sullins, J., McDaniel, B., & Gholson, B. (2008). The relationship between affective states and dialog patterns during interactions with AutoTutor, Journal of Interactive Learning Research, 19 (2), 293-302.
[4] Dzikovska, M.O., Isard, A., & Bell, P. (2011). Adaptive Intelligent Tutorial Dialogue in the BEETLE II System, Proceedings of the Artificial Intelligence in Education Conference. Vol. 6738 LNCS, 621.
[5] Lane, H.C., Noren, D., Auerbach, D., Birch, M., & Swartout, W. (2011). Intelligent Tutoring Goes to the Museum in the Big City: A Pedagogical Agent for Informal Science Education, Proceedings of the Artificial Intelligence in Education Conference. Vol. 7926 LNCS, 309-318.
[6] Abascal, J., Aedo, I., Cañas, J., Gea, M., Gil, A.B., Lorés, J,; Martínez, A.B., Ortega, M., Valero, P., Vélez, M. (2001). La Interacción Persona-Ordenador, Jesús Lorés, Asociación Interacción Persona Ordenador (AIPO).
[7] ISO 9241-210. (2010). Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems. ISO.org
[8] Sauro, J. (2012). http://www.measuringusability.com/blog/effective-he.php. Last visit: January 25th, 2017.
[9] Pérez-Marín, D. & Salmerón, L. (2014). When Human-Computer Interaction meets Blended Learning and Heuristics are not Enough: Evaluating the usability with Field Observation and Interviews. In Press, Nova Publisihing.
[10] Alsumait, A. & Al-Osaimi, A. (2009). Usability Heuristics Evaluation for Child E-learning Applications. 11th International Conference on Information Integration and Web-based applications & Services, pp. 425-430.
[11] Dix, A., Finlay, J., Abowd, G., & Beale, R. (2004). Human-Computer Interaction, Prentice Hall.
[12] Tamayo-Moreno, S. & Pérez-Marín, D. Adapting the design and the use methodology of a Pedagogical Conversational Agent of Secondary Education to Childhood Education. In Press, IEEE 2016 International Symposium on Computers in Education (SIIE).

Silvia Tamayo Moreno is a Computer Engineer and has a degree in Business Administration and Management since 2011 by the U. Rey Juan Carlos and a Master's in Advanced Software and hardware Systems Research since 2012. She was awarded the prize for Best Master's Thesis by E-Madrid Network in 2013 for the work " Propuesta de Desarrollo Centrado en el Usuario de un Agente Conversacional Pedagógico para la Compresión Lectora de Ejercicios de Matemáticas a Nivel Escolar". She is currently doing a PhD, and is passionate about technology, research and innovation and works as an engineer big data.

Diana Pérez-Marín received the Ph.D. degree in Computer Science and Engineering from the Universidad Autónoma de Madrid in 2007, and her degree in Computer Science by the Universidad Autónoma de Madrid in 2002. She is currently a lecturer and researcher at the Department of Computing Languages and Systems I of the Computer Science Faculty of the Universidad Rey Juan Carlos in Spain. Her PhD has been published in 2010 by VDM Verlag and it is available with ISBN 978-3639207200. She has published more than 40 papers in international journals such as Journal of Educational Technology and Society, or Computers and Education, and participated in eight national and international research projects. She has been awarded a prize to their novelty of the teaching possibilities of Internet in 2007 and as project leader in 2008. In 2010, she organized the First International Workshop on Adaptation and Personalization in E-B/Learning using Pedagogic Conversational Agents (APLeC) together with Dr. Ismael Pascual-Nieto and Dr. Susan Bull, collocated with the international conference User Modeling, Adaptation and Personalization (UMAP). Currently, she is also co-editor of a Special Issue on Automatic Free-text Evaluation in the International Journal of Continuing Engineering Education and Life-Long Learning (IJCEELL) to be published by Indescience at the end of 2010. She is member of the Artificial Intelligence in Education (AI-ED) Society, LEMORE, LITE, and ADIE.