The design of

a storytelling and learning by teaching activity with a social robot as teachable main character, for primary school language learning

Created by: Gijs Verhoeven
Created for: Creative Technology graduation project | University of Twente
Supervised by: Dr. Mariët Theune & Dr. Alejandro Catala
Handed in on: 06-07-2018

Abstract
This research is about the design of an activity for primary school second language learning. The activity is a combination of storytelling and learning by teaching, with a social robot as teachable main character. Both storytelling and learning by teaching with a social robot have proven to be beneficial for primary education. However, the combination has not been researched yet. This could have unknown benefits by combining the beneficial aspects of both approaches into one activity. First a concept was developed based on literature research and (game) design principles. In an iterative process of rapid prototyping, user testing, and design improvement, the concept was further developed. Then a high-fi prototype was built which was tested on a primary school. The twenty two children had learned two words on average. They also had fun and interacted with the system as desired. However the reason for specific learning behaviour could not convincingly be explained in terms of the design choices made.

<u>Table of contents</u>

Remerciement
Yes, that is a French word. First of all I would like to thank my two supervisors. They supported me very well during this project. Communication with them was very clear. On the one hand, they gave me the freedom to design what I wanted. On the other hand, they gave me useful feedback and tips when that was needed. Secondly, I would like to thank my uncle. Together we solved some problems I was facing during the project. Thirdly, I would like to thank the voice actor. Because of her willingness to help, the robot now has the perfect voice that suits its role. Fourthly, I would like to thank the UT daycare and the children there. I was welcome to do two user tests with a paper prototype. Thanks to this collaboration, the design has made big steps forward. Lastly, I would like to thank the primary school where I have done my final user test. Communication with them went very smoothly and they helped me with getting the parent's consent for example. The children participated with pleasure and listened well to me.

1. Introduction

Robots are slowly but surely integrating into our society. The number of applications increases as technological development advances. One possible application that has caught researchers' attention is primary education. Educational robots are a subset of educational technology, where they are used to facilitate learning and improve educational performance of students. Robots provide an embodiment and the ability to add social interaction to the learning context and thus an advancement on purely software-based learning, say Mubin et al. [1]. For this graduation project, a coBOTnity surface-bot is used, which can be seen on image 1. The robot consists of two parts: a base and a tablet. The base is a round block, about 20 cm diameter, on wheels. The tablet forms the face and "brain" of the robot. The robot has some functions: Its position on the field can be tracked by a camera, it can recognize lines on the field, its face can display emotions and say some words, tangible objects can be recognized via a QR-like code and it can "carry" objects by displaying them on the screen. Although not highly intelligent or advanced, the robot can be seen as a social robot. Additional functionalities can be programmed as part of the graduation project.

Image 1 coBOTnity surface-bot on a playground

One way to integrate robots into primary education is via storytelling. According to the *National Storytelling Network, s*torytelling is defined as "the interactive art of using words and actions to reveal the elements and images of a story while encouraging the listener's imagination". In [15], Yamaç and Ulusoy show the value of storytelling in children's development of skills like linguistic skills, communication, logical thinking and creativity. Robots have successfully been used by researchers in storytelling activities. For example in [5] Krzywinski and Chen designed a collaborative, tabletop storytelling activity with a robot as main character and successfully enhance collaboration among students in the form of telling a story together, passing tangibles and evoking discussions about the plot. In such an activity, the robot can enhance the story by displaying emotions and behaviors based on the story, such as is done by Druin et al. in [6]

Another way to integrate robots in education is via learning by teaching. This means that someone learns by putting in effort to teach someone else. There is rich evidence for the effectiveness of this pedagogical approach, both in terms of learning outcomes and motivational effects, say Lindberg et al. [3]. In this approach, technology can be implemented as teachable agent (TA). This means that some technology plays the role of agent that needs to be taught by the child. In human-human interaction, this would be another person. Important to note is that initial research on learning by teaching with a TA was done with virtual TAs instead of robots. According to a research done by Chase et al., students from 10 to 14 years old seem to take responsibility for their (virtual) teachable agents and spend more time on their learning activities this way than when they learn for themselves [4]. 10 to 11 year olds even treat their TAs as social entities by attributing mental states to them. In recent years there is growing interest in having a robot as TA instead of a virtual agent, given that this possibility becomes more feasible as technological progress advances. In [7], Knapp et al. (as cited by Werfel) discuss how robots, being physical entities, might have a greater impact on children than virtual TA's by using their non-verbal communication capabilities, like posture, movement and pointing. This can help in the acceptance of the robot by the children and increase engagement and focus. Moreover, a robot may be more suitable for physical tasks like learning hand-writing than a virtual TA. The exact advantages of a robot TA instead of a virtual TA are not yet clear however.

Besides how to implement them, another variable within the application of robots in education is the subject that is taught. While many could be chosen, certain subjects, like language learning, are more popular amongst researchers. Kory and Breazeal in [8] argue that children's language development is not just about exposure to words, it is also about the dialogic context, about communicating meaning and having a social interaction that happens to use words to communicate. Such a dialogic context would require interactivity, social cues and the shared roles of speaker and listener. A storytelling activity can be a reason for such a dialogic context to arise. Besides this, they argue that robots can be beneficial as an intuitive physical interface for interaction, using easy interpretable social cues like speech, movement and gaze. Of course the type of social cues used is dependent on the type of robot used and its technical capabilities. According to Chang et al. in [9], children are not as hesitant to speak to robots in a foreign language as they are when talking to a human instructor. In addition, students can practice with the robot as often as they want. The robot will not get bored or run out of time, unlike a human instructor.

This graduation project is about educating the future generation and is part of the ongoing research on how educational robots can be integrated to enhance learning activities. Previous research has shown that robots can successfully be integrated in primary school education in the context of a storytelling activity or a learning by teaching activity. It has also been proven that children can benefit from a robot while learning a language. However, a combination of these approaches has not been researched yet. Therefore in this research a new approach is taken by combining above-mentioned, proven to be beneficial, approaches into one learning activity. This combination could lead to further improvements in primary education by emphasizing the beneficial aspects of each individual approach and combining those into one learning activity, leading to a possibly better whole.
In other words, a robot will be used as teachable agent for learning by teaching and main character for storytelling at the same time, with learning a language as subject. French is chosen as language mainly because of the researcher's affinity and recent experience with it. More formally this research's main question is:

Main question: *How to design a storytelling and learning by teaching activity with a social robot as main character and as teachable agent, for primary school children to learn French as second or third language?*

Note that here the choice was made to make the robot the main character. Like this, the robot could be seen as tool for the child in its play. Another option would be as learning companion. As tool, it would probably need less complex social features than as learning companion because it would more passively follow the child's play. In other words, there is less complex active input needed from the robot. As tool, technical requirements for the robot would be less complex. In addition, the role as main character sets the focus of the activity on the robot. Obviously, this is desired, to focus the child on the storytelling and learning by teaching that the robot supports via its social behaviour.

This subject leads to several sub-questions about the different aspects of the activity that need to be designed.

Sub-question 1: *How to design the storytelling part of the activity?*

This question is about the setting of the story, other characters next to the robot in the story, the balance between given player freedom and narrative structure [2] and all other necessary design choices regarding this aspect. It also addresses the choice for tabletop or another kind of setup.

Sub-question 2: *How to design the learning by teaching aspect in the activity?*

This question focusses on all important aspects of the implementation of learning by teaching. For example how the teaching students will get their knowledge about the French, given that they have no or little previous experience with this language, and given the possible time constraints when testing the activity in practice. In addition, how this knowledge can best be transferred from teacher to the teachable agent, given the designed environment and robot specifications. An option here is to use tangibles representing words. Also how many students best can be involved in one run of the activity and what roles they should take on within the activity in relation to the robot. For example they can all be teachers and work together, it can be a solo activity or one student can have the role of controlling the robot while another student teaches.

Sub-question 3: *What are the required technical capabilities of the robot to fit its role as main character as well as teachable agent, and how to implement these into the system?*

This question addresses the technical aspect of the activity, which is centered on the coBOTnity surface-bot. An important factor for design decisions, in sub-question one and two, is if it can timely be implemented in the surface-bot. In this sub-question, all these design choices will be evaluated and translated to technical requirements for the robot. Then, these requirements will be realized by adjusting existing functionalities and adding, if necessary, extra functionalities.

Sub-question 4: *How to set up, execute and evaluate the experiment?*

This question includes practical matters like the role of the teacher, the introduction of the robot to the children, contacting a primary school, research ethics, informing the parents and the place of the experiment. In addition, this question will address how the evaluation of the experiment will take place. For example, learning by teaching might be evaluated by testing the children’s French vocabulary before and after the activity. If allowed by the teacher and the parents, additional questions could be asked to the children, and maybe to the teacher(s) as well, to evaluate aspects of the activity.

Sub-question 5: *What influence do age and gender of the children have on the activity?*

According to ter Stal [2], children between six and eight have enough skills to tell a structured story and also still use toys in their story. This age range seems ideal for storytelling with a (toy) robot. However when adding learning by teaching and a new language, it may be that children of older age, like ten to twelve, are more suitable for the activity as a whole. These children may be more capable of taking the responsibility of being the teacher and may also be more ready for learning a new language, given that they have already experience with learning English as second language. In addition, the influence of gender on the activity will be researched. Maybe girls of this young age are better capable of taking the responsibilities of being a teacher than boys. Maybe boys are more creative with combining learning by teaching in the storytelling activity. There could also be no significant differences.

The report follows a main chapter structure of ideation, specification, realization, testing, results, conclusion and discussion. The sub-questions will be addressed and answered throughout this structure in the report. Answers will be found in different ways. The first way is via literature research. Existing work will be investigated and recommendations and success and failure factors of these works will be taken into account. In other words, literature research will form a solid base of knowledge on the subject and a source of advice for the experiment. The second way is by designing something, trying it out on other people, preferably children, and evaluating and improving it and repeating the process. Then in the conclusion all these sub-answers will be summarized and used to answer the main question. Finally, in the discussion, the whole research will be discussed and ideas for further research will be given.

2. State of the art

This research addresses several topics and combines them. These topics are social robots, storytelling, learning by teaching, primary school education and language learning. In the past there has already been done much research on the combination of several of these topics, pointing out positive and negative sides of these specific setups. A state of the art research on these works will provide a solid base of knowledge that can be used to make well-founded decisions during the design and execution of this research's experiment.

2.1 A robot in a storytelling activity for primary school children

Image 2 a child does an activity on the tablet with the DragonBot

The combination of storytelling with a robot for primary school language learning has been researched before, for example by Kory and Breazeal in [8]. They adapted the robot's language level to the children's level in a native (L1) learning activity and found that strategically matching the robot's language skills improves learning outcomes for the child. This strategy makes use of scaffolding as underlying theory. This means that someone learns by introducing him/her to knowledge that builds forth on what he/she already knows about it [10]. In this experiment's case, the robot was slightly better than the child, making use of more advanced words and story structures, which positively influenced the child's skills, who adapted him/herself to the peer. This setup used the robot as learning companion however (image 2), not as main character in the story. Nevertheless, some useful conclusions were drawn. For example the importance of the robot's language level, based on the children's level. Moreover this research shows the benefit of a social and interactive environment, created via storytelling, for language learning.

Image 3 Two DragonBots

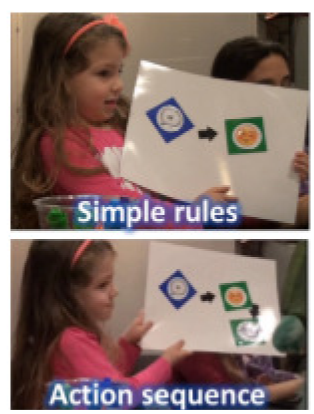


Image 4 child with tangible stickers

The robot used in this research was the DragonBot (image 3). This fluffy toy-like dragon robot was designed to interact with children. Because of this reason, it is regularly being used in experiments like this. Another example where it is used is in [11] by Gordon et al. They use the robot as tool that needs to be taught how to socially interact. The children create a concatenation of tangible stickers displaying commands that look like programming concepts, as can be seen in image 4. This research concludes that the use of tangibles is intuitive and appealing for children. The successful use of tangibles in a storytelling activity with a robot is also shown in [5]. Here Krzywinski and Chen research social skills learning with a tabletop storytelling activity. Their small robot on wheels, image 5, can be pushed and pulled by the children with tangible tiles. They can also draw additional characters and weather effects in the form of tangibles. This installation worked very well to enable collaboration amongst the children in the form of passing tangibles, discussing the plot and helping each other. The tabletop setup eased this process and the tangibles were intuitive. However they also addressed an important negative point on the use of technology in the activity, which is technological constraints. During the experiment, the robot often did not exactly go where the children wanted and they needed to adjust it as a result. This made the interaction more complicated and frustrating. In addition, the children wanted to express more behavior and emotions with the robot than it could. In other words, the robot was not advanced enough to fluently support the narrative. This is an important point: bad ease of use, as concluded, can drastically reduce the fun and engagement of the children and therefore possibly also reduce educational benefit.

Image 5 Krzywinksi and Chen's robot

This was also concluded by ter Stal in [2]. She made an interactive storytelling activity with a robot as main character and focused on the effect of emotional behavior of the robot on the child's story. She found that giving the robot emotions results in more creativity and more causal reasoning within the child's story, instead of just a follow-up of actions. Also the story included a larger variety of emotions and story summaries were more often described from a third person view, suggesting that the robot is seen as separate entity. However, as just said, she also pointed out the many technological challenges that remain, like more autonomy for the robot, support of multiple themes, typing input reduction and the use of multiple robots that interact. These factors could in the first place allow children to come up with more complex and varied stories. In the second place, by letting the robot decide things on its own, the children can be steered towards a certain scenario or effect, which may stimulate their creativity and storytelling skills to fit the robot's behavior into the existing plot.

A possible solution to technological constraints is the use of tele-operation. This means that the robot is controlled by a person from a distance and it reduces the need for advanced and/or autonomous technology to allow a smooth interaction. The child is not aware of this, unless the effect of telling it is part of the experiment. DragonBot in [8] and [11] is successfully controlled like this. In [12], Kory and Breazeal take a closer look at the effect of teleoperation on the experiment and suggest that it does not matter for the interaction if the robot is teleoperated or autonomous. One could say that it is better to use teleoperation to make a decent activity, then to let deficient technology limit it. Of course this is only true if the experiment focusses on the interaction itself and uses the technology merely as a means.

2.2 Second language learning for primary school children

Research on second language learning (also called L2 learning) for children specifically has also been done. For example in [13] by Tonzar et al. There the effect of so-called word-word learning versus image-word learning is investigated. Word-word learning means learning L2 words by showing the L1 translation. Image-word learning means learning L2 words by showing an image of the meaning. In this context Potter et al. (as cited by Tonzar et al. in [13]) introduce two different models of word processing. The first is the word association model which means that L2 words gain access to concepts via L1 words. The second is the concept mediation model which means that the L2 words themselves are linked to the concepts. The second model is considered to be better because it allows for a broader use and recognition of the L2 words. After comparing and testing word-word versus image-word learning, Tonzar et al conclude that image-word L2 language learning is more effective. It appears that this way, learned words stay longer in the memory because the words are coupled to the concept, instead of to the L1 word. In other words, image-word language learning goes hand in hand with the concept mediation model of word processing. Besides this it was confirmed that cognate words are easier learned than non-cognate words both during word-word learning and image-word learning. An L2 cognate word is a word which is similar to the L1 word, in meaning and way of writing and/or pronunciation. For example in English-French: Letter and lettre or adorable and adorable. Non-cognate words are apple and pomme.

A popular research topic in L2 learning for primary school children is the implementation of a social robot as tutor. An example is the L2tor (pronounced 'el tutor') research project which is funded by the European Commission. The project aims to design a child-friendly tutor robot that can support L2 learning by socially interacting with the children. The research project has led to some interesting insights on language learning with a robot, even though the robot is used as tutor instead of as teachable main character. According to their insights Vogt et al. present guidelines for designing a social robot in this context in [16]. First they address age differences. They found that children on average have to be at least 5 years old in order to be engaged and not be scared of a less cute robot like NAO, the robot they used, image 6. Second, they address target word selection, meaning what and how many words to teach the child per activity. They state that words should be taught in groups of the same subject (e.g. supermarket products or animals in the zoo) and within their context. This means that supermarket products should be taught in an imaginary visit to the supermarket. In addition, they show the importance of repetition to learn an L2 word and its pronunciation. For a five-year-old, to learn six new words in a session of twenty minutes, each word needs to be repeated about ten times.

Image 6 NAO interacts with a child

Third, regarding the child's engagement, they found that the content of the feedback does not matter, whether it is positive and constructive, or only negative. On the other hand, they found that the difficulty level is very important as either a too easy or too difficult exercise will drastically reduce children's engagement. Finally, they mention the importance of introducing the robot to the children in the class, before the experiment sessions. Children who are a bit afraid of the robot could get more confident when seeing their peer's confidence. They also state how introducing the robot as peer could reduce anxiety and make sub-optimal interactions due to technical constraints more accepted by the child. Moreover, they state that introducing the robot as peer, who needs to learn the language, could lead to learning by teaching.

2.3 A robot as teachable agent in a learning by teaching activity

While previous research has proven that robots as well as learning by teaching have many beneficial effects in a wide variety of applications [7], the effectiveness of robots used as teachable agent in learning by teaching is a less researched topic. In [7], Werfel discusses this topic elaborately. He points out possible advantages of a robot as teachable agent in the future, when technology and research has further advanced. For example the robot could gather data and analyze the teaching behavior of the student, from which it could identify topics the student covers less elaborately. This can mean that the student does not know the specific topic well yet and the robot could anticipate by asking questions about the topic. Continuing on this, the data collected this way could in the long term be used to identify different teaching patterns. Then the robot could specify its learning behavior by adjusting it to the specific teaching behavior of the student. In other words, it is the challenge to design the robot in such a way that it best helps learning for the student.

Learning by teaching was achieved by Hood et al. in [20]. They used the NAO robot as teachable agent to improve children's handwriting, as can be seen on image 7. The robot wrote wrongly on purpose in the beginning. It improved after the children showed him how to do it. The children taught him until they were satisfied with the robot's performance. Especially technical challenges were overcome during this research. For example making the NAO robot capable of handwriting, given its limited fine motor capabilities.

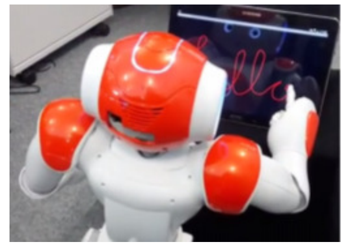


Image 7 writing NAO robot

The discussion among researchers in the field of teachable agents, however, is whether a robot as TA is really better than a virtual TA. The above-mentioned advantages of an artificial TA could also be achieved with a virtual one. Lindberg et al. investigate this in [3]. They let one group of children teach mathematics to a robot, and another group to a virtual TA. They conclude that both options work, image 8 and 9. However, in the case of the robot, much can still be improved on the technological side. For example learning a physical act, like writing, requires a very specific mechanical robotic setup. On the one hand robots are better as TA for physical activities, on the other hand, they are often still constrained by technology. Mentioned in [7] is the possible solution to use a robot head with a virtual body. This way there is still a physical and interactive entity, but it is not constrained in its body movements. Lindberg et al. mention how the robot definitely elicited more attention and enthusiasm amongst the students, compared to the virtual TA. However this enthusiasm decreased as the activity progressed.

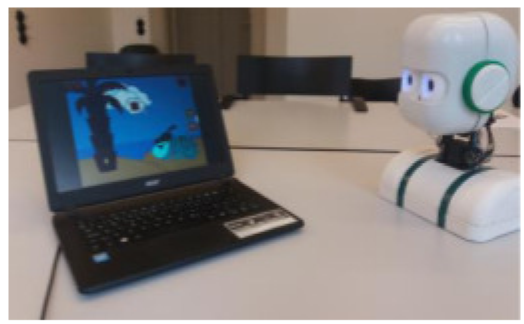


Image 8 Lindberg et al.'s robot TA

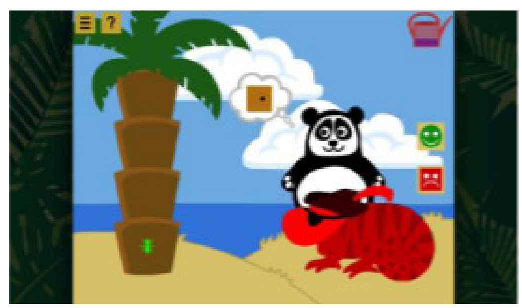


Image 9 Lindberg et al.'s virtual TA

The increased enthusiasm for the robot compared to the virtual TA was also mentioned in [14] by Rosenthal et al. They warn that this enthusiasm may take away the attention for the actual tasks that need to be done. They tested the difference in linguistic alignment and perception of the system for a robot TA, virtual TA and recorded speech in an L2 learning activity. Linguistic alignment is the effect that one person adjusts its language usage depending on the person he or she is talking with. For example with accents and level of complexity. The researchers found no difference in these aspects for the different TA's. In addition, they did not find a difference between computer-spoken text samples and human voice recorded samples. Besides this, they suggest that a virtual TA may be better to focus on conversations, thus language learning, but that a robot triggers more interactional features of using a language, like calling someone by their name. Finally the research concludes that an advanced social robot with human-like voice is not necessarily needed in language learning by teaching.

2.4 Conclusion

To conclude, previous research has drawn both positive and negative conclusions about storytelling and learning by teaching activities with robots, and about language learning. For example it was found that giving the robot emotions results in more complex and better structured stories and a tabletop storytelling setup is convenient for collaboration among the children. The use of tangibles in an activity was also evaluated positively by researchers. Tangibles were intuitive and appealing for children and seem a good design decision. They could also be combined in some way with the image-word learning method, which was proven to be more effective than the word-word method for learning a second language. Continuing on L2 learning, research pointed out that words should be taught repeatedly and in coherent packages within their context. The form of robot feedback does not seem to matter for the children's engagement but the exercise difficulty does. It should not be too hard or too easy. Additionally it is important to first introduce the robot to the children in the class and to frame it as a peer. The latter could evoke learning by teaching. An important point of attention is the technological constraints. As long as technology is not yet advanced enough, activities could either be teleoperated or simplified to fit the technological possibilities. Anyway it is important to not let them negatively influence the child's experience with the activity. Especially robots as TA could still benefit largely from further advancements in technology.

Investigating previous researches has also given insight in what is not yet known about the different topics in this research. First, a robot in a storytelling language learning activity was previously used as learning companion. This research will use it as main character of the story, which is normally more popular for social skills learning. Second, the exact benefits of using a robot TA versus a virtual TA are not clear in a language learning activity. This research will give new insights in these specific implementations of a robot by combining them into one storytelling activity. The storytelling environment using a robot as teachable main character may have unknown benefits for primary school children's second language learning.

3. Ideation

This chapter describes how a concept for an activity was developed, based on findings of the literature research in chapter two (the state of the art). This concept is based on combining elements of storytelling and learning by teaching that contribute to L2 learning. Then, based on insights in game design and design in general, elaborate brainstorming was done to come up with a concrete idea that builds upon this concept. The source for these (game) design principles was the book *the art of game design: a book of lenses* by Jesse Schell [17]. Out of several ideas, one was chosen. These several ideas were different sets of design choices that each formed a coherent whole.

3.1 Ideation based on the state of the art

The possible educational benefits and characteristics of storytelling and learning by teaching, were listed in a notebook. On this list, the elements that L2 learning could benefit of were highlighted and became clear. This gave an indication of what elements would be required for the activity in order to benefit from both storytelling and learning by teaching to achieve a better whole and it led to the following concept. A less sketchy version of this list can be seen on image 10. The red part may need some explanation. During storytelling, complex stories could arise that the system cannot support (by means of the robot's interactional qualities for example). Next to this, the robot may be not very socially capable in general. This can be solved via tele-operation, which is often used for learning by teaching anyways. Less smooth social behavior also helps with framing the robot as 'in need of help', because it can actually do less, or do it clumsy.

Storytelling	Learning by teaching	L2 learning
- cause/effect creation	- active learning	- image-word learning
- table top	- robot in need of help	- learning in context
- tangibles	- discussion	- learning words in groups of the same subj.
- robot story enhancement with social behaviour	- tele-operation	- repetition
- complex stories		- communicating meaning
		- dialogs

Image 10 List of beneficial aspects

In this concept the robot, as main character, is in an imaginary world, portrayed on a table top. Objects in the world are represented by tangibles. Children make up a story about the robot in this world. The story involves the objects and the robot has emotions based on what happens to him in the story. In order to better understand this world and the objects, the robot wants to know more about the language of the world, in this case French in France. The children know more, in a way that will be discussed later, and will teach the robot the French words that he wants to learn based on what happens in the story. In this way, the children will learn French by repeatedly teaching French words and in a story context to a robot. Because the robot is social, dialogs and interaction may occur. In other words, the world on the table, its objects as tangibles and the robot as main character with emotions allow for a story to arise. This is a context in which words of certain subjects can be learned in a meaningful way. In order to make the children learn something, they teach the robot. The robot is in need of help and wants to learn. This creates dialog and interaction for L2 learning. Important repetition is achieved by making the robot forget things and/or by making him encounter the same objects multiple times. How this concept is based on the literature research is explained in terms of storytelling first, and then learning by teaching.

3.1.1 Storytelling

In storytelling, when the robot is the main character, the story can take place on a tabletop surface. The robot can move around and go to different places on the table (representing an imaginative world). This allows for collaboration amongst the different children doing the activity and forms a clear focus point for the children: the robot in the story world. Collaboration can happen in the form of discussion about the plot of the story and team work to control the robot and other characters and objects on the table. An example of this is [5] by Krzywinski. Looking from an L2 learning perspective, this table top world forms the context in which the words can be learned. There could, for example, be a supermarket or a zoo represented on the table, where the robot encounters supermarket products and zoo animals. The robot could then learn these words in the second language. This means L2 learning in context and in word groups of the same category, which was shown by Tonzar et al. in [13] to have added value over random words learning without context. The objects/to be learned words in the world could be represented by tangibles. Also in [5], tangibles have proven to be understandable and an intuitive way of interaction with the world. For L2 learning, these tangibles go hand in hand with the effective image-word learning method, where a word is represented by an image of its meaning, which was also investigated by Tonzar et al. in [13].

The robot can enhance a story with emotions and behaviour. The robot's emotions are the consequence of a certain event in the story, and may form the cause for the next event in the story that the child creates. In this way, the interplay between the robot and the world is given meaning. A simple example: The robot goes to the zoo where he encounters a lion. The lion escapes and threatens the robot, which scares him. However, before going to the zoo, the robot bought a big bottle of water in the supermarket. He uses this to scare off the lion, since lions/cats do not like water. The robot is happy again and continues his zoo visit to the monkeys. There the monkeys want the banana that the robot just bought… and so on. In this example, a sequence of events happens that give meaning to the objects and emotions in context. This could work well for language learning given the concept mediation model, which states that L2 words should be linked to their concept instead of to their L1 word. This model was established by Potter et al. (as cited by Tonzar et al. in [13]). Logical thinking and creativity of the child are stimulated in this way, because it is stimulated to create a story that makes sense to the robot in the world and its objects.

3.1.2 Learning by teaching

The children teaching the robot can be stimulated by making the robot in need of help. The robot alone cannot make sense of the world around him and needs some external source of information. In this case, that information should come from the children in some way. This need for help could be realized in terms of L2 level. It has been shown by Kory and Breazeal in [8] that the robot's language level has influence on the children's behaviour. In language learning with a robot, giving the robot a worse language level than the child leads to teaching behaviour by the child. In other words, the robot's L2 language level should be lower than the children's, and the robot should want to learn more, in order to understand the story world around him. A logical design choice would then be to let the story happen in France. There everything is in French and the robot does not understand that. The children do, in a way that will be discussed later. They may then feel responsible for and concerned about the robot. In this way, interaction and dialogs could arise. It has been shown that interaction and dialogs can contribute greatly to L2 learning by Kory and Breazeal in [8]. Learning is not just about exposure to words, but about communicating meaning and having social interaction in the context of the story. The robot, being social, can make this interaction happen. The words are given meaning by the story.

Another important factor is repetition of the words. The child will remember words better when he/she has seen them several times during the activity, instead of only one time. To realize this, the robot should ask about the same object many times. This can make sense if the robot encounters the object many times during the story. More importantly, he should forget what he learned. This way, it is a more logical action for the robot to ask again about the same word.

3.2 Ideation based on (game) design principles

The book *The art of Game design: a book of lenses* [17] introduces a number of important aspects of a game. All these aspects need to be considered and well designed in order to create an interesting and engaging experience. These relevant aspects are the resources of the player, his possible choices, surprise factors, rewards, freedom, goals, punishments, challenges, (game) elements, the story, character and elegance. While this research is not exactly about the design of a game, the book explains that the game itself is a mere medium to create an experience. An experience is certainly being created here. This experience happens in the mind of the player, and therefore, is very subjective. However with proper design, an experience can carefully be crafted in the player's mind. 'Character' means that the activity should have something funny, unique or strange, something that characterizes the specific activity. An example of this is Mario being an Italian plumber. This does not serve any purpose in the games but adds character to it. Elegance in a game means that a simple set of rules allows for a wide variety of game scenarios. This is achieved by only having elements in your game that serve multiple purposes. Many things can be considered about each aspect with respect to the experience. For example there can be many different goals for the player: build a nice world, create a great adventure, learn many words, mess around, help the robot, be as fast as possible or be the expert.

During the brainstorming process, several mindmaps were made with the above-mentioned aspects as central points. Possible (combinations of) solutions were created and written down in the mindmaps. Some ideas were then more elaborately brainstormed upon. Note that brainstorms happened based on all the aspects, but not all of them let to interesting ideas worth mentioning. Only the most interesting ideas will be described here.

Looking at player choices, the first interesting idea was to let children create the tabletop world themselves by placing tiles on the table. On these tiles, different locations, for the robot to go to, could be depicted. There was also an idea to let the children make roads or rivers in the world, via which the robot would travel. This idea would fit to the children's imagination of building your own (fantasy) world in which exciting stories take place. This could possibly work well with letting the child choose which objects/words to add to the world. Later, this whole idea was abandoned since it would probably focus the activity too much on world creation instead of actually interacting with the robot and L2 learning. Then, looking to player expectation and surprise, the idea arose that the child would either not know which words belong to which location, or not know which location belongs to which words. Each location that is being chosen would add some new objects to the world. This idea was abandoned for the same reason, it would place the focus too much on something else than L2 learning.

Possible rewards for the child were elaborately brainstormed upon. There could be some form of customization in terms of objects, the world, colors or something unique in the end. In order to achieve the great feeling of completion, there could be some end goal. Rewards could also be given in the form of new resources. For example more words, locations, side-effects or space. The robot could earn points, gain more powers, gain access to new content and say things like “good job!”. In a similar fashion, brainstorming has been done on factors like punishment and player cooperation.

A closer look was taken on the character of the experience. Different entities for the robot were thought of and the idea came up to make the robot a cheese. A Dutch cheese that has heard about all these amazing cheeses in France. He wants to find out more about France and the cheeses there so he goes on an adventure to find out. A possible goal for the cheese would be to not get eaten, learn about the French cuisine or how cheese is being made. The children could then be a cheese as well, a French cheese. As French cheese they could explain words to the Dutch cheese. Having a cheese head, the cheese would be not so smart, so the children would have to teach the words many times to him. This idea was a possible way to go. The activity would have character and make sense at the same time. The cheese going to France would be a nice starting point for the children to come up with stories. Given that the cheese is not smart, learning by teaching could also be realized.

An idea came up to make the robot a water droplet that starts in the mountains and flows all the way through France and ends in the sea. During his journey, he would encounter a variety of interesting locations. Another idea was to make the robot a sad dictionary that wants to learn more words. Then the idea came up to make the robot a Dutch mole that had dug a tunnel all the way to France. When he came up to the surface, he was fascinated by his surroundings and wanted to learn more about it. The children would have to teach him directions in French, to guide in which direction he would have to dig to go to a certain location. Blinded by daylight, the children would have to teach him about his surroundings, because he could barely see it for himself. This idea was also a possible way to go. It has a reasonable story and a reason to teach the robot.

Brainstorming was done from a learning by teaching perspective, instead of the storytelling perspective. In order to force learning by teaching behaviour, learning by teaching would have to be a requirement to get the rewards and to meet the goal of the activity. It would be a requirement if the robot cannot learn for himself, maybe because he has no eyes or forgets things very fast, like a gold fish (is the common belief). Another option is when the robot does not fully understand the world, like an infant, a baby or a caveman that has been transported to modern times. Then the idea arose of a grey blob that wants to be colored. The child can teach him a word. When learned, the blob gets (partially) the color of the learned word. However, the colors fade away over time, representing the memory, so the word needs to be learned again, in order to regain the color. Colors can be made interesting in several ways, think of special colors like gold, mixed colors, changing colors, moving colors or rainbows. This idea stood out from the others because of its elegance. In this case the colors, for example, form the reward, create surprise, trigger repetition by representing the memory and can be a cause for the robot’s emotions to change as well as for how the story continues. This idea was also an option, given the required learning by teaching and elegant way in which many important factors were integrated in the design.

Out of these three ideas, the cheese, the mole and the blob, the blob was chosen to continue with and to further develop into a low fidelity paper prototype. The blob was chosen because the idea includes the most aspects, that contribute to a meaningful and interesting experience, elegantly in one activity. Especially the many colors as a reward, and at the same time representing the memory, seemed an elegant solution to trigger learning by teaching repeatedly. However, the animation of the blob on the screen getting colored would have to be appealing and spectacular as a reward for the children. The only downside of this idea was that there is little reason for storytelling. A blob is rather characterless, and therefore there seems no convincing reason to go to France, no meaningful motive for a story to arise. To solve this problem the blob was changed to something else grey. Options were a mouse, wolf, cloud, parrot or elephant. The elephant was chosen simply because it seemed the most fun idea, also giving the activity some more character.

4. Specification

This section describes how the chosen concept was further developed in an iterative process of prototyping, testing and improving. In this process, two paper prototypes were made. The second paper prototype was tested twice on a local daycare. Paper prototypes were made to identify design flaws, unanswered design questions, and the limits of the design. Examples of questions identified and answered here are ‘How should the activity prevent certain misuses?’, ‘What role will the elephant’s text and emotions have?’, ‘How will the child itself know French translations of the words it wants to teach?’, and ‘What will be a clear end goal of the activity?’. Note that many aspects of the activity were not designed yet at the start of this phase. Only gradually during this phase, problems were identified and solved, and design choices were made. The consequence of this is that not all interactions/aspects are clearly described from the beginning of this chapter, because they were simply not designed yet. For example the interaction of how the child ‘teaches’ the robot.

4.1 The first paper prototype

This first low fidelity paper prototype consisted of four elements: the robot, five locations, fifteen objects (three per location) and some color for each object to put on the robot when the word is learned. The number of locations and objects was chosen rather arbitrarily. However it was kept in mind that the number of different words introduced did not have to be so large. Children need time and repetition to learn some words properly, as was learned from Vogt et al. in [16], so it is useless to introduce them to a wide variety of words when there is only limited time to learn them. That is, while there is only limited time to play with the activity. As can be seen on image 11 and 12, the robot, represented by the round, yellow paper, is surrounded by the five locations and three words per location are represented by a piece of paper in the color of the object. The locations with their words and the color:

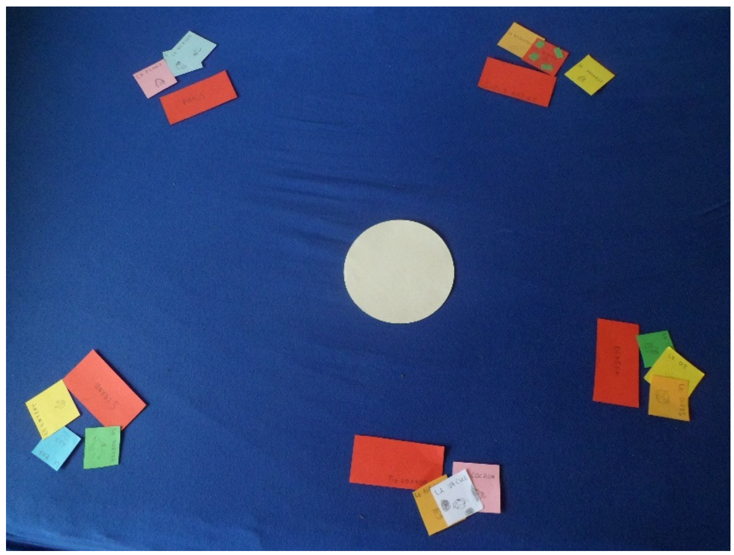


Image 11 First paper prototype top view

Image 12 Close up on the farm and its objects

- The supermarket
    - The cheese – le fromage – yellow
    - The apple – la pomme – red and green
    - The baguette – la baguette – orange
- Paris
    - The man – le homme – blue
    - The woman – la femme – pink

    - The car – la voiture – red
- The farm
    - The pig – le cochon – pink
    - The cow – la vache – white and black
    - The farmer – le agriculteur – orange
- The mountains
    - The bear – le ours – orange
    - The gold – le or – yellow
    - The cactus – le cactus – green
- The beach
    - The child – le enfant – yellow
    - The turtle – le tortue – green
    - The water – le eau – blue

Note that the colors did not always exactly match the object. For example an orange bear. This was purely because of a lack of differently colored paper. Also note that in French it is mandatory to write for example l'enfant instead of le enfant. This was left out because children of a young age may get confused by this apostrophe. The specific words were chosen based on the possible value in a story, simplicity of the French word and characteristic color of the word. Note that also a cognate word was added, 'le cactus'. Later in the process a second cognate word was added, 'le bebe', which will be discussed later. By adding these, the influence of cognate words could be observed. It was stated by Tonzar et al. in [13] that children learn these more easily. The locations were chosen based on possible value in a story and on variety. With the 'value' is meant to what extent the object or location makes the child want to use it in its story, to create a story that may be exciting, or surprising. In other words, words or locations that possibly add fun and diversity to the activity. The words were written in French on the objects because it is in France, so everything is in French. The locations were in Dutch, purely to remember what the location was. On the final prototype, these locations would be represented by images.

The concept was evaluated by own playing and experimenting with the paper prototype, acting out possible scenarios and trying to find troublesome situations that users also could encounter. Moreover, the holographic design method, described by Schell in [17] was applied here. While playing with the prototype, the researcher observed his own acting. He then tried to find a reason for this specific acting and why no other actions or interactions were taken. This way, possible natural interactions were identified. An example can be seen on image 13. Naturally, the objects that the researcher wanted to teach to the robot, he moved these in front of the location. On image 13, this happens with the turtle and the water. This interaction was not pre-designed, but the researcher observed himself naturally doing this.

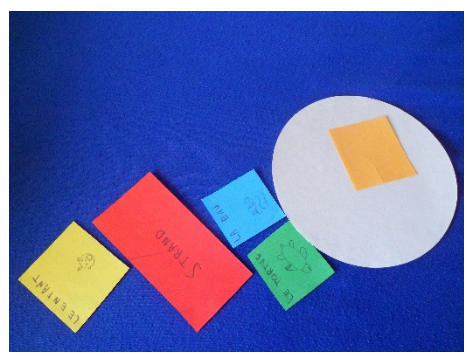


Image 13 Robot encounters two objects

This playing and experimenting was done in the following way. The robot started in the middle, and from there it was moved to one of the locations. Interactions, like word-teaching were not specifically designed yet and here this holographic design approach was used. To teach a word, it was put in front of the location. When taught, a piece of paper of the same color as the taught object was put on the robot. During the activity, the robot collected several taught words this way. Sometimes a word was forgotten by removing the piece of paper. To relearn, the robot had to go back to the location of the forgotten object. Some storytelling was done as well, mostly as reason to go to a different location. This was however very hard to mimic from a young child, given the big age difference and different levels of overall development. The role and content of the robot's speech and emotions was not clear yet at this point in development.

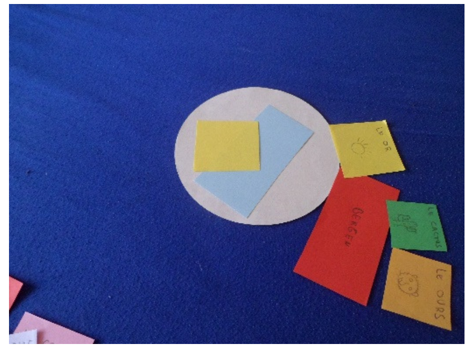


Image 14 Robot encounters gold

Many points that needed some more attention were identified with this simple paper prototype, via the just-described playing. On image 14 for example, the robot encounters gold. However, it feels not interesting to learn the word since the robot already acquired the yellow color.

1. How does the robot recognize the correct object that the child wants to teach?
2. Are the robot's emotions based on the story or on the word learning process?
3. What shape will the tangibles get?
4. What does the user interface look like?
5. How does a child teach a word to the robot in an intuitive way?
6. How does the child know the French translations of the words in order to teach them?
7. What things does the robot say to steer the learning activity?
8. How does repeating the same words not get boring if the robot needs to go back to the location of the forgotten word all the time?
9. What happens when the robot encounters two or more objects at the same time?
10. Having the same color for different words, it is not motivating to learn the word when the color is already acquired via another word. Colors should be divers and unique.
11. Make sure that the child cannot teach the robot many words in a very short time, without creating any story. This is not a proper way of word learning that the activity should allow. Like this, attention may be more focused on trying to seek boundaries of the system, than actual teaching.

The current setup of the real robot was experimented with (image 15), to get more insight into the already existing abilities of the system that could maybe solve some of the above-listed problems. This was also useful to get used to starting up the system, which is a procedure consisting of many steps. Via the real robot setup, it was found that the system could track objects based on a QR-like code and could act based on that. For example, the robot could move in the opposite direction of an object representing an alien. This tracking system could be used to recognize the 15 different objects and react based on that. For example change emotions to happy or worried when encountering a child (the story object) alone on the beach. This could also be used to make sure that the child teaches the correct object to the robot. That is the object that is closest to him, or at least in some verifiable reach. However, then still there is the problem of having more words close to the robot. The child could still choose the wrong word. The word would be wrong if it is not the word that is being used in the story, or not the word that the robot asks for.

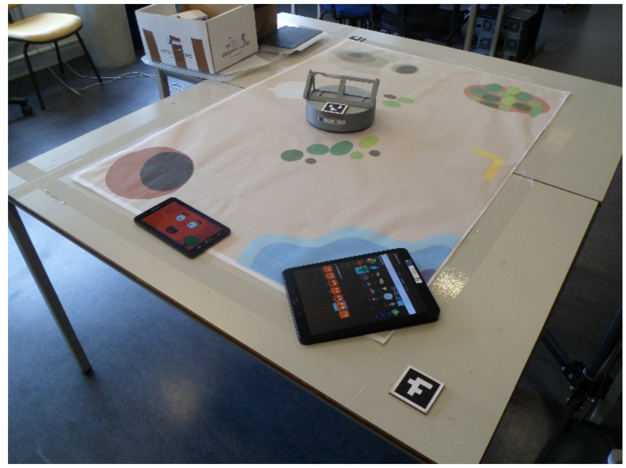


Image 15 Ideation with the current setup

The existing system's tangibles are cubes with an image on the sides and the track code on top. This shape would also be convenient for this activity. It is better than a 2D representation because the code as well as the image can be presented on an object of realistic size compared to the robot and the table. Having the tracking code and an image both on a 2D surface would require big tangibles. This size may get out of proportion compared to the robot and/or the locations.

4.2 The second paper prototype

With these above-listed issues in mind, a second, more advanced paper prototype was made. This prototype used more different colors and carefully selected images instead of quick drawings. The prototype consists of: Five locations, fifteen objects, fifteen colored things to represent a learned word, the robot in four different versions (each with another emotion), and the user interface (a tablet app faked on paper). See image 16 through 19. Some minor changes are that 'the water' was replaced with 'the sea' and 'the child' was replaced with 'the baby'. This to have an easier French word (la mer and le bebe) and a possibly more interesting object for a story. Paris was changed to 'the big city', since children may not know Paris at all or relate it to things that, when used in the story, cannot not or hardly be supported by the system, like Disney land.


Image 16 Top view of second prototype

The interface was made in a way that seemed easy to understand. Other interface designs were sketched as well but did not seem clearer or better in any significant way. On the left side is an overview of all the French words with an image, on the right side is an overview of the map. The child can do two different actions on the interface. Firstly, on the right hand side, it can click on the robot-icon and click on a destination, then the robot goes there. Secondly, by playing with the prototype, an intuitive way for teaching the word to the robot was found. The child drags the word from the overview to the robot icon on the map and drops the word. If it is the correct word, the color animation plays on the actual robot, meaning that the word was learned. This seemed a convincing way to teach a word ("showing" it to the robot). It also links dropping the word on the robot to the coloring animation. Note that with this interface another big problem was solved. This design allows the children to know the French words that they want to teach. It simply displays them in an overview. This way, the children know more about the French than the robot and they can now help him by teaching.

Image 17 User interface

By own playing with this second prototype it was realized that still many of the same problems from the earlier-described list occurred and that the solution may lie in the robot's sentences that it could say. From the above list, problems seven, eight, nine and eleven could be solved by carefully designing the sentences that the robot says and the timing when he says them. For example problem eight. In order to learn a word several times without having to go back to the same location, the robot could just say something like "Oh no I forgot what a cow is in French! Could you teach it again?". In a similar fashion, problem eleven that the child rapidly teaches many words without any story, could be solved. The robot could say text like "Ho wait, not too fast! I am still learning le fromage". The sentences could go hand in hand with the robot's emotions. He could be sad when having forgotten a word, and happy when he has learned one. At the same time, the emotions could still be used to support the story. For example amazed when he sees the sea.

Image 18 Close up on the mountains and its objects


Image 19 the four emotions of the elephant: normal, amazed, sad, and happy

Then a list was made with all possible sentences of the robot and when he would use them with what accompanying emotion. However, this seemed impossible to do without a clear list of how all the different aspects would come back in the activity. So first, this list was made to get some clarity. This forced the researcher to make some important decisions. For example to set the goal of the main character to ‘going on an adventure and learn French words’ instead of ‘getting colored’. The latter would differ too much from the designer’s goal, to teach the children some French words via learning by teaching in a story context. With ‘getting colored’ as goal, it was expected that learning by teaching behaviour would trigger less strongly, since children can get the colors on the elephant by dragging the words via the interface, without actually teaching anything. It is the desire of the elephant to learn words that should lead to the teaching. When teaching becomes the goal of the child, it may actually focus on the teaching, instead of on the coloring. It may put in more effort in teaching, for example by speaking out the words. Jesse Schell in [17] highlights how the goal of the main character should be the same as the goal of the player, the word-learning/teaching in this case, next to going on an adventure of course.

Another important decision was to let the elephant say story-related sentences during the activity. This was done to evoke storytelling behaviour. This way it seemed less likely that the children would just go to the locations, teach the words, and continue, without any story context. As learned from the State of the Art research, the storytelling component in the activity adds several important factors that contribute to effective L2 learning. For example the use of the words in their context [8], which realizes word learning via the concept mediation model [13]. Additionally, without any told story, the activity would last too short and the words would be taught too fast to the elephant. This way there is no opportunity for repetition of the words to take place, which is important to remember words in the longer term. It was decided that the elephant would be able to say two things per object that it can encounter. It can either directly ask for a word’s meaning in French, or it can ask/say something related to a story, which can be seen in table 1. The first would take place if the child already makes up a story himself. The second if the child does not do that. This way the activity can be steered by evoking either some more storytelling or some more teaching. Note that, if this feature makes it to the final prototype with the robot, tele operation will be needed, since autonomously deciding what to say based on the current situation and past actions is out of the scope of this project and its goal. Also note that the child is not forced to tell a story, it is only strongly guided in that direction, by the elephant’s different texts, goal, and availability of locations to go to and objects to use. The activity is made suitable and attractive for storytelling.

A description of the idea in terms of Schell’s [17] aspects of an experience:

Goal of the main character: Go on an adventure to France and learn French words
Goal of the child: Go on an adventure to France and teach French words (not necessarily as a story character itself, but still present in a way. It is hard to tell at this point how the child sees itself within the activity). Note that the coloring was not a child’s goal anymore, because it could move the child’s focus too much to the coloring itself, instead of the teaching. The colors were still present, but as a reward of teaching and to make the objects more distinguishable and to make the activity more colorful.
Elements: Five locations, fifteen words, the main character, the interface, the colors, the emotions
Choices: The plot of the adventure, what words to teach? What colors to give the elephant (by teaching what words)?
Challenge/success: Teach the words repeatedly, keep the main character happy

Rewards: Happy main character, colors on the main character, compliments by main character
Surprises: Plot twists mentioned by the main character, nice (combinations of) colors on the main character
Punishments: Sad main character and sad sentences.
Resources: The fifteen words that the child can teach and use in his story, the five locations.
Freedom/control: The child is meant to come up with a story, and preferably with the objects given. Not so much with imaginary objects that it forgets the given objects. However note that this storytelling freedom is less present in the child's experience when the child does not come up with a story itself and the pre-made story sentences are used often.
Cooperation: In this state, there is no real need to do the activity in duos. It triggers communication and discussion however.
Character: A Dutch elephant as main character that goes to France to learn French, and for an adventure, meeting and interacting with all kinds of strange objects for an elephant, like a car or a bear.
Elegance: Learning by teaching without the children realizing it. The color system, being a possible surprise, a reward, a choice and an intuitive way of repetition by fading away.

In table 1, all possible sentences for the robot are listed. Note that this is translated to English. For user testing, a Dutch version was used. More importantly, note how an extra repetition was put in the learn and re-learn sentences. The robot repeats the Dutch word as well as the French word. Besides for getting more repetition, another important benefit of this is that the child can hear the correct pronunciation of the French word. The robot actually pronounces the word correctly. Knowing that the children would probably not pronounce it correctly at first, this was added. It may seem strange to the children that the robot immediately pronounces it correctly. It could hurt children's image of the robot as student that knows less than they do. It will be found out during the user test if this is an issue or if the children mention this while interacting.

E = elephant
X = certain object in Dutch
Y = certain object in French

Table 1 – Robot sentences, first version

When	What	Emotion
Activity starts	Hello! Let's go on an adventure in France and learn some French words! I need your help with that! Where do we go first?	normal
E encounters object X	Interesting! What is X in French?	normal
E learns a word	Ahaa, so X is Y in French! Thank you for teaching me that!	happy
E is being taught too fast	Hoo wait not so fast! I am still busy with X and Y.	Sad
E forgets a word	Oh no, I forgot what X is in French! Can you teach me again?	Sad
E relearns a word	Oh yeah, X is Y in French! Thank you for reminding me of that!	happy
Child teaches E the wrong word	Yes? Is that really the correct translation?	Sad

E wants to go somewhere else	Oké, where shall we go now?	normal
E has learned three words	Wow I have already learned three words! Thank you for your help!	happy
E has learned six words	Wow I have already learned six words! Thank you for your help!	Happy
E has learned nine words	Wow I have already learned nine words! Thank you for your help!	Happy
End of activity	What an adventure was that! Thank you for teaching me all these French words. I can't wait for our next adventure.	Happy
E encounters the bear	Whoo a bear! Would the bear hide something in his cave?	amazed
E encounters the gold	Real gold! Maybe I can buy something with this gold!	happy
E encounters the cactus	Hee a cactus! Maybe I can eat that! I am hungry.	normal
E encounters the cheese	Hmmm this cheese smells much better than Dutch cheese! How would they make it?	amazed
E encounters the apple	Hmm I am hungry but I have no more money!	sad
E encounters the baguette	Hmm a typical French baguette! I would like that!	amazed
E encounters the baby	A lonely baby?! Where are its parents?	amazed
E encounters the sea	Let's go for a swim! It is so warm here in France	happy
E encounters the turtle	The turtle seems so lonely, he could join me on my adventure!	normal
E encounters the man	The man seems so happy! Why would that be?	normal
E encounters the woman	The woman seems so happy! Why would that be?	normal
E encounters the car	What a nice car! Maybe I can borrow it	happy
E encounters the farmer	Maybe the farmer knows some nice places to visit in France	normal
E encounters the cow	What a friendly cow, she could be my friend!	happy
E encounters the pig	Hmm the pig is all dirty of mud! He needs to be washed	normal

4.3 First user test with second paper prototype

The second paper prototype was tested at the daycare on the university of Twente campus. Children of about 8 years old were selected as target group. As described in the State of the Art, children aged between 6 and 8 are old enough to make up a structured story, and young enough to still use toys in their story [2]. For this activity, children could not be younger than 7 because reading skills are required (to read the French words on the interface). There was no clear upper limit on the age. This would need to be found out by trial and error. It could be that children older than eight are simply not impressed by the design and goals of the activity. Then they could feel less engaged and less motivated to achieve those goals. However, this age limit could in principle also be nine, ten or older. Maybe the border is not so clear and some children of ten are still engaged while others of eight are not.

For this test, four children, aged eight and higher and whose parents consented, were present at the daycare at the moment of the test. In the end only one child wanted to participate. This was a girl of

nine years old. The paper pieces were laid down like in image 16 and the elephant with his goal were introduced. This girl was not particularly engaged in the activity, but also not bored. In terms of storytelling, she was rather passive and did barely come up with story elements herself. Questions asked by the robot (Would the bear hide something in his cave?) were often answered with “hmm, I do not know…”. However, she wanted to use the gold to buy the supermarket products (after the robot mentioned he was hungry but had no money) and she figured out that the lonely baby on the beach was watching the turtle (after the robot asked why the baby would be alone). In other words, she linked some of the provided objects to each other after the robot encountered and mentioned them. She did not do this before the robot had said something about it.
There was no clear end of the activity. At a certain point, the researcher just stopped it by putting the elephant in the middle of the table and saying what the robot would say at the end.

In terms of learning by teaching the activity went decently. In the end the girl had taught the elephant nine words. She also re-taught a word when the elephant forgot one, and tried to pronounce the words. She was not only matching images but actively busy with the words while teaching them and interacting with the interface. However, she was struggling with the pronunciation of the French words. This improved after the elephant had pronounced them correctly. Right after the activity, she was asked which French words she remembered. She was able to reproduce six words, with images of the fifteen words spread out before her. Interesting to mention is that she remembered many words that she had used in the story: the turtle, the baby, the gold, the sea, the cheese and the bear. However she had been on holiday to France several times so she already knew the cheese and the sea in French. Interaction with the user interface went smoothly. After showing two times how to move the elephant and teach him words, she was able to do this.

Even though the prototype was tested on only one child, who might simply be too old to be engaged, some points for improvement could be identified that would probably increase the engagement of children:

- Make sure there is more storytelling
- Add a story end goal
- Use words that are easy in French

To create more storytelling and to add an end goal, the following solution was found: The elephant has a concrete end goal in France, something more concrete than just “going on an adventure”. To achieve this goal, interaction with the objects is required. Knowing the French translation is needed to interact with the object. This way, teaching the words contributes to achieving the elephant’s goal. This goal is also the child’s goal. The objects are linked to each other from a story perspective, so they are used in their context. For example: The end goal is to have a picnic on the beach with the elephant’s new friends. To achieve this goal, he needs to make friends in the first place. Additionally, he needs to buy food in the supermarket. He has no money so he goes to the mountains to get the gold, which is hidden in the bears cave. The bear can be paid with honey, which can be bought at the farm. But the farmer is busy with washing his pig, so the elephant… and so on. This way, the elephant will need to know some French words to achieve his goal. He should know the French word for honey in order to buy it from the farmer for example.

Without any story, it would be just teaching the robot words because he asks for that. Other than the elephant's interest in French, there would be no reason to teach him French. If the robot happens to need French words in the story, it would make more sense to teach him those. He needs to know French in order to achieve his goal. Jesse Schell mentions how players will almost always do things that contribute to achieving their goal [17]. That would be learning by teaching French in this case. Besides that, the story would place the objects in their context, which may be better way of language learning than without any context, argued by Kory and Breazeal in [8]. On the one hand it may be coincidence that the girl knew exactly the words she had been using for story elements. On the other hand, this might show the effectiveness of learning words in context. Additionally, the few times that the girl made up story elements, based on the robot's questions, made her smile. So the story can make the activity more fun.

The girl was (obviously) quite struggling with the French pronunciation. As said above, this improved after the robot had said the words some times. Nevertheless, some extra attention should be paid to the words that are chosen in terms of difficulty. Especially when taking into account that this girl was already nine years old while other children will be eight or even seven.

As a final note of this section, it is important to realize that the lack of engagement was also caused by the fact that the prototype was made of paper. Robots elicit much enthusiasm and engagement for children as was learned during literature research in [7] from Knapp et al. (as cited by Werfel). The final prototype, using a tablet, a robot and some animations, is almost guaranteed to be more engaging.

4.4 Improvements to the second prototype

According to the points for improvements, as stated above, the prototype was adjusted. Many of the story-related sentences were adjusted to form one coherent story with an end goal (see table X). This end goal was to go surfing on the beach (location) and it was made clear to the child via the introduction text of the elephant. Arrived at the beach, the elephant was hungry and first needed to go to the supermarket. There he had no money so first needed the gold from the mountains. This gold was guarded by the bear and was only given in trade of honey. This honey was given by the farmer once you would bring his pig back that was wandering around in the city. Note that several objects changed their starting location in this improved version. The car was moved from the city to the mountains, the baguette (now honey) was moved from the supermarket to the farm, the pig was moved from the farm to the city, and the cactus was moved from the mountains to the supermarket. Besides that, the baguette was replaced by the honey. This was needed for the story. The baguette was less interesting story-wise and it was a difficult word to pronounce as well. Story sentences for some other objects were also adjusted (See table X). For example it was made clearer that the baby on the beach had to be brought back to his parents, that the cow makes the cheese, and that the turtle wanted to join the adventure.

In addition to the story-related end goal, also an end goal in terms of word learning was added. A progression system was added on the main character, as can be seen on image 20. Now there was a clear goal of eight words that had to be taught. The elephant would now have to be moved to other places together with this underlying sheet. This whole was meant to be displayed on the robot character screen in the final prototype.

Image 20 word-learning progression system

This improved prototype was tested with by the researcher himself and quickly was concluded that the story was complex. In the way it was made now, the elephant would need to visit all places twice. Once on a place, it would become clear that the elephant first needed some object from another place each time. One could say that the story would be played backwards first, to find out about all requirements. The elephant would start at his end goal, the beach, then go to the supermarket, the mountains, the farm and the city, and then do everything again but with the required objects. He would get the pig, the honey, the gold and then the food. To fix this issue, the elephant would need to start in the city. There he would find out about the pig, get honey at the farm, and so on. However, knowing the end goal from the beginning, a child would likely not go to the city first. The problem was solved by perceiving the surfing not as an end goal, but simply as a guided starting point to make the children enthusiastic. An elephant that wants to go surfing is rather strange and would make the children curious to find out how this would turn out. This idea was based on the interest curve theory, pointed out by Jesse Schell in [17]. On image 21, a desired interest curve for an experience can be seen. Point B is called 'the hook' and is meant to grab the player's attention from the beginning and get them excited to find out more. The unusual start goal of the elephant would contribute to establishing this hook. The shape of the rest of the curve was expected to be achieved via the already existing elements in the activity. For instance the teaching of words and the progression of the story. If this would turn out to be not the case, the design could be changed accordingly in a later stage of the project. Especially the G peak could not happen, this would depend on how the child's story would evolve.

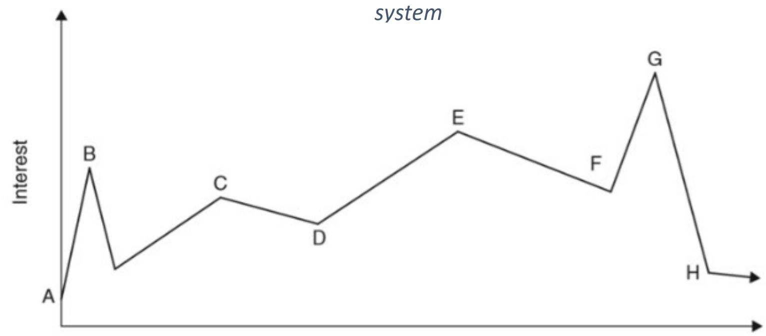


Image 21 Schell's interest curve

4.5 Second user test with improved second paper prototype


Image 22 prototype setup for user test

Because the first time only one child participated, and because the prototype had changed again, a second user test was arranged with the daycare on the University of Twente campus. The test setup can be seen on image 22. This time eight children participated, some in duos, others alone. There were four boys and four girls. Two aged six, one aged seven, three aged eight, one aged nine and one aged eleven. From a design perspective, there was no clear preference yet for either alone or in duos, so both were accepted and tested with to see what would happen. There was a preference for eight-year-olds, based on the literature research, but it was not yet confirmed that this would be an ideal age, so also other ages was tested with. Children were not put in duos, some preferred that themselves. Two children participated alone, six in duos. After the test, some questions were asked to the children:

5 What did you like about the activity?
6 What did you not like about the activity?
7 What do you think I wanted to achieve with this activity?
8 Which French words do you still remember? (show the images)
9 Age, gender

The eleven-year-old child was alone and participated first. He was not particularly amazed by the activity, but was not bored either. He did not do much with the story elements told by the robot, nor did he come up with his own story. He visited all the places, however, and taught the elephant the words. Then a duo of a seven-year-old boy and eight-year-old girl participated. The highlight of this session was that they started to move the physical objects around. For example they moved the pig to the farm. Afterwards they answered to the questions that they liked to move the tiles around. Then two girls of eight and nine participated. They were visibly amused by the activity and played the activity in the way it was designed. They moved around and taught words based on the story elements. Then two boys of six participated. They were clearly more intrigued by the activity. They came up with their own story and taught the elephant the words for their story. They often neglected the robot's story sentences, or barely heard them, because of their own story. They also moved the elephant around very fast on the tablet, to see how the (paper) robot would respond. Finally an eight-year-old girl participated. She also came up with some own story elements and started to reason why the baby was left alone on the beach, for example.

In general, the end goal of teaching eight words worked very well. Without explanation, the children understood that after having taught eight words, the activity would stop. The improved story sentences worked very well to keep the activity going and to establish reasonable cause-effect story elements on the one hand, and still give them freedom to come up with their own story on the other hand. All children went to the beach at first, because of the elephant's text. In other words, the newly added mechanics worked as intended and had a positive influence on the experience. Several words were remembered afterwards, that were also used in the story, but often the same, easy words like la mer, le ours, le or, le miel and le bebe. Pronunciation of the French words went surprisingly well.

Some things went less well. The children spoke out the French word, when the elephant asked for it, but then they did not immediately teach it to him. However, this could be caused by the prototype being from paper. When the tablet actually works, the teaching mechanic may be less easy to forget. Also when the elephant says the text, instead of the researcher, it may be more obvious to teach, instead of only to reply to the question. Another issue was that the children taught the elephant some words from other locations than where the elephant was at that moment. For example, he is in the mountains and says that the bear wants honey in exchange for gold. Then the child teaches him honey and takes the gold or the bear, probably assuming that the honey is available once taught. On the one hand this can be seen as an issue, because the child does not take the desired action as a result of the robot text. On the other hand this behaviour is not necessarily wrong, since the child is free to do whatever he wants, and the story elements are only a guidance to keep going. Anyway, there was some confusion about how the robot can move an object or acquire it story-wise. This was confused with acquiring a word learning-wise. Children taught words in order to acquire it for the story instead of interacting with the physical objects. As said above, only one group did this correctly by moving the physical objects.

Based on the children's answers to the questions afterwards, some things can be said. Many children liked the freedom of the activity. They liked that they could choose where to go and with what objects to interact. In addition, they liked to teach the elephant. Some were aware that the activity was meant for them to learn the French words. They mentioned that they liked to learn a language in this way. The fact that it is a paper prototype and that there is no real tablet and robot was the only negative point mentioned. In general, the children were enjoying the activity. Some thought the maker's goal was to teach the elephant, some thought to teach the elephant and the student, some thought to teach the student. Opinions were most divided on this question. One could wonder if all children really understood this question.

Some things were executed a little differently during this test, than intended for the technical prototype. The robot's emotions were not changed for example, he always had the normal face. Because of the paper prototype, the children's empathy for the elephant was less triggered, so changing emotions would likely not cause any changes in behaviour. Besides this, the test's goal was to evaluate (core) design choices, the emotions is not a major one of them. The hassle of constantly changing them would probably disturb the flow of the activity and therefore lower the engagement more than increasing it. In the actual prototype, the emotions will be adjusted with less effort, so it can be executed properly. Moreover, sometimes elephant texts were not spoken out by the researcher. Either because they were forgotten, or found to be redundant. The sentence "oh, so X is Y in French" after having learned a word, was often forgotten or skipped. This may have negatively influenced the word learning process of the children since this sentence forms an important form of repetition of the translations. It was often skipped because the child had already continued its adventure so the correct timing was gone, because it did not expect this sentence. In the actual prototype, sentences will be said more consistently and well-timed. Most important, per session, the elephant forgot a word only about three times in total. This was not done as often as intended because there seemed no proper timing for it, similarly to skipping sentences. It felt not right to disturb the children in their play. In the actual prototype, this will probably be some timed event that happens regularly. However, as learned from this test, forgetting a word could disturb the children's play and might therefore be programmed to happen at certain moments when there is less action. This will be discussed in a later section of this report. Executing these aspects a little differently than intended has some negative effects on the representativeness of the test. Especially since one of this test's goals was to evaluate the flow of the activity with the new story-related robot texts. Skipping sentences about word-learning may have a positive influence on the flow of the activity story-wise.

4.6 Final concept

As stated above, a major point of confusion was the vague interaction with the physical objects. It was not clear that this was allowed and there was no clear place on the robot to carry an object. One boy wanted to take the cow so he put the tangible on the elephant, but this looked a bit strange and he removed the cow again. This confusion resulted in story hints not being followed, because it was not clear how to move/trade objects. It also resulted in words being taught when the robot was not at that specific location. As said, this is not necessarily wrong, given player freedom, but still it does not always make sense that the elephant learns words he has never encountered. This is not the intended way in which the story is linked to teaching and learning words.

To solve this problem, an obvious way for the robot to physically carry an object was created. Because from now on the actual prototype was being built, this solution was not made for the paper prototype. As can be seen on image 23, the grey robot base has a marker on top. This marker is needed to track the robot's location. As will be discussed in a later section, the tracking feature was not used in the final prototype. This spot on the robot could then be used to carry a tangible. The child would have to put it there him/herself. Exact realization of this feature will be discussed in a later section of the report as well.

Image 23 marker on the robot in the existing coBOTnity system

This new feature would emphasize the affordance of the robot to carry an object and to move it to another location. It would strengthen the link between the story and the physical world because now the story hints could be executed in the real world. The real world would therefore more convincingly represent the (state of) the story. Also the teaching action would less likely be linked to acquiring an object for the story. Now the children would interact with the tangibles more, instead of only with the tablet. The benefits of using tangibles, as described in the state of the art section, would be realized more since the tangibles would actually be used. In addition, one more robot sentence was added to prevent that the children would start to pile up multiple tangibles. “I can only take one object at a time with me.” Stacking tangibles is not desired. For the child this could mean that the elephant can collect them and this may become a goal for the child which does not necessarily add to storytelling or word-learning. The ability to collect them can also easily escalate, since the child may want to collect them all and build a big pile of tangibles on top of the elephant. This may be funny, but it does not add to activity’s goal.

Some other minor changes were made. The ‘amazed’ emotion was removed. This emotion was not often used, only some times when certain objects were encountered. It was not used to support the word-learning aspect of the activity. Now there were three emotions left: normal, happy and sad. These three were often used, both for the story and the word-learning. In addition, some robot sentences were adjusted. See table 2 for an overview of the final sentences. As will be discussed in a later section, the sentence after learning a word and after relearning a word was made the same. Also the learn and relearn sentences were rephrased and the maximum words to teach was set to eight. The sentences of the man and the woman were changed to make them more obviously linked to some story. Their initial sentence (the (wo)man looks so happy, why would that be?) was too vague. Note that also in this table the sentences were translated from Dutch to English, since the activity had taken place in Dutch.

Table 2 – Robot sentences, final version

When	What	Emotion
Activity starts	Hello! Let’s go on an adventure in France! I would like to go surfing. I also want to learn some French words but I need your help with that! Let’s go to the beach!	normal
E encounters object X	X, can you teach me what that is in French?	normal
E (re)learns a word	Ahaa, so X is Y in French! Thank you!	happy
E is being taught too fast	Wait, I cannot learn words so fast.	Sad
E forgets a word	X, what was that again in French? Could you teach me again?	Sad
Child teaches E the wrong word	Yes? Is that really the correct translation?	Sad
Child wants to take multiply objects	I can take only one thing with me.	Sad
E wants to go somewhere else	Oké, where shall we go now?	normal
E has learned four words	Wow I have already learned four words! Thank you for your help!	Happy
Child wants to teach ninth word.	I cannot remember more than eight words!	Normal
End of activity	What an adventure was that! Thank you for teaching me all these French words. I can’t wait for our next adventure.	Happy

E encounters the bear	Whoo a bear! The bear gives me gold in exchange for honey!	happy
E encounters the gold	Real gold! Maybe I can buy food with that! Or something else.	happy
E encounters the cactus	Hee a cactus! Maybe I can eat that! I am hungry.	normal
E encounters the cheese	Hmmm this cheese smells much better than Dutch cheese! How would they make it?	normal
E encounters the apple	Hmm with gold I would be able to buy this apple!	normal
E encounters the honey	Hmm honey, bears would like that!	normal
E encounters the baby	A baby alone?! Let’s bring it back to its parents! Where would those be?	Normal
E encounters the sea	Aaah the sea, let’s go for a swim! It is so warm here in France.	happy
E encounters the turtle	The turtle seems so lonely, he could join me on my adventure!	happy
E encounters the man	The man would like to walk in the mountains.	normal
E encounters the woman	The woman wants to sunbath on the beach.	normal
E encounters the car	What a nice car! I can drive everywhere with that!	happy
E encounters the farmer	The farmer gives me honey in exchange for bringing back his pig	happy
E encounters the cow	What a friendly cow, I think she makes tasty French cheese.	happy
E encounters the pig	A pig in the city? He has to be brought back to the farmer!	normal

5. Realization

This chapter describes how a high-fi prototype was built for the activity. This consisted of three tablet apps that were programmed in the Processing language. Processing was imported in Android studio to develop apps with this language. These apps could communicate with each other via internet. A technical schematic of the general working of this setup can be seen on image 24. There is an app for the robot face, an app for the child UI, and an app for the tele-operation UI. The child UI, being a client, can send a 'word' message to the robot, being the server. This happens when the child teaches a word. The tele-operation UI, being a client, can send a specific integer 'i' for each button, when that button is pressed on the UI. For some functionalities of the system, the server also sends a message back to the clients. Requirements per app are listed below. Note that these are the main requirements. Additional minor (quality of life/ease of use) functionalities are not in the list. These will be discussed later in this chapter.

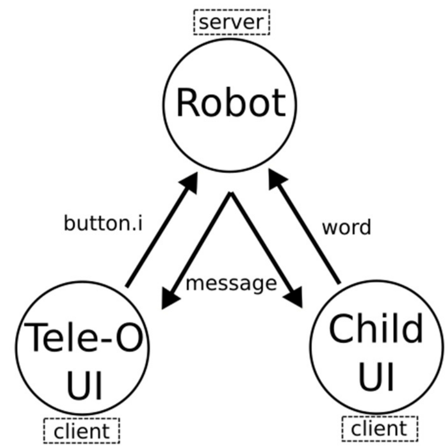


Image 24 schematic of internet communication between apps

Child UI:

10 The child can teach one of fifteen words to the robot by dragging something, that represents the word that is being taught, to a robot image on the tablet app.
11 When the child releases this representation of the word on the elephant, and a word is taught to the robot this way, a small colorful explosion is triggered on the place of the elephant image.
12 An overview is visible of fifteen tiles. There is a tile per word. On each tile, the French word is displayed, alongside the image of the word that is also being used for the tangible of that word
13 A map which looks similar to the physical playground is visible on the UI.
14 The image of the elephant is visible on this map
15 The image of the elephant can be moved to different locations on this map.
16 This movement happens via one tap to select the elephant, and a second tap on a new location to send him there.
17 Whether or not the elephant is selected, is graphically visible
18 The image of the elephant cannot be moved to other locations on the tablet than on the map
19 The app does not allow that the child teaches words very rapidly
20 The app does not allow that the child rapidly moves the elephant image to different locations.
21 One half of the screen is used by the overview, the other half is used by the map
22 All distances and sizes are relative to the screen size
23 When a message needs to be send via internet, send it only once
24 Send and receive messages over the internet, as client

Tele-operation UI:

25 All desired audio fragments can be selected and sent via a button for each fragment.
26 All desired emotions can be selected and sent via a button for each emotion.
27 Word teaching for the child can be turned on and off, to prevent misuse

28 Send and receive messages over the internet, as client
29 All distances and sizes are relative to the screen size
30 When a message needs to be send via internet, send it only once

Robot face:

31 Show the face of the elephant
32 Show the learned words
33 Show a colorful explosion when a word is learned, on the place where the word appears on the screen
34 Play the correct audio sample that is being commanded via the tele-operation UI
35 Represent the elephant's emotion and (state of) learned words according to the state of the activity, during the activity. In other words, update correctly according to the incoming commands during the activity.
36 Do not allow word learning, when this option is turned off via the tele-operation UI.
37 All distances and sizes are relative to the screen size
38 Do not allow the same word to be learned more than one time
39 Do not allow more than one word to be learned on the same spot
40 When a word is forgotten and re-taught, make it appear again on the same spot as it originally appeared.
41 Forget words on a regular basis
42 When a message needs to be send via internet, send it only once
43 Send and receive messages over the internet, as server

5.1 The initial plan

Image 25 two surface-bots from the coBOTnity system

For the realization of the final, high-fidelity prototype, the initial plan was to use and expand the current coBOTnity surface-bot system. Two surface-bots can be seen on image 25. This system has several components. Firstly, there are some apps, made in Android Studio. For example there is an app for the face of the robot, and an app to control this face. Most important, the emotion of the face can be changed, audio fragments can be played, objects can be displayed, and the robot can be sent to another location. Secondly, there is the Zumo robot base with wheels (image 26). This allows the robot to move. It communicates via Bluetooth with the tablet on top. This can be controlled via the just mentioned app, or, thirdly, via the smart navigation system. This system works with a camera that hangs above the playground (image 27) and tracks the marker on the front of the robot. This data is sent to a computer that displays the playground with the up-to-date robot location. There is a controller app that also displays this and via this app the robot can be sent

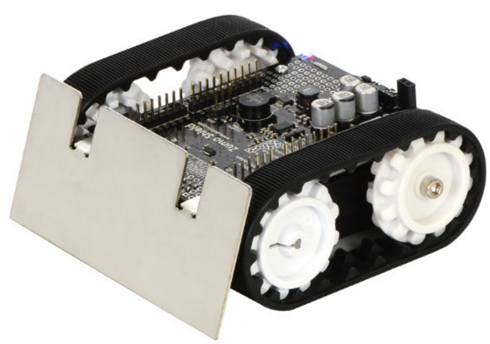


Image 26 Zumo robot base

to a new location. In both cases, using the robot control app and using the smart navigation system, the several components communicate via a robot operating system (ROS). In this system one device is the master, and other devices are clients. The clients can send information to the master and the master can send information to the clients. The robot face app is the master when the smart navigation is not used. When used, the computer is the master.

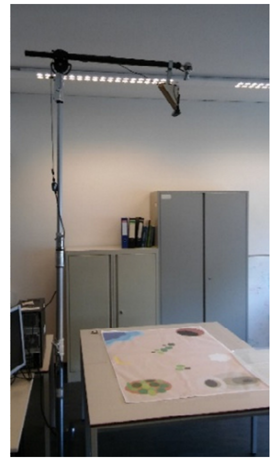


Image 27 Camera setup

As said, the initial plan was to use this system for this graduation project's final prototype. After all, the idea was developed for this system. However, in the end the existing system is not used. It could not simply be used, it had to be expanded or adjusted in many ways:

Child UI app

- Teach words to the robot. Once taught, an image of the learned word should appear on the robot face. So commands to get images on the screen had to be communicated.
- Send the robot to a new location using smart navigation. Robot navigation commands had to be generated and communicated.
- The robot location has to stay updated on the Child UI app, to show its current position.

Teleoperation app (will be discussed later, mainly needed for proper robot speech timing)

- Activate audio fragments. Communication with robot face app needed.

Robot face app

- At least eight objects should be able to show up on the screen, instead of four currently
- These objects should slowly fade away, instead of work via on/off
- Learning a word should be accompanied with a colorful animation

To realize the desired result, these new apps had to be made within Android Studio. They would partly integrate existing code, for example to activate audio fragments, and would partly need new code, for example to show eight objects instead of four. Development progressed very slowly because of a lack of experience in Android programming. Due to this there would be a high risk that the prototype would not work as desired during the user test. A prototype for the user test that represented the concept as imagined would be vital in order to draw valuable, scientifically correct conclusions. Because of this, it was decided to take a different approach to make the prototype.

5.2 The new plan

It was found that the Processing programming language could be imported as a library into Android Studio. Android and Processing are both closely related to the Java programming language. This way, Processing could be used to develop apps for the tablets. This approach was then used to realize a new system of three apps, specially made for this prototype, and not connected to the existing coBOTnity system. Using Processing to develop the apps and incorporating them into the existing system was also an option, but then still many of the above-mentioned issues would remain. They are mostly caused by not enough understanding of (the structure of) the coBOTnity system and not having the time to thoroughly figure it out.

The built system was purely a network of three Processing sketches. The Zumo robot base was not used, so the robot could not move itself. Implementation of this base would bring along several technical challenges. Autonomous movement was not strictly needed to test the concept so adding this base had low priority and eventually there was no time left to look into this. Other than this, no sacrifices with regards to the desired system were made by not using the coBOTnity system.

5.2.1 The child User Interface

This is the app that runs on the tablet that is controlled by the child. Via this app, the child would be able to teach words to the robot and to navigate the robot to a new location on the playground. The UI app sends to the robot app which word the child teaches to the elephant, so that the taught word is being displayed on the robot app. Note that the navigation part of the UI was developed, but not used, because the Zumo robot base was not used in the end. Two basic requirements apps were made, according to some of the requirements from the above list. The first one realized the teaching and robot moving mechanics, the second one realized internet communication between two apps. These apps were made to ensure the possibility to develop the desired system in this way. Possible fundamental problems would then be identified as early as possible. This prevents that a lot of work is being put in some feature that turns out to not function. Given that importing Processing in Android Studio is not a very common approach, it was worth the time to check if like this, for example internet communication between apps is even possible in a time-efficient way. All programming was first done in the Processing environment and each time a certain amount of progress was achieved, the code was copied to Android Studio. This was done because of familiarity with the Processing environment.

5.2.1.1 Basic functionalities implementation

In Processing a sketch was made with core functionalities of the user interface (Image 28). The red face represents the robot and the green square represents an object tile. The robot could be moved by dragging it to a different location. This represents the action of sending the robot to a different location on the table. A green dot could be dragged from the square to the robot. If released upon the robot, it would smile (Image 29). This represents the action of teaching the robot a word.

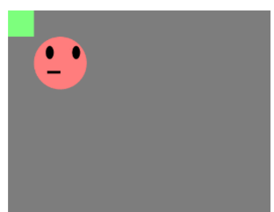


Image 28 Processing sketch

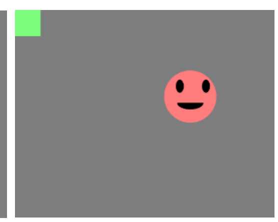


Image 29 Robot has moved and learned a

The technical requirements realized here:

44 The image representing the robot can be selected on the screen by clicking on it
45 By clicking somewhere else, this image changes location and is not selected anymore
46 After one time, this image can be selected again and the process can be repeated
47 A dot with the same color as the tile can be dragged anywhere from the tile
48 When not released on the image, the dot disappears and nothing happens
49 When released on the image, the image smiles, representing that something happens after releasing the dot.

Note that the interaction of clicking on the robot and clicking somewhere else was specifically chosen instead of dragging the robot. As was observed during the paper prototype user test, dragging the robot could make the children move the robot everywhere very quickly. This is undesired behaviour. With clicking, the robot will always move straight to a new location. Then still, the robot can be rapidly clicked to many places. In a later stage of the UI, this was prevented by adding a timer to the movement that would allow a movement only once each two seconds. This seems still a short period, but sometimes it happened during the paper prototype user test, that the child wanted to slightly correct the robot's position. For example, one centimeter closer to a location right after having moved him to that new location. A two second timer seemed the best period, after varying and testing several values. Obviously, also the interaction of dragging a word to the robot to teach him was specifically designed, but this has already been discussed in the conceptualization chapter.

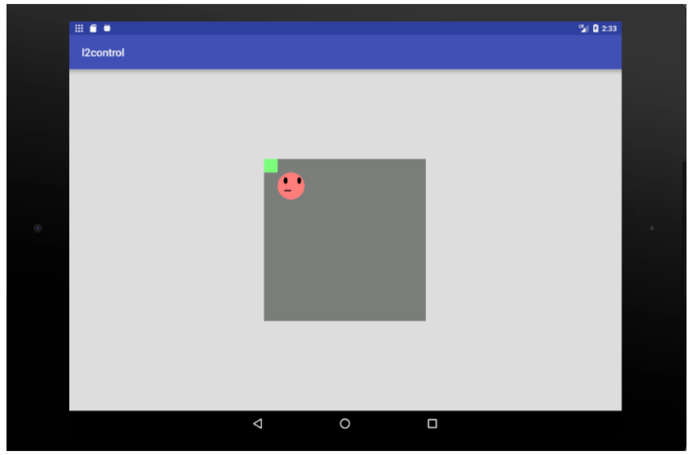


Image 30 Processing sketch in Android Studio system

This Processing sketch was then put in Android Studio, in order to test the possibilities of Processing integration in Android Studio (image 30). On the image, the app runs on an emulated tablet from Android Studio. This means a virtual tablet created by the software in order to test apps when a real tablet is not available.

A second Processing sketch was made to test communication over the internet between two apps. This prototype can be seen on image 31. It shows two separate Processing sketches. The right sketch is the client. The left one is the server. This client-server setup works the same as the master-client setup of the robot operating system. When the mouse was pressed in the client window, a white square and some text would appear on a random location on the server window. This particular setup makes use of 'localhost' as IP-address to which the clients should listen and thus does not really make use of the internet. It communicates internally on the laptop instead. The real challenge was to actually use two tablets, run one sketch on each tablet and communicate via the internet. A TP-link wireless router was used (image 32) to establish a local network. Distributed IP-addresses to the tablets were fixed. This was done to make sure that the server tablet would always get the same IP-address. Then this address could be hardcoded into the client's code as IP-address to which it should link. The two sketches were run as apps on the tablet and internet communication was successfully established. Next, a third

Image 31 Prototype to test internet communication

Image 32 TP-link wireless router

tablet was introduced as second client with similar code as the first client. Now two clients could send data to one server and this represented the desired system. Note that in the desired system also the server sends messages back to the clients. This was not tested on the tablets with these Processing sketches in android studio, only in Processing.

The following technical requirements were now met:

50 Internet communication between two clients and one server on tablets, making use of the Processing programming language.

51 A single mouse-pressed/touch event that leads to a single message sent by the client that leads to a single action executed by the server.

5.2.1.2 High-fidelity prototype

According to the (well-tested) design of the paper prototype, and building on the core functionalities prototype, the high-fidelity prototype of the Child UI was made. A first version can be seen on image 33.

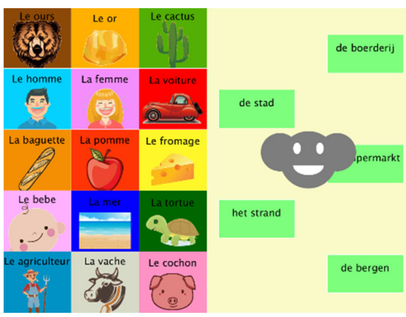


Image 33 First version

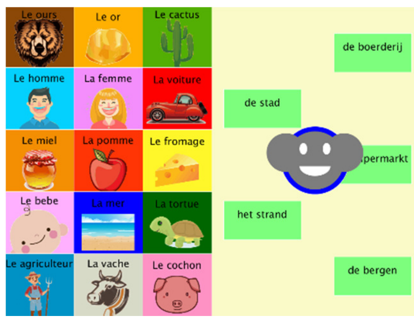


Image 34 Robot image is selected

Image 35 Robot has moved, and unselected

From left to right, the process of selecting and moving the elephant can be seen. Having selected the robot is visually made clear via the blue circle (image 34). Basic interface design principles teach that the state of an interface should always be clear. On image 35, the robot has moved-image has moved and is not selected anymore. Note that the elephant image and the green squares are placeholders. These were not immediately implemented because they are not strictly needed, contrary to the fifteen object images. Out of precaution that too many images could make the app crash, they

Image 36 Final version

Image 37 Cactus being dragged

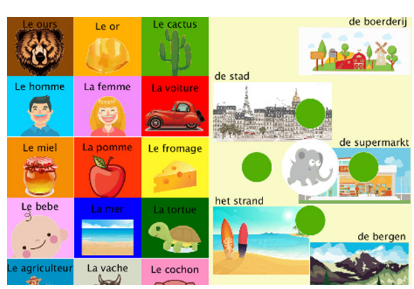


Image 38 green dot released on elephant

Image 39 blue dot released on elephant

were left away. Later this seemed no issue at all and they were added. This can be seen on image 36. Also note that on the left image 33, there is still the baguette, while on the others there is the honey. This is because the images were not taken at the same moment within the development process. Note that some of the fifteen images do not fully fit on the screen, which looks a bit sloppy. These screenshots are taken from the Processing variant. It turned out that on the tablets, the images fit well on their tiles.

On image 37, a green dot can be seen that is being dragged to the elephant, in order to teach him 'Le cactus'. For each tile, a dot of the same color as the tile can be dragged. On image 38, the dot has been released upon the elephant image. Teaching a word is followed by a simple and visually pleasant explosion animation of matching color. This can also be seen on image 39 for 'La mer'. The explosion was added for two reasons. Firstly, it visualized that a word has successfully been taught to the elephant. In the core functionalities prototype, this was visualized by the smile of the robot. As will be discussed later, sometimes it is not possible to teach the elephant a word. In such a case, releasing a dot on top of the elephant will not trigger the explosion and the dot will just disappear. Secondly, the explosion adds a fun factor to the teaching interaction. It makes the act a little rewarding, but not too much. Some big explosion all over the screen could make the child too enthusiastic about it. Then the child could want to quickly teach him more words in order to trigger more explosions. This would disturb the pace of the activity. The explosion is a reward, but kept simple so that triggering the explosion does not become a goal on its own.

Besides this, an internal timer of five seconds was implemented on the act of teaching a word. In other words, at maximum speed only once each five seconds a word could be taught. This still seems very fast but the period is reasonable. Firstly, when the child does actually abuse the teaching, the five-second break is enough time for the tele-operator to completely disable word teaching, as will be discussed later. Secondly, most audio fragments lasted about five seconds, so it would be reasonable for the child that no action can be taken if the robot is talking. Thirdly, making it longer could lead to frustration if a child does not understand why the mechanic does not work. This would need to be made visibly clear on the UI. However, this would result in an extra feature of the UI that would have to be understood by the child. This increases complexity and could lead to misunderstanding, which makes implementing it not worth it, given the low frequency at which this feature would actually be useful.

The locations were placed in such a way that they would somewhat represent France geographically. This is most clear with the big city (Paris) in the north-west, the mountains in the south-east and the beach in the south-west.

It was specifically chosen to use a simple dot as representation of a word when teaching it. Other options that were considered were to drag the image itself, or a dot with the image inside. A reason to use the image would be that the child sees the image more often and therefore might learn the French better. In addition, a link via the image might be clearer, or better remembered, than only via the color. These possible advantages were unlikely to be significant. In the first place, the child will drag the thing, so the finger is on the image, so the image is not fully visible. In the second place, the eyes will probably be focused on the destination of the interaction (the elephant) and not on the image itself. In the third place, drawing more images would burden the system more. No significant reason was found to not stay with the simple dot.

5.2.2 The Teleoperation User Interface

A teleoperation user interface (image 40) was made for the researcher to control certain parts of the robot during the activity, without the child knowing about this. No basic functionalities prototype was made this time because the app would only consist of buttons that would send a message via the internet to the server, that is, the robot face app. This mechanic was already successfully implemented in the Child UI and could be repeated.

Four different things can be teleoperated with the interface:

1. The robot’s speech
2. The robot’s emotions
3. The possibility to turn teaching on/off
4. The possibility to show certain numbers on the robot face app (discussed later)

Teleoperation was used because implementing these features to let the system execute them autonomously would take too much time. The robot would have to decide what to say next, based on choices of the child in the activity. For example if the child does not do any storytelling and only teaches words, the robot should sense this and say story-related sentences. In addition, it should know when some sentences are not applicable anymore. For example the pig may already be at the farm because of the child’s own story. The pig’s story sentence does then not make sense anymore. The camera tracking system would then be needed to track all the object’s locations. Also some sensor would be needed to know when the child tries to put more than one object on the elephant. Moreover, the robot would need to time his sentences in such a way that interaction is not disturbed. For example the ending sentence should not be started in the middle of some storytelling. All these things would require more technology and a lot more programming while a tele-operating human would be able to do them with minor effort. In the literature research it was found in [12] by Westlund and Breazeal that teleoperation is a common technique used by researchers to mimic autonomous, complex (social) behaviour of robots. Also from a time perspective, teleoperation is very efficient. It allows certain features to be experimented with before any time is spent on the actual implementation.

The only sentence that was programmed to be said automatically, was the ‘Oh, so X is Y in French. Thank you!’ after every time a word was (re)learned by the elephant. Doing this saved fifteen buttons on the Teleoperation UI for sentences that were often said, and on fixed moments. The buttons on the UI are structured as follows and have the following meaning:

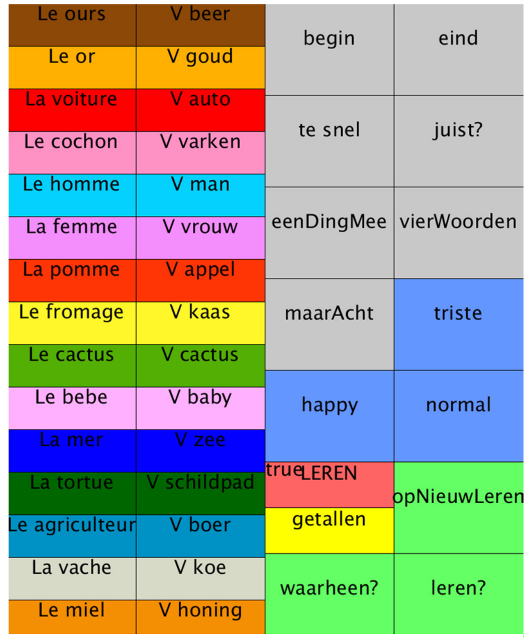


Image 40 Tele-operation UI

The first column on the left triggers the Dutch word, expressed in an enthusiastic/surprised way. For example if ‘Le fromage’ is pressed, the robot says ‘de kaas!’. These buttons have to be combined with the green buttons on the bottom right. The green buttons are buttons that are used often. For convenience they were put together on the bottom right. This way the tablet could be hold with one hand on the bottom right, to trigger these sentences, and the left hand could hover over all words on the left. With the ‘leren?’ (learn?) button, the elephant asks if she can learn a word. With the ‘opNieuwLeren’ (learnAgain) button, she asks to learn again, after having forgotten a word. With the ‘waarheen?’ (where to go?) button, she asks where to go now, which is less often used, but still frequently. In practice, each time a word on the left was pressed, and then either the learn or the learn again button. The words in the left column are in French for a technical reason that will be explained later.

The second column on the left triggers the story-related sentence for each word, for these sentences, see table 2. The V stands for ‘verhaal’ which means ‘story’ in Dutch. The grey buttons on the right are the sentences that are not, or only once, used per activity. For these sentences, also see table 2. They are the starting sentence (begin), the ending sentence (eind), being taught too fast (te snel), taught wrong word (juist?), carry only one thing (eenDingMee), already four words learned (vierWoorden) and can only learn eight words (maarAcht). These buttons are placed in the most inconvenient spot, because the tablet is being held on the bottom right. The blue buttons do not trigger a sentence, they change the elephant’s emotion instead. Because the word ‘sad’ was disliked by the researcher, the French ‘triste’ was used instead. Via the red ‘LEREN’ (learn) button, the child’s teaching can be turned on and off. In practice this is always turned on (recognizable by displaying ‘true’). This button can be used when the child starts to abuse the teaching mechanic and teaches words very rapidly, or words that have nothing to do with the story. The teaching can be made impossible, the face can be made sad and the elephant can say that she cannot learn words so fast or ask if that was the correct translation. When deactivated, also the explosion animation on the Child UI would not trigger. In practice the LEREN button was also used to quickly check if the internet communication was working properly, right before the start of the activity. The ‘true’ would show up as a result of a response message sent by the server. So when this true would show up, it was known that the client and server could successfully send and receive a message.

The final yellow button called 'getallen' (numbers) has no influence on the activity. It is purely for logging purposes and was added in a later stage, right before the final user test, when was concluded that this would be necessary. As will be explained later, a logging feature was added on the robot face. Each learned word received a counter that would count how many times the word was (re)learned. Via the getallen button, these counters could be displayed on the robot face. They would not be displayed during the activity and would be turned on afterwards, to take a photo of the screen and log the numbers in this way. In contrary to the LEREN button, no visual feedback on the UI was given whether the numbers were visible or not, but this could simply be seen on the robot face.

Two additional features were added to the UI in a later stage of the development process. The main purpose of these features was to improve activity progress management for the tele-operator. The improvements were identified during some final evaluations of the fully working high-fidelity prototype, some days before the user test. The first feature can be seen on image 41. The UI starts an internal timer once the app is started. After ten minutes, this red rectangle would appear in the top right of the screen, notifying the tele-operator that ten minutes have passed. This way it could be made sure that each session of the user test would last about ten minutes. Note that this implementation is not optimal since attention needs to be payed that the UI is started up once the children start, and not much earlier. Then the timer would already be running for a while. This feature was added in a little hurry but could be improved by making the timer start once the begin button is pressed instead of on start up.

Image 41 Red rectangle after ten minutes

A second feature that was added can be seen on image 42. On the left of this image, part of the robot face can be seen, which will be explained later. On the right, the teleoperation UI can be seen partly. Each time the elephant would (start to) forget a word, a black rectangle would show up on the word that is being forgotten. This way the tele-operator can still know which word was forgotten, after it has completely disappeared on the robot face app. This is essential in order to ask the right word to teach again.

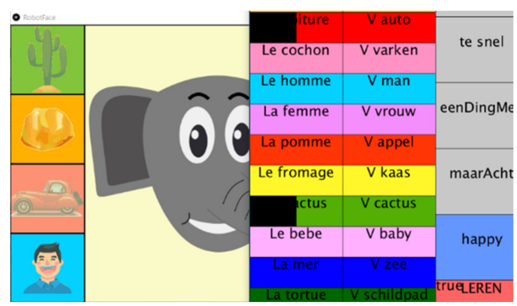


Image 42 Black rectangles for words that are forgotten

Finally, note that this UI was made mostly to be functional and not to be aesthetically pleasing. For that reason, certain aspects of the UI may not be following common UI design principles. The UI was made with focus on effective usability by the tele-operator, who is also the app developer in this case. If this app were to be made for a target user, it would be designed with greater care.

5.2.3 The robot face

The robot face app is for the tablet that is part of the robot. It represents the elephant. On start-up, this app looks like image 43. On each side of the face there are four spaces. This represents the progression system, as was also implemented in the paper prototype version. Eight words can be taught and each time a word is taught, it appears on the screen in the first empty slot. This goes from top left to bottom left, and then from top right to bottom right. On image 44 can be seen how the elephant has learned five words during the activity. Each time the elephant (re)learns a word, an explosion animation is played. This can be seen on image 45. The explosion is realized by a particle system and is of the same color as the learned word. Being a reward of successfully having taught a word (which is an important aspect of the activity), the explosion is made big and exaggerated, in order to feel rewarding.

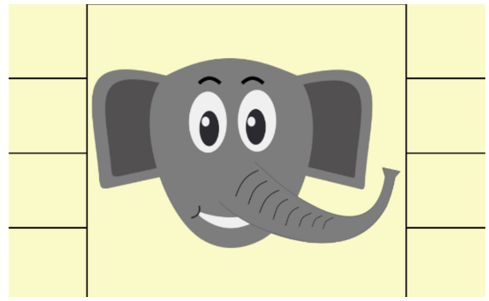


Image 43 Robot face at start-up

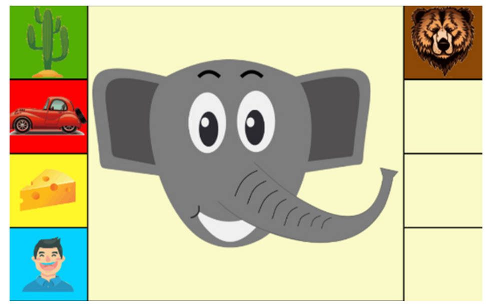


Image 44 Elephant has learned five words

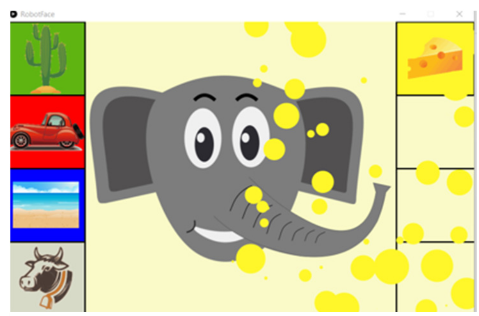


Image 45 Explosion after learning a word

The elephant can only learn a word when the child drags and drops a dot on his UI to the elephant image on that UI. A word can only be taught if not already eight other words have been taught. When the child tries this, the tele-operator can make the elephant sad and make her say the according sentence. Once a word is taught, it is not possible to teach another word on that same spot of the robot face. The word gets assigned to the spot. Also when a word is forgotten and the spot is completely blank, only the forgotten word can be taught there. This has been done to focus on these eight words during the activity, instead of all fifteen. In the literature research [16] it was found that a five-year-old needs six words to be repeated 10 times in twenty minutes, in order to remember them. Knowing this, it seemed reasonable to focus on eight words for an eight year old, but not on fifteen, given that each word should be repeated several times during the activity and there is a limited amount of time. In addition, the storytelling takes time as well. Those periods should not be interrupted too much with constant word learning and forgetting.

As shown earlier, forgetting a word is visibly shown by the word slowly fading away. In fact, the word does not fade away but a rectangle is slowly appearing on top of the image. On image 46, the car and the cactus are fading away. It takes about ten seconds for a word to completely disappear. The word can be re-taught to the elephant as soon as the fading starts. When a word is relearned, it is shown in full color again. For this, the rectangle on top would be made transparent again. The black rectangle on the teleoperation UI is gone too, as can be seen on image 47. Note that in the final version, only one word could be forgotten at the same time. First, each word would start to fade away independently of the other words' state. However, after about five words learned, words would start to fade simultaneously and there would always be some word fading away. In this situation, the child would have to constantly reteach the words and the storytelling would be heavily disturbed. It was also demotivating to see the robot face with several blank gaps where words have disappeared despite of all the teaching effort. This mechanic was changed so that one word can be forgotten at a time. All learned words have an individual timer, randomly between thirty and forty seconds. The word with the lowest value starts to fade away first. As soon as one word fades, it is made impossible for the other words to start fading. Once the forgotten word is retaught, all word's timers are reset and given a new random value. Again the word with the lowest time will start to fade, and so on. Besides this, it was added that the robot cannot forget a word anymore once the closing text of the elephant has been played. First the robot would keep forgetting words after the end of the activity, this reduced the rewarding feeling of completion significantly. Completion as a reward was described by Schell in [17].

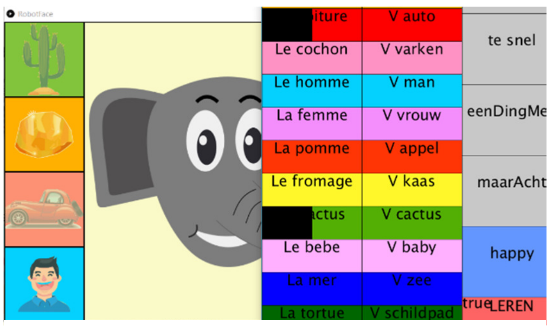


Image 46 forgetting two words

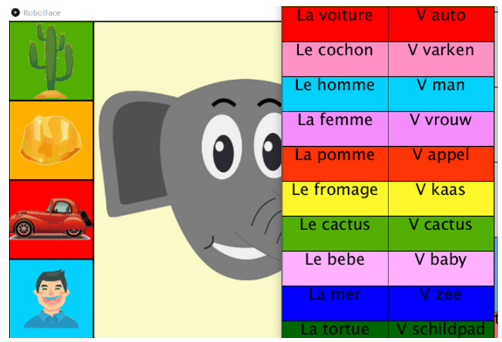


Image 47 words re-learned

The three different emotions of the elephant can be seen on image 48, 49 and 50. Normal, happy and sad respectively. These faces were made in Inkscape 0.92.2. To clearly distinguish between the emotions, and to keep the design simple at the same time, the ears, mouth, and eyebrows were varied with. Also tears were added to the sad face. Varying the trunk was experimented with but was kept the same in the end. It made the mouth less visible and/or looked not naturally. To make the elephant look cute, the eyes were made big and a cheek was added near the mouth.

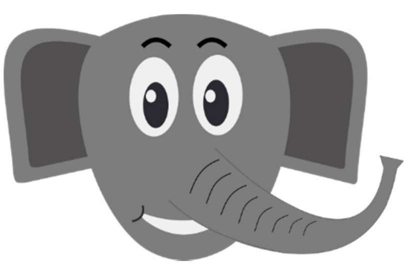


Image 48 'normal' emotion

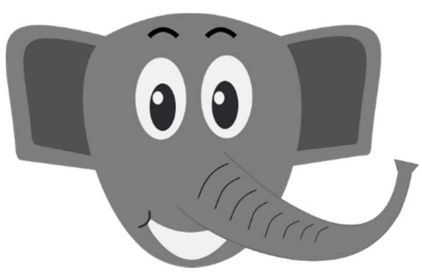


Image 49 'happy emotion'

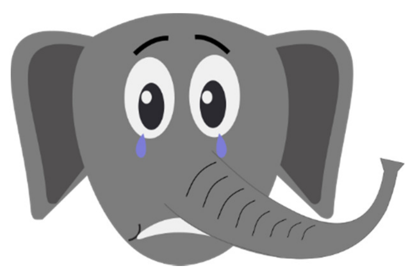


Image 50 'sad' emotion

The logging numbers, as discussed earlier, can be seen on image 51, when made visible. This is a typical example of how the robot face would look like when the activity has ended. Eight words are learned and many of them several times. Each single word can only be forgotten twice. In other words, the on-screen counter's maximum value is three. It can be learned for the first time, and then relearned twice. This maximum was added later. After experimenting with this mechanic, it became clear that the

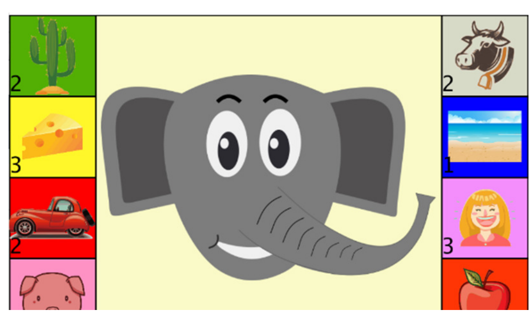


Image 51 logging numbers displayed

elephant constantly forgetting the same word would become annoying and unreasonable. It would also result in other words being repeated less often.

On image 52, the voice actor can be seen in action. With table 2 in front of her, she recorded all audio fragments separately, using an MP3 audio recorder app on the smartphone. This person was chosen because of several reasons. Firstly, she offered her help after getting to know more about this graduation project. Secondly, she has French family and can speak French very well. Thirdly, she acts on theatre, so she is familiar with role-playing and imagining to be the character. Fourthly, her voice was well-suited to frame the elephant as a peer of the child. In other words, her voice does not sound too adult-like. As was learned from the literature research [16], introducing the robot as peer could reduce anxiety and make sub-optimal interactions due to technical constraints more accepted by the child. Moreover, introducing the robot as peer who needs to learn the language could lead to learning by teaching.

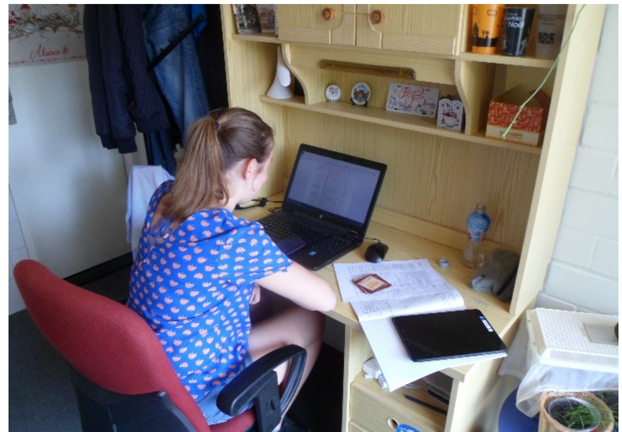


Image 52 The voice actor in action

5.2.4 The system of the three tablet apps

All development of the apps has been done in the Processing environment. Each time a certain amount of progression was made, or when a problem was solved, the code was copied into the respective Android Studio project. Not all code could simply be copied. Some parts were not supported by Android Studio and had to be figured out again. On image 53, the setup to work with the tablets and Android Studio can be seen.

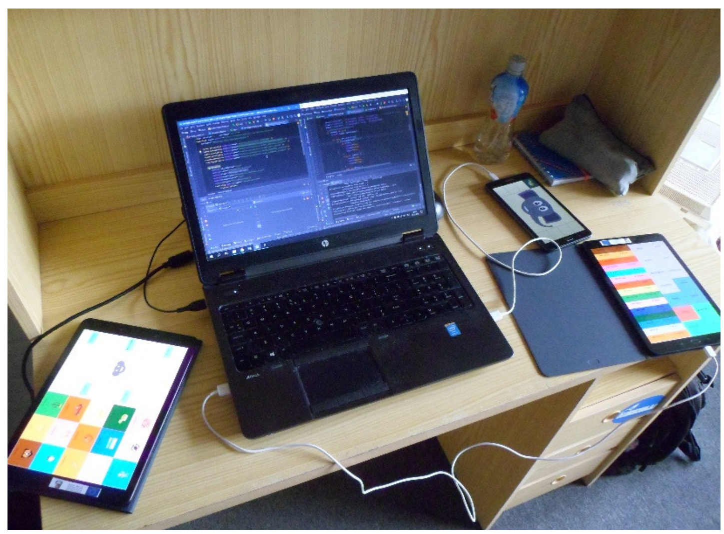


Image 53 working environment

An example of a Processing implementation that caused errors in Android Studio was the use of the Processing Sound library. It is used to play all audio fragments of the robot. This library was not supported in Android Studio. Other libraries were tried and also did not work. Eventually, the 'playSound' function of the coBOTnity system was used. This function makes use of the Android Mediaplayer library, a common way to handle media in Android. Another cause of problems was displaying the images on the tablets. In Processing, images can simply be put in the same folder as the sketch. This did not work in Android Studio. Instead, it was figured out that the images need to be put in a certain folder on the tablet, and that the file path to this location needs to be saved in a string. This string could then be put into the Processing loadImage function as usual. Besides this, the pixel density on the tablets is higher than on the laptop (on which the Processing sketches were made). Hardcoded values needed to be adjusted. For example the particle speed of the explosions and the text size. However, most values are coded with respect to the screen size, so this was a minor problem.

On image 54 can be seen how the tablets are being tested with the paper prototype. This has been done extensively to identify (minor) flaws in the system. This shows that the final prototype was realized based on many tests, iterations and improvements. For example the yellow 'numbers' button on the teleoperation UI is not yet present at this moment.

Image 54 Tablet testing with paper prototype

5.2.4.1 Technical explanation of the system

The different possible interactions with the system can be explained in terms of messages being sent and received via internet communication. On image 55, the first interaction with the system during an activity is displayed. This is the action of pushing the 'LEREN' button on the tele-operation UI that allows teaching for the child. When the button is pushed, the number forty is sent to the robot. This value can be explained. All buttons on the UI are individually defined in an array of Button objects (called allButtons). In a for-loop through this array, each individual button is available in the form of allButtons[i]. When a button is pressed, this 'i' is stored in a 'message' variable and sent to the robot on mouse release. The 'LEREN' button is number forty in the array. When the robot receives this forty, it executes 'learn = !learn' and sends 'learn' as message back to the clients. Each client executes 'learn = !learn' when 'learn' is received. On the tele-operation UI, 'true' is being displayed when learn is true. On the child UI nothing visibly represents the state of the 'learn' boolean.

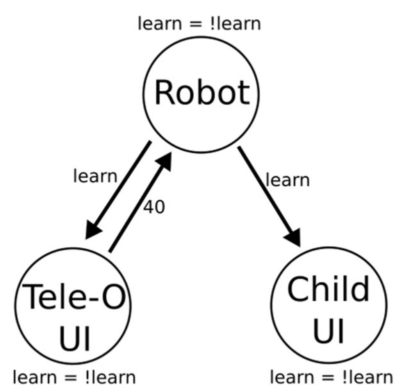


Image 55 turning 'learn' on

On images 56, 57 and 58, the actions of teaching a word, forgetting a word and re-teaching a word respectively can be seen. Image 59, 60 and 61 are put in as a reminder of how the apps look like. First a word needs to be taught. Once the child drops a colored dot of a word on the elephant (on the UI), the 'word' parameter of the specific Tile object is passed in a message and sent. Each of the fifteen squares on the left of the child UI is a Tile object with its own properties. The 'word' parameter is the text, the French word, on each Tile. On the robot app, if 'learn' is true, if at least one of the eight spots is still open, and if that word is not already learned, the elephant will learn the word. Among several actions that happen internally at that point, the 'tag' parameter is set equal to the incoming 'word' message. So now the specific learned word spot on the robot app carries a tag of the French word that is learned there. Other actions that happen at this point are the correct audio Sample being played (of having learned a word), an explosion being made at the correct location, the image showing up on the spot with colored background, Booleans being set on true and values being stored in variables for later use, just like 'tag'.

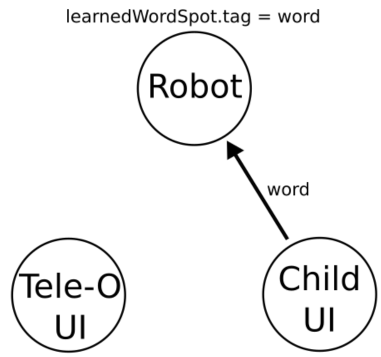


Image 56 teaching a word

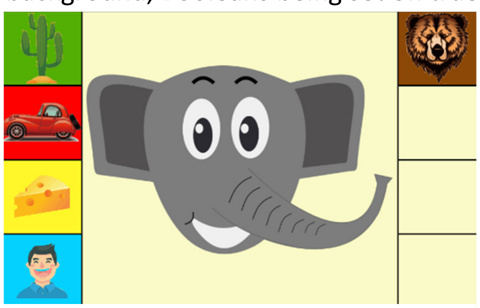


Image 59 robot face

Image 60 Tele-operation UI

Image 61 Child UI

For each spot of the learned words on the robot face, if the activity is not yet finished, if no other word is being forgotten, if the word is not already two times forgotten, if a word was learned on that spot, if the timer has expired, the word will start to forget. At that point, the tag of the word being forgotten will be sent by the server. On the tele-operation UI, for each button of the fifteen buttons of the left column, if the incoming tag message equals the button's 'word', its 'forgotten' Boolean will be set to true and the black rectangle will show up as a consequence. Note that a server sends its message to all clients, not to one client specifically. However, in this system there is often only one of the two clients that responds to a certain incoming message. Also note that with this, the reason for the words of the left column being in French is clear. In fact, the Dutch word is not known in the system. On the Child UI, the Tile object's 'word' parameter is the French word. This word as the parameter's value is the same that is being displayed on the tile. This 'word' is sent by the child UI and saved as 'tag' on the robot face when learned. This 'tag' is then being used to send back again to the clients when a word is being forgotten or

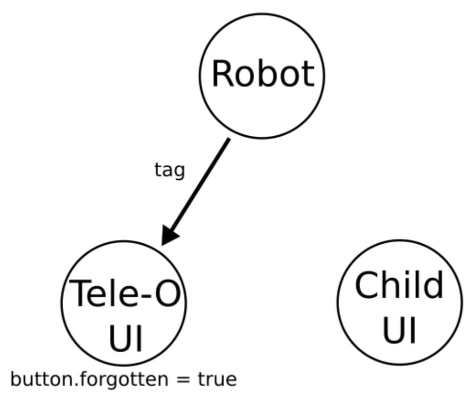


Image 57 send to tele-o UI which word is re-taught. On the tele-operation UI, the ‘word’ parameter of the Buttons is the Buttons’ text and this is being compared to this ‘tag’. Therefore, the ‘word’ on the Button had to be in French, just like the ‘tag’, in order to be comparable. In other words, the words that are being sent via internet, are the same words that are displayed on the tiles and buttons.

When the child re-teaches a word to the robot, the ‘word’ is again being sent. On the robot face, if this incoming 'word‘ is equal to the ‘tag’ of the word that is forgotten, the word is relearned by the robot. At that point, the tag is sent by the server. On the tele-operation UI, if a Button’s word is equal to the incoming ‘tag’ and its ‘forgotten’ Boolean was true, ‘forgotten’ is false again. The black rectangle disappears as a consequence. On the robot face app, the counter of the word that count how many times the word it taught, goes one up.

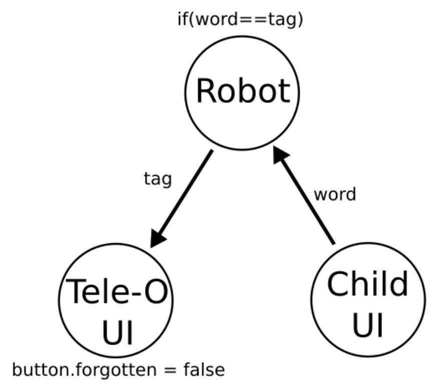


Image 58 Re-teaching a word

Finally, image 62 shows the general working of the tele-operation UI. In a for-loop through the allButtons array, it is being checked if a button is pressed. If that is the case, that Button’s position in the array is sent to the server. This is a unique number i for each button from 0 to 44. Two different things can happen on the robot face as a result of a button push on the tele-operation UI. Firstly, a Boolean can be toggled. There is a Boolean for each of the three emotions, to toggle the logging numbers, to mark that the closing text has played and to toggle the possibility to learn words. Secondly, an audio fragment can be played. In a file, all audio samples are saved in the same order as the buttons are put in the allButtons array. For example, button number five should trigger audio fragment number five in the list. This way, the sent number i can immediately be used in the audioSample array to play the corresponding fragment.

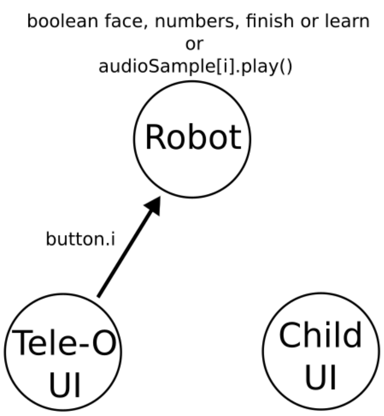


Image 62 Tele-o UI working

5.2.5 The playground, the tangibles and the final prototype

The playground and the tangibles were made. Each location was enlarged and printed on A3 size paper. All locations can be seen on image 63. To make them more robust, they were glued on one mm thick cardboard. This can be seen on image 64. The corners were strengthened with tape.

Image 63 all locations

Image 64 Close up locations

The tangibles were made by laser cutting eighteen (three spare parts) wooden tiles. Each tile is four mm thick and eight by eight cm large. The same sort of paper and cardboard as for the paper prototype's tangibles was used to put on top. The upper paper, the image and word, is six by six cm large. The tangibles can be seen on image 65. The tangibles were made this size to be not too large for the robot to carry them, but still be somewhat in proportion to the location's size.

Image 65 the tangibles

As can be seen on image 66, both sides of the tablet were masked with a piece of thin cardboard. This was done to mask the tablet's buttons to prevent that a child would accidentally push one during the test. By hiding the buttons and the tablet's logo, it looked less like a tablet. This may contribute to the child seeing the robot as a whole, instead of a tablet on a base. The robot's marker, previously needed for the tracking functionality, was masked as well with the same cardboard. On the image this is not visible because the

Image 66 masked logo and buttons

robot carries a car with him (It can be seen on image 67 however.). This feature is made clear with two small wooden blocks glued on both sides of the robot. As is done here, a tile can be put between these blocks. Note that the robot is sad because he forgot a word.

Image 1 Final setup

On image 67, the final setup as a whole can be seen. On image 68 can be seen how also the camera was tested, its position and sound quality in particular.

Image 68 camera testing

6. User test with the high-fidelity prototype

A user test was done in group 4 of a Dutch primary school. The twenty two children were eight years old on average. Testing happened systematically and rich data was collected. This data was extensively analyzed and discussed. The goal of this user test was primarily to proof the concept of the designed activity. This was determined by the learning behaviour of the children. It was also observed if the children had fun and if they interacted with the system as desired.

6.1 Preparations for the user test

Preparation for this final user test already started in the first month of the project with an informal visit to the researcher's former primary school. This school in particular was chosen to visit mostly because of familiarity with some of the school's teachers and positive memories. This visit had several goals. Firstly, to get a bit more familiar to young children. After not really having been in contact with young children for years, it was hard to estimate how children grow up over the years and what they are capable of at each age. Secondly, to discuss the assignment with an experienced teacher. At that point I had no idea yet what to do. Insights and tips from this experienced person could help a lot in choosing directions based on the target group. Thirdly, to learn more about the children's familiarity to technology, robots in particular. This was discussed with the teacher as well. In addition, the next day the children had their weekly technology lesson, organized by 'Digihelden' (see the reflection report in appendix 2). Digihelden is a small company. It is run by three people who are experienced primary school teachers, and who also know much about modern technologies. Each week, they visited the primary school and gave a special lesson, dedicated to technology, in each class. This was visited as well, the next day. Due to Digihelden, the children were surprisingly familiar to robots, programming concepts and other modern technologies. Fourthly, to informally ask around for a possible user test in the future.

The school showed interest and saw no problems in hosting the user test. Much later, the school was asked more formally, with a detailed e-mail about the assignment, if they wanted to participate. Their reaction was positive. Besides the good connections with the school, this school was also chosen because of the children's familiarity to robots. On this front, the school is ahead of many others. In the first place, this familiarity makes that the children better understand what a robot is. This mostly reduces possible anxiety and misuse. In the second place, this project's robot would fit well in the existing curriculum and the school could be seen as 'the school of the future' where technology education is integrated in several ways. According to the stated target group of this project, group 4 of the primary school was chosen to do the user test with. These children are mostly eight years old.

Ethical permission was requested to, and obtained from, the ethical committee of the faculty. Then an information flyer and consent form were made for all parents of group 4's children. These (Dutch) documents can be found in appendix 1. They were sent to the group's teacher who gave them to the children to take home. Several days later, already many of the forms had been signed and returned. In the meantime, regular contact with the group's teacher took place via e-mail. For example, to arrange a specific date for the test. In the end, 22 out of 24 children had their parent's consent to participate in the test. The other two also participated, but no data was recorded. The day before the user test, the school was visited. A location for the test was chosen, the school's gym. The activity was built up and showed to, explained to and approved by the teacher.

6.2 The user test

The main goal of this user test was to proof the concept. In other words, to proof that the children would learn some of the words. Besides learning, another part of this main goal was to observe if the children liked the activity, had fun, and if they interacted with the system as desired.

Image 69 user test setup

The whole user test was done in one day. Before the lessons started, the setup was built up again. The setup can be seen on image 69. The approach was as follows: first the children made a 'pre-session test' to test the Children's current knowledge of the French words. Then the activity took place. One day later, the children made an 'after-session test' to test their knowledge of the French words again. Children had to put their name on both tests, so the individual scores could later be analyzed per child and compared to the actions during the session. The pre- and after-test were identical with one exception. On the after-test, four additional questions were asked about how the children experienced the activity. They could fill these in for themselves. These questions were:

52 What did you like about the activity?
53 What did you not like or like less about the activity?
54 In your opinion, was the activity too hard, too easy, or exactly right? Why?
55 Were there any things you did not understand during the activity?

As said, 22 children participated in the user test, nine girls and thirteen boys, five aged seven, fifteen aged eight and two aged nine. All children participated in duos instead of alone. This was primarily chosen because 22 individual sessions could not be finished on one day. This was important to let all children do the after-session French test at the same time, one day later. This was more convenient and consistent than spreading the test out over two days.

A list of all materials used:

56 5 location plates
57 15 object tangibles
58 1 robot, grey base + tablet
59 1 child UI tablet
60 1 tele-operation UI tablet
61 1 Sony handyman camera + tripod
62 1 TP-link wireless router
63 1 smartphone (to take a photo of the logging numbers on the robot screen after each session)
64 1 A4 paper (to note per child, the session number, name, gender, age and role (discussed later))
65 24 pre-session tests
66 22 after-session tests

In the pre-test, the two children without consent also participated. Their data was not used. In the after-test however, both these two children were not present that day, so they did not participate.

As was learned from Westlund et al. in [18], during the literature research, it is important to introduce the robot in the class before the test. As said earlier, this may decrease possible anxiety of the children for the robot. This was started with on the day of the user test, alongside a very short introduction to the activity. Then all children made the pre-session French test. The test had to be made by drawing a line from the French word to the Dutch word-image combination. A typical pre-test can be seen on image 70. Note that the name field has been cut off for privacy reasons. Also note that five distraction words have been added to the French words list. These are Le arbre, Le autobus, Le loup, La table and Le velo. It was made clear that the result of this test would not count for the children in any way and that it was not strange at all if no words were known. This pre-test was held to filter out the pre-knowledge of the children and the guesses. As will appear later, the cognate words 'the baby' and 'the cactus' were often known/guessed correctly in the pre-test.

Image 70 pre-test

The duos for the activity were made by the teacher, rather arbitrarily. Not fully randomly however, for example the two children without consent were put together and paricipated first. During each session, first the names, gender and age were noted. Then the children were asked to take place in front of the Child UI tablet. The children could at no moment see what the researcher was doing on his tablet. Everything the children had to know was then explained: the setting (France) of the adventure, the objects, their freedom of choice to go and do whatever they wanted, how to move the robot, how to carry and move an object and how to teach a word. Then the children were asked to choose one of two roles. One role was to control the tablet (and teach the words), the other role was to move the robot and the tangibles. This choice was noted down. Then the camera was turned on and the children were told that the robot should soon start automatically. They were also told that the researcher would be taking notes on his tablet and that no attention should be paid to him.

Several screenshots of the camera's recordings during the sessions can be seen below.

Image 71

Image 72

Image 73

Image 74

Image 75

Image 76

On image 71, the children are listening to the robot. On image 72, the child puts the honey on the robot and discussed with the child behind the tablet. On image 73, the children are busy with exchanging the turtle for the bear. On image 74, the child teaches the robot a word and the other child watches. On image 75, the robot is being moved from the beach to the city. On image 76, the child swaps the honey and the car.

Twenty of the children participated in the morning, four in the afternoon. After each session, the children were thanked and were asked to send the next duo. In the meantime, the logging numbers were turned on and a photo was taken. Then the setup was prepared for the next session. The next day, the after-session French test was held in the class. Again, it was made clear to the children that their results would not count in any way for them. A typical after-test can be seen on image 77 and 78. After the test, they were again thanked a lot and they were informed about the tele-operation.

Image 77 after test

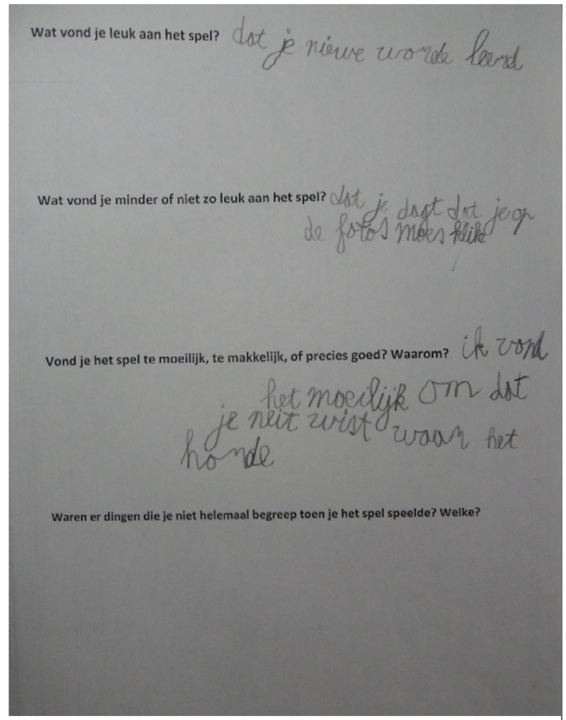


Image 78 after-test open questions

6.3 Results

In this section all user test results are shown and discussed in specific sub-sections. The available data consisted of twenty two pre-tests, twenty two after-tests and eleven screenshots of the robot face with the logged number of times a word was taught to the robot. Data analysis was done in Excel 2013. While the main goal of this user test was to proof the concept (to proof that the children would learn something), other more specific hypotheses were also made, based on the expected influence of several design choices on the learning behaviour. This was done because of the availability of the data to analyze this. First these hypotheses were stated, then the results were shown and discussed per sub-section. These sub-sections were made based on possible ways to divide the data in groups to compare. For example, the results of the tablet-role children versus the results of the robot-role children. Next, the answers to the open questions of the children are discussed as well as the researcher's own observations. Finally, conclusions are drawn and suggestions to improve the prototype are given.

6.3.1 Hypotheses

Several hypotheses were made before the test.

1. Children score better on the after test than on the pre-test
2. Cognate words are more often learned than non-cognate words
3. Tablet children learn more words than robot children
4. Many of the learned words are taught to the robot during the activity
5. Often-repeated words during the activity are learned more often than less-repeated words
6. Tablet children learned more words that are not taught to the robot than the robot children
7. No significant difference is expected between boys' versus girls' learning behaviour.

6.3.2 General results

In table 3, the most general data can be seen.

67 Pre score is the number of correct answers per child on the pre-test

68 After score is the number of correct answers per child on the after-test.

69 Pre lines is the amount of lines drawn per child on the pre-test

70 After lines is the amount of lines drawn per child on the after test.

	Pre score	After scor	Pre lines	after lines
Average	1,68	3,73	4,86	7,82
SD	1,29	2,47	3,56	4,08
Percentag	34,58%	47,67%		

Table 3 general results in numbers

Subtracting the pre score from the after score gives an average of 2,05 more words that have been done correctly on the after-test compared to the pre-test. It can be said with 99% confidence that the children learned words. A confidence interval of (1,01;3,09) can be made which indicates with 99% confidence that the children learned in between 1,01 and 3,09 words on average. Regarding the drawn lines, it can be said with 99% confidence that on average more lines were drawn on the after-test than on the pre-test. A 99% confidence interval shows a range between 0,62 and 5,30 more lines drawn on average. The accuracy of the drawn lines has increased from 34,58% to 47,67%. So not only more lines were drawn, there were also more correct lines. The accuracy is calculated by dividing the lines by the score.

On image 79, an overview can be seen of the amount of words correct on the pre and after test, per child. The amount of words learned is the difference between the two columns, that is red – blue. It can be seen that most children have learned at least one word. Only four children had equal scores before and after doing the activity. No children scored worse on the after-test. One child (19) scored very well, later was learned that this child has an Italian mother and speaks Italian at home. This fact probably has greatly contributed to this achievement. French and Italian may be partly cognate. The effect of cognate words will be discussed later.

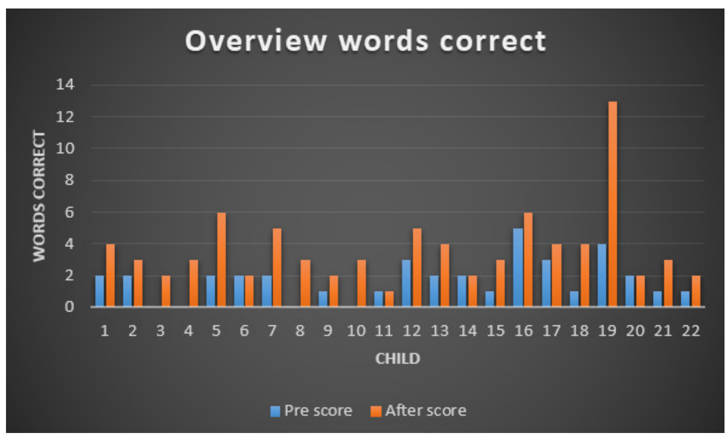


Image 79 general results on a graph

6.3.3 Results based on gender and role division

The children can be divided in sub-groups based on three different characteristics that were noted during the user test: the age, the gender and the role during the activity. Additional conclusions can possibly be drawn based on a more detailed look at the results per characteristic. It was decided to not draw conclusions based on age because the relative age difference within this group of children is very small. 15 out of 22 children are eight years old. 5 children are seven years old. Besides this small number, many of these 5 would turn eight within two months after the experiment took place. 2 children are nine years old. Besides the small number, one can argue that these children are nine in this group because they probably had to redo a class because they showed below-average school results.

This makes it very unlikely that they represent nine-year-olds in general in terms of school-related activities.

In table 4, the results based on gender can be seen. 'M' stands for man, 'V' stands for woman. No significant difference between the boys and the girls can be concluded based on the results. As can be seen, the difference between words learned (after score – pre score) is small, as is the difference between the increases in drawn lines. Note that also no specific difference was expected, the data was only collected out of interest and to possibly find unexpected differences. A 95% confidence interval for the difference in correct words pre and after, per gender, would be (-1,13;2,41). Since the left value is negative, no positive difference can be guaranteed at 95% confidence. This means that it cannot be guaranteed that either boys or girls will always score better than the other. It stays negative for 80% confidence level as well and lower confidence levels would start to lose their significance so these were not investigated. What can be seen however, is that the standard deviation of the after score for the boys is much higher than for the girls. This means a bigger variation in the amount of words learned for the boys. The data confirms this. On image 79, the children that learned above-average amounts of words were mostly boys. For example children 5, 7, 10 and 19. The girls scored more equally.

	Pre score	After scor	Pre lines	After lines
mavg	1,92	4,23	5,08	8,23
vavg	1,33	3,00	4,56	7,22
mSD	1,44	3,09	3,68	4,34
vSD	1,00	0,87	3,57	3,83
mPerc	37,88%	51,40%		
vPerc	29,27%	41,54%		

Table 4 results per gender

More interesting results can be found when looking at the role that the child had during the activity in table 5. As said, there were two roles, the tablet role ('t') and the robot role ('r'). It can be seen that the tablet children had more words correct on the pre and after test than the robot children. It can be said with 95% confidence that the tablet children had a higher score on the pre-test. An F-test on equal population variances gives 90% confidence that equal variances can be assumed. This actually needs to be the case in order to apply the T-test on the difference of two sample means. It can be said with only 50% confidence that the tablet children scored better on the after-test. In addition, equal variances cannot be assumed with 90% confidence, so this statement is significantly weaker. Presumably, this is caused by the high standard deviation of the tablet children on the after-test.

	Pre score	After scor	Pre lines	After lines
ravg	1,09	3,09	3,45	7,64
tavg	2,27	4,36	6,27	8,00
rSD	0,94	1,45	2,73	3,93
tSD	1,35	3,14	3,85	4,40
rPerc	31,58%	40,48%		
tPerc	36,23%	54,55%		

Table 5 results per role

Surprisingly, the tablet children also scored higher on the pre-test. The children were allowed to choose themselves what role they would do during the activity. This seems to have caused that the children with higher pre-test results chose for the tablet role during the activity. The higher scores on the pre-test can have several reasons. Firstly, one might say that these children are generally the more intelligent ones. Because of this, they could be more confident with the more difficult-looking task of controlling the tablet, and therefor chose it. Another explanation is that these children were more enthusiastic about the activity before it had taken place. They might have paid more attention during the pre-test and thus scored better. It is reasonable to say that paying more attention leads to better results given that le bebe and le cactus can easily be guessed correctly, when noticed. Then they might have wanted to do the (arguably) 'cooler' task of controlling the tablet and quickly claimed it in front of the child that cared less.

Surprisingly, it can be seen that both the robot and the tablet children have learned about two words on average. It was expected that the tablet children would learn more, because they had the tablet in front of them and interacted with it. On the tablet, all French words and an image of the meaning were displayed nota bene. A significant difference can be seen in increase of lines drawn on the pre- versus the after-test. The robot children show more growth in amount of lines drawn than the tablet children, can be said with 80% confidence. An F-test for equal population variances was executed and it can be said with 90% confidence that equal variances can be assumed. This big growth in amount of lines can be explained in several ways. Firstly, after having done the activity, the robot children could have become more enthusiastic about the activity and then participated in the after-test more actively. Secondly, the robot children could have become more convinced of their French knowledge and drawn more lines. The tablet children showed less growth because they were probably already enthusiastic/confident. Note that drawing more lines does not per se mean that more words were done correct. As can be seen on the percentages on image X, both groups show an increase in line accuracy, but this growth is bigger for the tablet children. One could argue that the robot children somehow thought to know more words than they actually did. This may be explained when examining the data more detailed.

In short, it can be said that an average of two words was learned. However, the reason for this learning cannot be explained in terms of gender or role. It cannot be convincingly said that boys scored better than girls or the other way around. It can be said however, that boys showed a larger spread of amount of words learned. It cannot be said that children with the tablet role learned more than children with the robot role. It can be said that children with the tablet role showed higher scores absolutely, but not relatively. The correlation between the role and higher scores can possibly be explained in terms of initial enthusiasm, intelligence, or a combination of both. The robot children drew significantly more lines on the after-test than on the pre-test. This growth could be explained in terms of increased enthusiasm and/or confidence after having done the activity. The increase in drawn lines by the robot children is only slightly visible in line accuracy increase. The tablet children show a bigger increase in line accuracy. The robot children possibly thought to know more words than they actually did.

6.3.4 Results based on taught words

The data can be examined in greater detail by only looking at the words that were taught to the robot by the children. These eight words are called 'selection' from now on. It is expected that a major proportion of the words that the children learned, was part of this selection during their session. This is expected because these words have been taught to the robot. This means that the tablet child probably has paid attention actively to this word while teaching it. Additionally, the word has actually been used by the children during the activity, in the elephant's adventure. For these reasons, it is expected that both the robot children and tablet children have learned these words more, and that the tablet children learned an even bigger proportion than the robot children since they controlled the tablet and did the teaching.

Now when looking at the data, all the above hypotheses are not confirmed. As can be seen in table 6, an average of about one more word was done correctly on the after-test. That is, 'S after score' minus 'S pre score'. 'S' stands for 'selection' here. Given that this selection is eight out of fifteen words and that a total average of two words was learned, the fact that one word of this selection was learned does not indicate that a big proportion of the learned words comes out of this selection. As can be seen, there is also no significant difference in the learned words between the tablet role and the robot role.

	S pre scor	S after sco	S pre lines	S after lines
Average	1,18	2,32	2,77	4,55
SD	1,05	1,91	2,22	2,50
Percentag	42,62%	51,00%		
ravg	0,91	1,91	1,82	4,36
tavg	1,45	2,73	3,73	4,73
rSD	0,94	1,64	1,89	2,54
tSD	1,13	2,15	2,20	2,57
rPerc	50,00%	43,75%		
tPerc	39,02%	57,69%		

Table 6 results for the selection

Something interesting can be seen when looking at the line accuracies. On image 80, the pre- and after-test line accuracies can be seen per role, for all words and for the selection. It can be seen that the line accuracy for the robot children actually decreased for the words in the selection, compared to the pre-test. This decrease, however, is mostly caused by the very high pre-test accuracy and not because of a low after-test accuracy. This can be said because the after-test accuracy is about equal for the selection and the words. The high pre-test accuracy means that the few lines the children drew, were actually accurate lines. One might say that the robot children made the test more securely, only drawing lines to words they really knew, or that were obvious guesses. Why this only happened to the selection words, may be coincidence. It cannot be explained at this point at least.

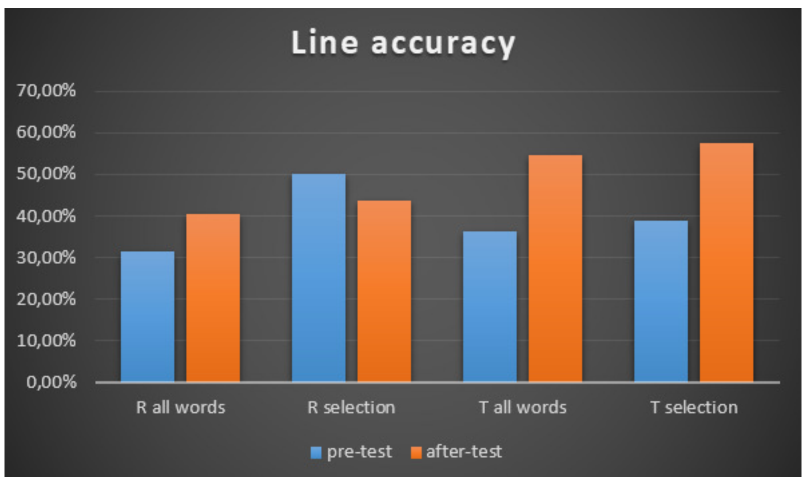


Image 80 Line accuracy

The earlier-mentioned suggestion that the robot children possibly thought to know more words than they actually did, on the after-test, may be true, but is not caused by the design of the activity. If that was the case, the line accuracy for the selection words would arguably have been lower than for all words. The robot children were interacting with the robot. They heard the French and Dutch words it was saying, but maybe did not pay full attention to this while playing. Also the difference in French pronunciation and writing could play a role here. In this case the children could vaguely remember what they had heard, but not being sure of the answer. This should have let to a lower line-accuracy for the selection than for all words. This is not the case, however.

When looking at the tablet children the increase in line accuracy is about equal between all words and the selection. This counters the above hypothesis even further. The tablet children did not only not learn the selection words better, they also did not draw more accurate lines to those words compared to the whole.

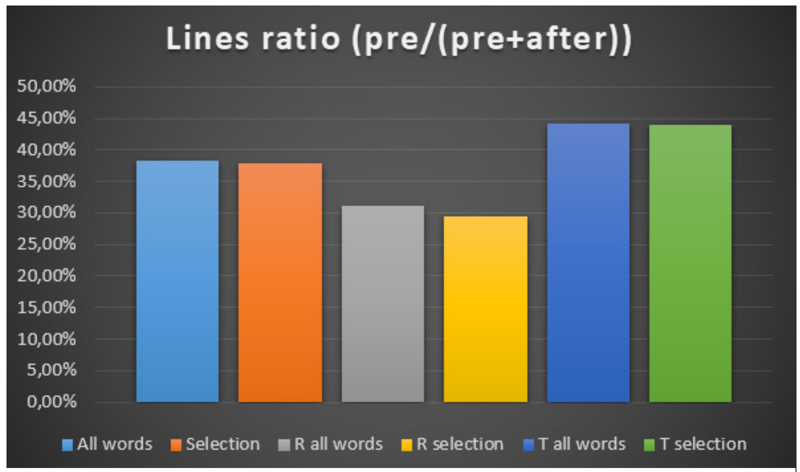


Image 81 Lines ratio

On image 81, the lines ratios can be seen. This means the amount of lines drawn on the pre-test, divided by the total amount of lines drawn on the pre- and after- test, on average, per child. In other words, the lower this value, the bigger the relative increase in lines drawn on the after-test. According to the above hypothesis, this lines ratio should be lower for the selection. This would mean that a bigger increase in drawn lines has occurred for the selection than for the whole. This would mean that the children have gotten more confident about the words in the selection and at least thought they knew more of them. As can be seen, this is not the case for all children together, as well as for the robot and tablet children separately. This further counters above hypotheses.

In short, in terms of selection versus all words, no significant differences can be spotted. The hypothesis that many of the learned words would be in the selection is not true. Also the hypothesis that this proportion would be even bigger for the tablet children is not true. Even no significant differences in line-draw behaviour are spotted, except for the unexplainable high line accuracy for the robot children on the pre-test.

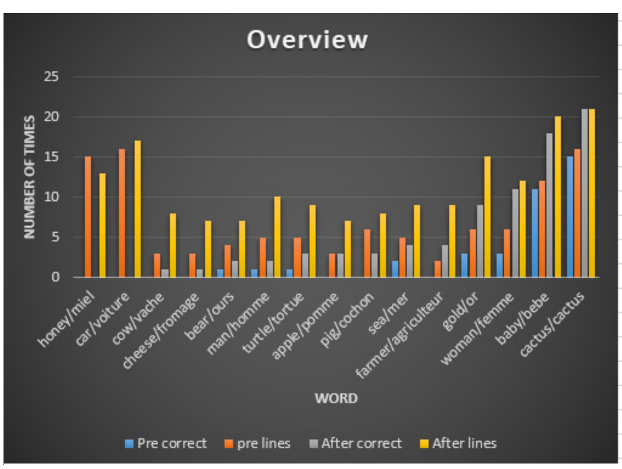


Image 82 overview per word

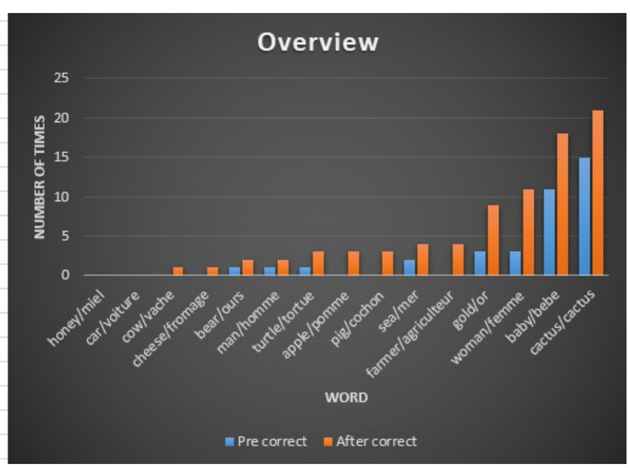


Image 83 Overview per word, correct words

6.3.5 Results based on individual words

The data was also investigated per word, instead of per child. Two overviews can be seen on image 82 and 83. On image 83, the drawn lines are not visible, in contrast to image 82. On both images the words are ordered by ‘after correct’. In this case, a ‘line’ means that a line is drawn to the particular image-Dutch word tile on the paper. Most interesting to notice on image 82 is that le miel and la voiture both were never done correct by any child even though many lines were drawn to these image-words. Here the influence of cognate-looking words is probably visible. The Dutch word for honey (honing) and the French word for man (homme) are somewhat similar in spelling. This is also true for the Dutch auto (car) and the French autobus (bus). A link between these words was made very often. In addition, auto and la voiture as well as honing and le miel are totally different from eachother. Other common mistakes were varken (pig) linked to la vache and man linked to la mer, likely for the same reason of being spelled somewhat similarly. It seems that the children based their answers partly on whether or not two words are cognate-looking.

More interesting than the lines is the words done correctly. This can more clearly be seen on image 83. In terms of times learned (after minus pre), the words can roughly be divided in three groups. The four most-learned words (or until cactus), the seven medium learned words (ours until agriculteur), and the four least-learned words (miel until fromage). Based on the findings about cognate words in the literature research [13], it is very likely that this is the reason that le bebe and le cactus are learned so many times. This is probably also the reason that they were done correctly so many times on the pre-test already. However, for the other thirteen words, a reason for this specific spread has to be found somewhere else.

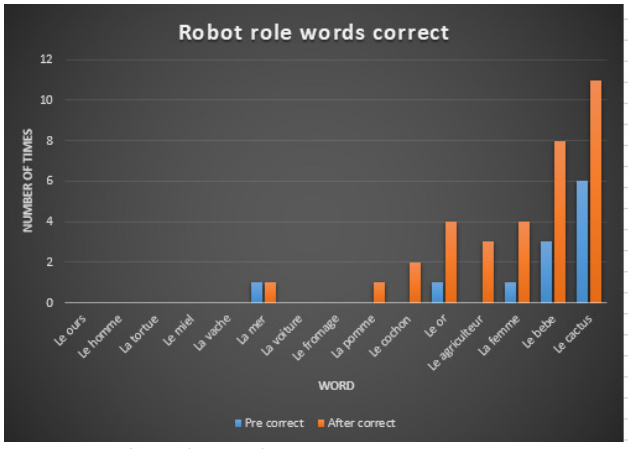


Image 84 Robot role, words correct

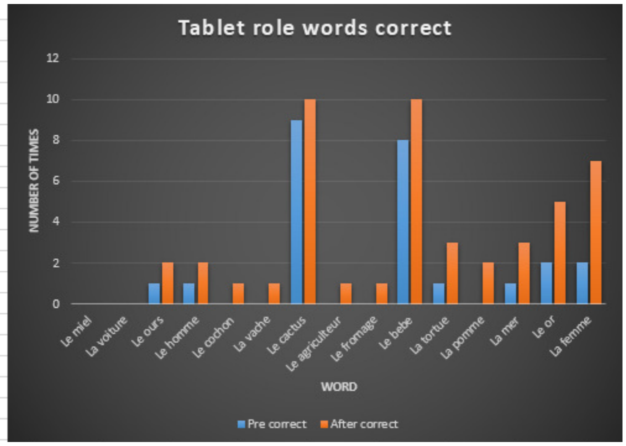


Image 85, tablet role, words correct

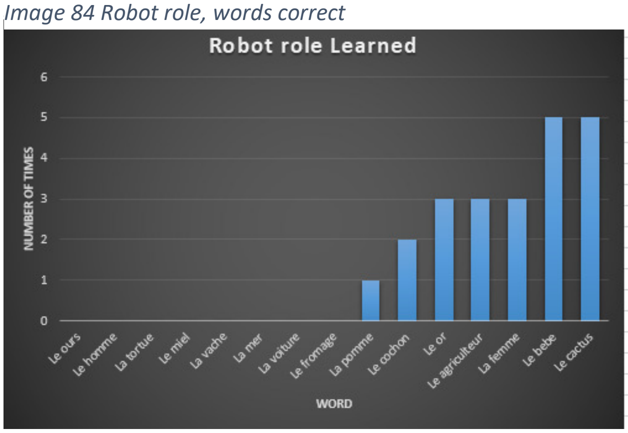


Image 86 robot role, words learned

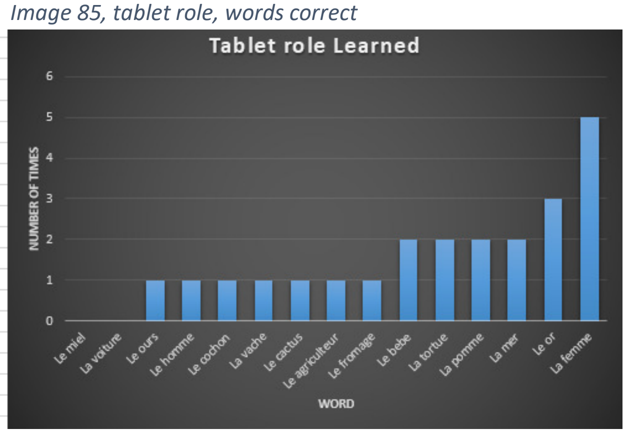


Image 87 tablet role, words learned

The amount of words correct and learned can be divided per role. This can be seen on image 84, 85, 86 and 87. Note that the words are not in the same order on the x-axis for the tablet and robot role because they are both ordered by words learned, per role. Now on image 86 and 87, a clear difference can be seen in what words were learned and how many times, per role. The tablet children show a bigger spread of learned words than the robot children. The robot children mostly learned le bebe and le cactus, while many of the tablet children had these words already correct on the pre-test, as can be seen on image 85. La femme and le or are amongst the most learned words for both roles. The reason for this different spread could be the roles themselves. The tablet children could have scored a bigger variety of words because they saw the words on the tablet and taught it to the robot. While the roles do not seem to have effect on the amount of words learned, they could still have effect on what words were learned.

To learn more about above spread, table 7 and image 88 were made. In table 7, 'W' is wrong and 'G' is good. Meaning done wrong or good on the after-test, if done wrong on the pre-test. In other words, was the word learned or not. The numbers are the amount of times a word was taught to the robot during the activity. These values are a summation of all words for all children. In other words, of all words that were not taught to the robot, 122 were not learned and 21 were learned. Of course, hypothetically, the amount of W3 would be smaller than the amount of G3. This would mean that words that were taught three times were actually done correct more often. This is not the case. In fact, as can be seen on image 88, the ratio G/(G+W) is for all four different cases more or less equal. This means that the amount of times that a word was taught has no significant impact on the learning of the children.

	0	**1**	**2**	**3**
W	122	38	39	49
G	21	8	7	10

Table 7 number of times good versus wrong on after test per number of appearances in the activity

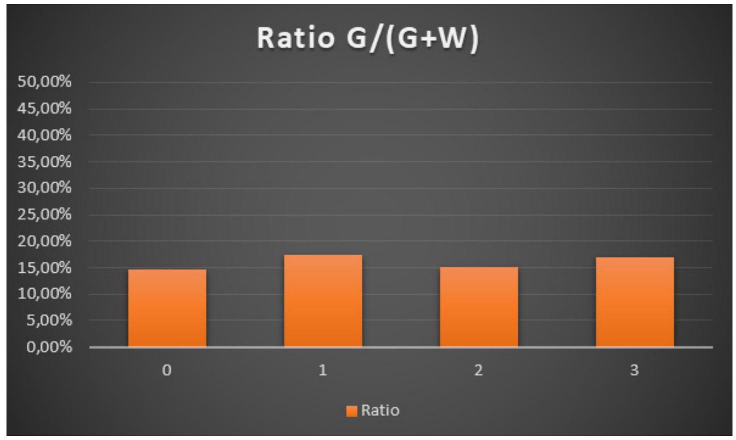


Image 88 good-wrong ratio per number of appearances in the activity

These numbers are for all children together. Per role, likely the most significant value would be the G0. These are the words that are done correctly without being taught to the robot. In other words, it is very unlikely that children in the robot role have seen or heard the French translation of these words during the activity. In contrast, the tablet children have been looking to all words on the tablet and could potentially remember a word even though they did not teach it. The hypothesis would be that most of the G0's were caused by tablet children. Counting the G0's per role gave 11 G0's for the robot role and 10 for the tablet role. This is of course against this hypothesis. Because G0 would be the clearest indicator in favor or against this hypothesis, and because other values would become insignificantly small, data analysis stopped here.

It can be seen that some words were learned more often than others, and that the learned words differ per role. However, these results cannot be explained with the design of the activity. Whether or not a word was taught to the robot, and how many times, seems to have no significant influence on what words were learned by what role. Other factors could be the reason for the learning of specific words. As earlier said, whether or not a word is cognate has an influence. Children seem to have based their answers often on how cognate-looking two words are. Other than that, some words simply seem to have caught children's attention more than others. For example le or and la femme. It might be that children recognized these words from their everyday life. There are certain brands and Dutch expressions that include French words, they are quite rare however. One might even argue that the position of the words on the test has played a role. As can be seen on image 89, 'or' and 'femme' are both very closely positioned to their translation. This may have caused that children remembered the translation more easily, because they saw the words next to each other. It could also contribute to 'fromage' being done correct only once, for example. However, there are more words close to their translation, like 'le ours', 'la pomme' and 'le cochon', that were not done correctly many times.

Image 89 after-test

A final thing to mention is the communication between the children outside of the activity. This could be a reason that children learned words that they did not see during the activity. They could enthusiastically tell to others what they learned, or test other's knowledge. Important to mention is that, right before the after-test, several children were saying to each other things like "I still know what baby is in French, le bebe!". It is very well possible that communication like this has influenced the knowledge of the children on the after test. One could say that the activity achieved its goal of teaching words by making the children enthusiastic about it and talk about it to each other. This may have been an important part of the learning.

6.3.6 Children's opinion and own observations

Besides the test results, empirical data was also collected in two ways. Firstly, the children had to answer four open questions on the after-test about how they experienced the activity. Secondly, the researcher observed the sessions and took notes, and looked back the camera recordings.

6.3.6.1 Children's opinion

The children's open questions were:

1. What did you like about the activity?
2. What did you not like or like less about the activity?
3. In your opinion, was the activity too hard, too easy, or exactly right? Why?
4. Were there any things you did not understand during the activity?

In terms of the open questions, the children's ability to read and write was slightly overestimated. Especially question three and four returned almost no valuable answers. For question three, most children did not answer the 'why?' part. The most given answer was 'exactly right, because I can learn something'. Some children said it was too easy or too difficult but did not support this answer. For question four, no useful answers and explanations were given. Children answered for example with specific words like 'the cow' or 'yes', or 'no'. Also no answer at all was given regularly.

For question one, many different things were mentioned by the children. The most common answer was 'everything'. More specific answers were 'learning new words', 'teaching the elephant new words', 'solving the tile puzzle', 'working with the robot', 'the French pronunciation by the robot', 'the elephant' and 'the robot'.

For question two, some things were mentioned as well. The most common answer was 'nothing'. More specific answers were 'that you thought that you had to control the elephant via the tablet', 'teaching the elephant again each time', 'dragging the circle to the elephant' and 'French'. Note that all these single more specific answers for both questions were given only once or twice.

Based on the above, you could say that the opinions of the children were positive overall. Besides the most common answers of 'everything' and 'nothing', many different positive things about the activity were mentioned by the children, and only a few negative things. However, one can doubt an eight-year-old's ability to reflect on the activity and fully understand and answer the questions.

6.3.6.2 Own observations

In general it can be said that most of the children had fun while doing the activity. They seemed amused by the whole. A smile was often triggered when the robot started to talk "autonomously". Especially the French pronunciation of the robot made the children laugh or smile often. Then the children also said (some of the) French words themselves, either before or after teaching it to the robot. Communication often took place between the two children in the form of discussing where to go and what object to let the robot carry with him. It was clear to the children how to interact with the system. Teaching the robot and moving the robot and tiles was done correctly after the explanation. Interaction with the system went more or less the same for all sessions. Not a single duo interpreted the activity significantly different from the others. Many different words in different orders were taught to the robot. There was no trend in what words were being taught.

The clear affordance that the robot could carry an object, combined with the explanation that this was possible, changed the way the children did the activity. Many children perceived the activity now as a puzzle where the objects needed to be replaced to their correct location. Moving tiles to other locations was done constantly and was a major form of interaction with the system. In several cases, (almost) all tiles were moved to another location at the end of the session. The children perceived the robot's sentences as hints to where an object needed to go. Interesting to note is that also the locations were seen as possible hints. The clearest example is the car. On the image of the city are some small red cars displayed. This was seen as hint that the car would need to go there. While this puzzle-like experience was not intended, it had a positive effect on the course of the activity. It was a clear goal and challenge for the children to find out where all objects needed to go. Storytelling still took place in the form of children talking to the robot/objects and reasoning why objects had to be moved to another location. One could very well argue that the puzzle-solving enhanced the storytelling, because reasoning for the puzzle was based on logical story-like relations between the objects. In other words, children came up with story-elements in order to justify their reasoning for the puzzle. A story plot can be seen as a puzzle that is being solved via progression in the story. When someone takes a book, he is curious about the plot, has many questions, and does not understand the events. While reading, this 'puzzle' becomes clearer. Somewhat in this way, a story and a puzzle were merged rather unintentionally in the activity. This became clear when many children interpreted the activity in this way.

The re-teaching mechanic of the activity worked well. Many children noticed the sad face of the elephant and/or the fading away of a word before the elephant started to talk about this. It was not told to the children that the elephant could forget a word. When this first happened, many children were a little surprised, not clearly positively or negatively. Some children then immediately started to notice the forgetting themselves, others after some more times. Children said things like "Quick, quick, quick, otherwise he will cry" or "He forgot something" or "He has to cry!".

A design flaw was identified during the user test. The end of the activity was unclear. It was not told to the children that the activity would take about ten minutes. As was actually intended, many children thought that the activity had ended once they had taught eight words to the robot. Looking to the tele-operator they sometimes said things like "there is no more space, what should we do now?" or "now we did everything". After not getting a response, they would then continue to play but often less motivated and/or clueless. Note that also many children continued after the eight words and did not think that this was the end. Either way, the actual end of the activity was often not clear. The children understood that the robot said some end-like text, but besides this one text, there was no other (visual) indicator that the activity had ended. They often only fully understood this after the tele-operator said that it was over.

Three things were often tried or assumed by the children. Firstly, several children tried to move the elephant via the right part of their interface. The function of this part of the UI was not mentioned, only the teaching was explained. On the one hand, it is good that the children understood the reason of this UI functionality without any explanation. On the other hand, it had somewhat bad effects since the children tried to drag the elephant which does not work. It needs to be selected instead. When tried to drag, the elephant moves only a small distance because it is deselected (and thus moves) as soon as the finger goes off the elephant image. While this hassle caused only minimal frustrations, it should have been explained before the activity. When the child continued trying to drag the elephant, the tele-operator said that this was not needed in order to move the robot.

Secondly, many children tried to teach the robot a ninth word when the robot had forgotten one of the eight words. They thought that a new word could be learned on the blank spot where the robot just had forgotten a word. The robot sentence that she could only remember eight different words made this clear to the children. The initial attempt was there nevertheless, as well as the minimal confusion when this did not work. Thirdly, children sometimes assumed that the robot had to carry an object in order to learn it, or in order to say something about it. Often the other child of the duo then said to his/her partner that this was not needed. It became also clear via the system itself because sometimes the robot would start to talk about an object that was not being carried. In addition, when a word was forgotten, the robot would start about this word where ever he was. Even though these two things were often assumed by the children, one could say that this does not necessarily require a change in the design. The system was able to make clear to the children that the behaviour was not needed or allowed.

Important to mention is that the children generally waited until the robot would say something, especially in the beginning. They let themselves be guided by the robot's hints. When the activity progressed, this waiting reduced and free play took over more. The children still listened to the robot's text but interpreted this more freely. Realism was then more neglected. For example in the beginning they would bring the baby perfectly to his parents in the city, but later these parents would travel to another location and the bear and an apple would come to the city in return. The puzzle aspect could have played a role here too. When exchanging objects, it is not always possible to exchange in such a way that both objects are on the right location after the exchange. The tele-operator's behaviour could also have influenced this waiting. Sometimes a robot sentence was played rather quickly. As a result, the children had not much time to think for themselves. It also occurred that the children's reasoning got interrupted by the robot's sentence. The tele-operator could have been a little slower with the robot-sentences, in order to trigger more storytelling by the children.

Finally, an external factor that may have influenced the amount of fun a child had, was the other child of the duo. Some duos seemed to be (good) friends, other duos seemed not to be friends. Amongst friends, there was generally much communication and collaboration. They had fun in doing the activity really together. Doing it together contributed positively to the amount of fun they had. On the other hand, some duos barely communicated and interacted with the system rather separated. One child moved the robot and tangibles to whatever place (s)he liked while the other child did some teaching every now and then. In this specific case, the tablet child was enthusiastic about the activity, but this decreased because the robot child did not care to cooperate much. Doing it together contributed negatively to the amount of fun in this case.

6.3.7 Conclusions and further improvements

Based on all results, conclusions can be drawn. Based on these, suggestions can be given to further improve the prototype.

It can be said that an average of 2,05 words was learned by the children of this user test. It can be said with confidence that the children learned in between 1,01 and 3,09 words. Also more lines were drawn on the after-test than on the pre-test, and more accurate lines. However, the reason for this learning cannot be explained in terms of gender or role. It cannot be convincingly said that boys scored better than girls or the other way around. Only the boys showed a larger variety in amount of words learned. It cannot be said that children with the tablet role learned more than children with the robot role. It can be said that children with the tablet role showed higher scores absolutely, but not relatively. The correlation between the role and higher scores can possibly be explained in terms of initial enthusiasm, intelligence, or a combination of both. The robot children drew significantly more lines on the after-test than on the pre-test. This growth could be explained in terms of increased enthusiasm and/or confidence after having done the activity. The increase in drawn lines by the robot children is only slightly visible in line accuracy increase. The tablet children show a bigger increase in line accuracy. The robot children possibly thought to know more words than they actually did, after the activity.

Looking at the taught words versus all the words could give more insight in this. However, in terms of selection versus all words, no significant differences can be spotted whatsoever. The robot children and tablet children learned both about one word on average from the selection. One out of eight is not a bigger proportion than two out of fifteen, so no more words were learned from the selection. Equal line behaviour for all words versus the selection on the after-test shows that the children also did not think to know more about the words in the selection.

When looking at the individual words, it can be seen that some words were learned more often than others, and that the learned words differ per role. However, these results cannot be explained with the design of the activity. Whether or not a word was taught to the robot, and how many times, seems to have no significant influence on what words were learned by what role. Besides the likely difference in intelligence and/or enthusiasm between the roles, other factors are likely the reason for this specific learning behaviour. For example, cognate words are learned more easily than non-cognate words. In addition, communication between the children about the words after the activity could have played a role, as well as the distribution of the words on the test and frequency of exposure in daily life.

On the one hand, words were learned. On the other hand, the learning was not significantly influenced by the design of the activity. The hypothesis that children would score better on the after test than on the pre-test is true. The hypothesis that cognate words are easier to learn than non-cognate words is true. The hypothesis that the tablet children would learn more than the robot children is false. The hypothesis that many of the learned words were taught to the robot during the activity is false. The hypothesis that often-repeated words are learned more often is false. The hypothesis that the tablet children would know more words that were not taught to the robot is false.

From the observations and open questions, it was clear that the children generally liked the activity and had fun. Except for the unintentional puzzle aspect and slightly vague ending, interaction with the system went as desired and was consistent among the many sessions. It can be concluded that a fun, understandable and educational experience was designed. It can also be concluded that some design choices failed to make a significant difference in terms of educational benefit, and arguably also were less fun. Mostly three aspects are meant here: the re-teaching of the words, making the limited selection of taught words, and restricting the children to a specific role. Further research will have to be conducted in order to convincingly justify that above aspects are less important. This will be discussed later.

It can be argued that the re-teaching was experienced as less fun. A couple of children wrote this down at the after-test questions. Additionally, it was observed during the sessions that this feature did not significantly cause more pleasure and/or excitement for the children in general. Also the restrictions of the eight words and the roles might be unnecessary. On the one hand, the eight words were a clear end goal. On the other hand, they conflicted with the artificial end of playing ten minutes. Of course, one could argue which way of ending the activity is preferred. Doing them both (like now) was no success. Removing (some of) these restrictions or making them less strict allows for new, interesting possibilities.

For example, what would happen if it was actually possible to teach a new word once one of the eight words is forgotten? What would happen if the colors would appear somewhere on the elephant herself, and there was no grid. What would happen if these changes were combined with a longer play time of twenty minutes? What would happen if nothing was told to the children about any task division? The way of interaction with the system can be changed drastically by simply playing with some restrictions.

It was concluded that external factors likely have played a role in the learning behaviour. Of course, also this will have to be confirmed by further research, as will be discussed later. Arguably an important factor for learning was communication between the students after doing the activity. This was likely caused by the fun they had, enthusiasm about the activity and the fact that they (thought that they) learned something of it. In other words, learning in a fun way triggers communication, which triggers even more learning. Now creating a fun activity may be at least as important as creating an educational one. The current design was for a substantial part focused on creating an educational activity. Design choices were made to get repetition of words, actively teach words to the robot, and the robot's goal was to learn words. This was all very obviously about word-learning. Even children discovered that it was about word learning for themselves, while this was never mentioned. This user-test's results suggest that a less educationally-focused activity may be at least as effective. Assuming, for now, that the external factor of child-child communication indeed plays a significant role for learning, this could be used in the design. The design of the activity could be focused more on the fun aspect, in order to trigger this communication.

Now the idea of the puzzle comes into play. This idea already somewhat proved itself to have a positive impact on the activity, in terms of engagement, child goals and challenge. The system could be designed to support this behaviour more. A puzzle could be designed in which all words need to be moved to another location, in order to reach some goal. The robot's texts could indeed be the hints for the puzzle. Even the locations could be used as hints. As was experienced many times during the design process of this activity, the story plays a major role in how the activity is perceived. The current story of having an adventure in France may not fit well to this puzzle-like design. Instead, some detective or a mystery might fit better. Many creative design options are possible here and this would require extensive brainstorming. Now only the rather obvious choices are mentioned.

Instead of redesigning a major part of the activity, existing parts could be improved. The current set of design choices is largely based on literature research. For example the learning by teaching, the limited amount of words a child can remember, and the importance of repetition. All these aspects were proven to have impact in other research, but show minor impact here. A more detailed look could be taken to existing literature on how these results were achieved. New, more detailed insights could be gained on how to successfully implement the above features so that they become more impactful.

Continuing on this, possibly a very effective way to improve the learning, is to take another approach to the learning by teaching aspect. An extensive learning by teaching method was developed by J.P. Martin [19]. In this method, the children do not only learn by teaching, but also learn something for themselves, to teach it to someone else later. In other words, the child knows that it will have to teach the subject to another child and prepares a lesson to do so. The child learns the subject, not with the goal of learning it just for itself, but with the goal of teaching it to someone else. In some way this could be realized in the activity. Currently, the children do not know beforehand that they will have to teach. The activity could be changed so that the children have to learn the words beforehand, in order to teach/use them in the activity later. A child could be very motivated to learn the words, knowing that it will need to teach them in the (fun) activity that follows. Of course, a feature like this needs to be carefully designed. For example, a design flaw that could occur easily is that more intelligent children learn more words beforehand and then achieve better in the activity.

If child-child communication indeed is an important factor for learning, this could be supported by the design, possibly in combination with the puzzle. Firstly, this could be supported during the interaction with the system. Secondly, the activity could be extended with a part focused on knowledge exchange between the children, after interaction with the system. For the first case, collaboration could be made a necessity in order to reach the goal. In the current design, this is not needed to reach the goal. As said, this leaded to some cases where children had less fun because of too little team work and communication. For the second case, the activity's goal could be extended to go beyond interaction with the system. The children would need each other's information in order to reach the goal. This could very well be combined with the just described change of the learning by teaching aspect. Also here, note that these are the obvious answers. A solution that is actually creative and elegant requires extensive brainstorming.

7. Conclusion

Conclusions on the project will be drawn by answering the sub-questions and the main question of the project. First the sub-questions will be answered. These answers will then lead to an answer to the main question. Note that specific design choices were made during this project. Some have shown to be good choices in this specific context. This does not mean that these are the best choices in general however, or that other choices will show worse results. The choices made contribute to this specific activity as a whole, but may not work equally well in other designs.

Sub-question 1: *How to design the storytelling part of the activity?*
The storytelling aspect was designed in such a way that the child could freely decide what to do, but within a given set of locations and objects. A certain balance was found between given story structure and player freedom. The basic ingredients for the story were given. The setting of the story was set, as well as possible objects to include in the story and the goal of the main character. The player was free to achieve this goal in his own way, making use of the objects and the setting. Even though it was not strictly necessary to use given objects and stay in the given setting, the fact of having them strongly suggested a certain limited set of options for the storytelling.

The story structure was designed to support the language learning part of the activity. It triggered behaviour that would contribute to achieving the story goal, as well as achieving the activity's educational goal via learning by teaching. The story was linked to language learning in the following way. Since the activity is about learning French, a logical setting for the story to take place is France. The presence of the French language is justified in this way. The robot was chosen as main character of the story, to set the focus on the robot. This character was chosen to be not French itself and visit France. This partly justifies the need to learn French. Only partly because one can visit a country without learning the language. To fully justify this, the main character's goal was set to having an adventure in France and wanting to learn some French as well. Now the robot, being the focus of the children, had the goal of learning French. The given story objects were set as desired words to learn for the main character. These objects were chosen based on possible value in a story and on difficulty of the French word. With the first, it is meant that firstly, the objects on their own would be interesting to use in a story, and secondly that the objects would relate to each other. This to trigger cause-effect reasoning in the story. This contributes to experiencing a word as a concept in its context of use, instead of just as a word.

To build upon the robot's characteristic of being a physical entity, the described setting and objects were made physical as well. This was put on a table top, to allow for overview and possible communication between multiple users of the activity. The robot was able to move objects to other locations, in order to link the physical world to the story. This could be done by putting an object tile on top of the robot and moving the robot to another location.

Sub-question 2: *How to design the learning by teaching aspect in the activity?*
Learning by teaching behaviour was realized by closely relating it to story aspects. The robot as main character was framed as peer of the children, who is in need of help during the story. The children were able to give this help in the form of teaching French words to the robot. Children could freely teach words to the robot, and the robot would frequently ask to teach a specific word based on the story

progress. A difference in French knowledge between the robot and the child was realized via the interface of the child. On this interface, an overview of the available story objects was visible, with their French translation. The interaction of teaching a word to the elephant was fully created via the interface. A colored dot, representing a word, could be dragged from the overview to the elephant image on the tablet. When a word was taught successfully, a rewarding explosion would appear on the robot as well as on the UI. The empathy for the robot by the child was used to trigger interactions based on a combination of robot speech and emotions. Next to teaching, also re-teaching words that the robot forgot, was implemented, in order to repeat words several times. Words were forgotten regularly, but not so many times that it would drastically disturb the story.

The activity was made for either one or two children. In case of two, each child would get one of two roles. One role was to use the child UI, the other role was to move the robot that could not move by itself, and the tangibles. So one child actually did the learning by teaching while the other had more of a general helping role. This role division was however mostly done out of practical reasons. In the activity, the children were framed as being on the adventure together with the robot, to increase engagement and stress the peer feeling.

Sub-question 3: *What are the required technical capabilities of the robot to fit its role as main character as well as teachable agent, and how to implement these into the system?*
Social features were given to the robot in order to communicate with the children for various reasons. The robot has three different emotions (normal, happy, sad) and can say a variety of texts. The texts can be divided in three groups based on purpose. There are texts that help the child with storytelling, texts that evoke learning by teaching, and texts that make clear to the child what behaviour is not possible or not desired. With these social features, the robot was able to fit its role as teachable main character. The robot was not able to autonomously decide what to say and what emotion to have. Tele-operation was added to the system in order to realize a prototype of the social robot that would be close to the desired system, in a short amount of time.

The high-fi prototype was realized in the Processing programming language that was imported in Android Studio as a library. This allowed for Android app development using Processing. The system consists of three apps that communicate with each other via a local network created by a router. There is the app that is the robot face, the app for the child UI, and the app for the tele-operator UI. The robot face app is the server and the two UI's can send and receive messages to and from the server, as clients. The robot face app displays the face of the main character, with a grid of the learned words on the sides. The child UI app displays the overview of available objects to teach, and a map of the playground with the robot on it. The tele-operation UI consists of buttons, each triggering a specific audio fragment or emotion. The tele-operator chose what the robot said with what emotion. The child taught the words to the robot. Which words were taught, and if a word was forgotten, was displayed on the robot face app. The technical requirement that the robot could be moved by commanding to go to a new location via the child UI was not realized due to time constraints.

Sub-question 4: *How to set up, execute and evaluate the experiment?*
The user test was setup and executed according to the following procedure. Dutch Children of group 4 (mostly eight-year-olds) of the researcher's former primary school were chosen to test the prototype

with. There were already good connections with this school and it has state of the art technology education. One group, instead of more, was chosen for practical reasons. Ethical permission and the parent's consent was obtained. The activity was showed to the teacher beforehand. The robot was introduced to the children in the class. The children made a pre-test to measure their pre-knowledge on the French words that would be in the activity and to filter out the guess-factor. For practical reasons, the children were put in duos for the activity. The activity took place outside the classroom, in the gymnastics room. At the start of each session, the main character's goal was told, and the possible interactions with the system were explained. The children chose themselves which of two roles they wanted to have. Gender, age, session number and role were noted per child. The sessions were recorded on camera. After each session, it was logged what words were taught to the robot and how many times. Each session, the children played ten minutes with the system. The researcher was the tele-operator. The next day, the children made an after-test. This test was exactly the same as the pre-test. In addition, some open questions were asked about the children's opinion on the activity. The children were informed afterwards that tele-operation had taken place.

The user test was evaluated in two ways. Firstly, data analysis of the pre- and after-tests has taken place. Secondly, the sessions were observed and watched back on the camera's recordings. Also the children's answers to the open questions were taken into account here. The data analysis focused on the achieved results in terms of word learning. Via data analysis, it was tried to explain specific learning behaviour with the design choices made. The observations and questions focused on the user experience and ways of interaction with the system.

Sub-question 5: *What influence do age and gender of the children have on the activity?*
Age and gender were noted for each child, however, no conclusions can be drawn about the influence of age. In addition, no differences were spotted between genders in the data analysis as well as in the observations. Even though some were seven or nine, the children's relative age differences were very small. Such a research on age would have to be done among several classes, instead of one, where children's developmental stage is not so similar. The focus of this user test was not to identify gender and age differences, it was only a possible interesting field. To state any convincing conclusions about this, more specific research would have to be conducted, where these factors are the only variable.

Main question: *How to design a storytelling and learning by teaching activity with a social robot as main character and as teachable agent, for primary school children to learn French as second or third language?*

It can be concluded that a fun, understandable and educative activity was designed. Design choices were made based on literature research, game design principles and some common sense. The storytelling part of the activity was designed to support the learning by teaching part of the activity. It triggered behaviour that would contribute to achieving the story goal, as well as achieving the activity's educational goal, via learning by teaching. To build upon the robot's characteristic of being a physical entity, the story world was presented on a tabletop with tangible objects that could be included in the story. With this set of given options, the child was free to decide how to achieve the main character's goal of going on an adventure to France and learning some French words. The robot, being social, was able to talk and show different emotions. The robot's speech and emotions were mostly used to give story-related hints about relations between objects, and to trigger learning by teaching. The child was able to teach the French translations of the tangibles to the robot via an interface on a tablet. To further enhance learning by teaching, the robot was framed as the child's peer, who is in need of help in terms of the French language. A knowledge difference between the child and the robot was created by presenting all available French words on the child's UI. A network of three tablet apps was made in the Processing programming language. The child's UI, the robot's face, and a tele-operation UI. Tele-operation was used to realize convincing interactive behaviour for the social robot. The robot's speech and emotion was controlled in this way. Images of all learned words were shown on the robot's face app. A user test was done on children of group 4 of a primary school, mostly eight-year-olds. User test data analysis concluded that the children had learned 2 words on average. However, the reason for specific learning behaviour could not convincingly be explained in terms of the activity's design choices. Different ways to possibly improve the activity were found with the user test.

8. Discussion
A variety of aspects of the project will be discussed in no particular order of importance.

According to the definition of storytelling that was presented in the introduction ("*the interactive art of using words and actions to reveal the elements and images of a story while encouraging the listener's imagination*"), one may wonder if the designed activity has realized real storytelling. For example, there was no one present in the role of 'listener'. In addition, one could argue that the 'elements and images' were already given by the activity and therefore not needed to be revealed anymore. In fact, it may be better to say that storytelling was used as inspiration for the activity, or that the definition was loosely interpreted. Regarding the activity, it may be better to talk about story parts, or story fragments, instead of a complete story created via storytelling. These fragments were created by the children to make cause-effect, conceptual relations between the given objects. Given that especially these relations were important for word learning, it could be said that there was no need to create full stories from a design perspective. It might also be much to expect from the children, to generate a complete story. Here also the question arises of when a story is complete.

The realization of learning by teaching can be questioned. The children were not prepared to teach. Only right before the start of the activity, they were told to teach. In addition, there was almost no time to create a bond with the robot (as student). These are two essential components of the earlier described learning by teaching method developed by J.P Martin in [19]. In this method, the child prepares, during his own learning, to teach to another student. While doing this, the child is motivated to put extra effort in his own learning, because it cares about teaching it properly to the student. This care comes forth of a certain empathy, or social bond with the student. In this method, this preparation before the teaching is where a part of the learning already takes place for the teaching student. In the designed activity, such a preparation phase was not present before the activity. Additionally, while a bond was being created with the robot via emotions and speech, the strength of this bond may have been too weak to change the child's behaviour with the created bond. One may wonder if the teaching during the activity was really done out of care for the robot, if the created bond contributed to more carefully teaching behaviour or if the children felt like they were teaching at all.

Longer-term interaction with the activity was not tested. A certain novelty effect may have played a role during this user test. In other words, the children could have been enthusiastic about the robot/activity purely because it was something new for them. This enthusiasm may have positively contributed to the achieved results. In addition, one might argue that the children simply wanted to explore all possible interactions with the system instead of really using specific things with a reason for the story, or to achieve the goal. If the children would be able to play once a week with the activity, for a period of two months, enthusiasm and/or results may decrease. This prototype was not designed to be engaging for a long period. Only a limited set of actions to take was available for the children. However, long-term engagement is essential for this activity to realize its goal, given that language learning is a long process. One could say that the proof of concept was not fully tested with this user test, since the activity's goal is long-term. On the one hand this is true because of the short test, on the other hand it is not. Firstly, the specific robot was indeed new to the children, but they had been using much more advanced robots already. Due to this, the novelty effect may not have had such a big influence. Secondly, during this short test, the children already focused much on the challenge of the puzzle and were captivated by this. Such a puzzle mechanic could easily be expanded with many new words or aspects that would keep it interesting for a longer period. When designing the activity for long-term use, (game) design principles to create an interesting long-term experience would become important. For example the activity would need to stay elegant, and not grow too much in terms of rules and mechanics. This would lead to complexity.

One can doubt the added value of the whole activity, given that "only" two words were learned on average and that the remembered words were often the easiest ones from the list. Further research will need to be conducted to test the effectiveness of the activity versus a standard word-learning method. For now, it can be argued that learning a word in the context of its use is being achieved via this activity. This is important to be able to use a word in practice, when a similar context occurs. Learning words from a list may link words to a wrong concept. For example a child can remember a word from a list because it would always be after a certain other word on the list. If learned in this way, it is unlikely that the child will remember this word if its actual context of use occurs in practice. However note that learning in context may be more important for constructions like "on top of" or "given that", instead of straightforward nouns. Constructions like these could be implemented in a later version of the activity. Besides this, note that the children only played for ten minutes and that the test was at least twenty hours later. These circumstances may result in low scores for any word-learning method.

Several remarks can be made about the user test. Firstly, this user test was focused on the proof of concept. Based on the data analysis, suggestions were made on the influence of several components of the designed activity, gender, and role of the child. These factors were researched mainly out of interest and because the opportunity was there. However, to make convincing statements, different user tests will have to be conducted that focus on a particular aspect. Secondly, the user results were influenced because the children were allowed to choose their role themselves. The results suggested that this has led to a division of the children based on intelligence and/or enthusiasm. Suggesting convincing statements about the influence of the role was made very hard this way. Also the fact that some words were put very close to their translation, could have had influence on the results. In a future user test, all these factors should be randomized, to remove their possible influence. Thirdly, one could doubt if these children were representative for average eight-year-olds, given their above-average experience with robots and modern technology. The novelty effect, misuse, or misunderstanding could have a bigger influence when a test is done on another school.

8.1 Future work

Based on the discussion of the user test results and of the project as a whole, some options can be given for future work. On the one hand, more detailed research can be executed with the current prototype. On the other hand, ways to possibly improve the activity are presented.

First of all, as already mentioned, the suggestions made based on the user test results can be researched in further detail. This would be important to verify that the suggestions are true. This cannot be said with the current user test because this test was very generally about the proof of concept. Several factors differed per session, instead of one, or no counter scenario was tested. For example it was suggested that repeating the words had no significant influence on the results, but no test was done without this repeat mechanic. A test would need to be done where a group of children is randomly split in two. The two groups would need to do the exact same experiment, but one with the repeat mechanic and one without this mechanic. For such a test, a bigger group of children would be needed. Having about ten children per group (five duos), could quickly become not significant. Besides this, one can seriously doubt how to design two identical activities with only one difference. The activity forms a coherent whole in which the different aspects have influence on each other. Simply removing the repeat mechanic would certainly affect other factors as well. For example it would allow a child to only teach a word to the robot eight times. The child could get frustrated because of this limit. It could also get bored after having taught these eight words. With the repeat mechanic, the robot would regularly interrupt the story with a question to reteach a word. If this would not happen, the story might progress much faster and finish earlier. It could also have a positive effect on the story because it does not get interrupted. Also the use of the robot's emotions would decrease, since especially the 'sad' emotion is often used during re-teaching. On the one hand, one can argue that all these effects are effects caused by removing this mechanic and that it is part of removing the mechanic. On the other hand, one can argue that, by removing this mechanic, so many factors are influenced, that one can no longer speak of two similar activities. This way, the effect of removing one mechanic would still not be measured.

To further explain last argument, removing an (arguably) bad mechanic in terms of direct educational value, can still lead to less educational value overall, because removing it negatively influences other important contributors to educational value. In other words, a mechanic can have educational value indirectly. For example, if the child gets bored quickly because it can only teach eight times, it has less fun and is less enthusiastic. This could lead to worse performance on the test due to less communication with other children and/or less motivation to score well. It could be tempting to conclude that the re-teaching mechanic does have positive impact. However, in fact it is its indirect influence that is positive. For example slowing down the pace of the activity, or allowing the child to teach a word many times. The re-teaching itself may be useless. A correct conclusion would be to replace the re-teaching with mechanics that maintain the indirect positive effects, and add direct positive effects on top.

One could say that it may be impossible to test the effectiveness of a single mechanic purely on its own, without influencing other aspects of the activity. Now one may wonder why you would even want to test this, arguing that the activity as a whole is what matters. If this is the researcher's reasoning, it would be better to speak of testing different versions of the activity, instead of testing the influence of a single mechanic. Testing different versions is possible. In this case, single aspects can be evaluated and compared within the activity as a whole. This would still need to be done with great care, given the many relations between all aspects. It would still be very hard to talk about the influence of a single aspect because all these relations may be hard to understand.

Knowing the above, a different version of the activity could be designed without the re-teaching, but with replacement for the indirect effect of removing this. The child should still be able to teach words to the robot more than eight times. The pace of the story should not drastically increase because there are no interrupts anymore. The robot should still be able to show its emotions in a valuable way. These replacements themselves may each have their direct and indirect influences on the activity. However, this may be the best way to realize similar activities, with and without re-teaching. Just removing re-teaching could lead to two more different activities.

In a similar way, other aspects of the activity can be looked into in further detail by developing a different version of the activity without this aspect. For example the restriction of eight words. From the results, it appeared that taught words were not remembered more often. It would be interesting to see how the activity changes without this limit. Again, this limit cannot be removed without influencing other aspects. The current robot face only shows eight learned words. This could be changed to possibly show all fifteen words. However when doing this, the influence of re-teaching decreases because each single word will be forgotten less often. There will be a bigger spread of words that are all re-taught once, instead of eight words that are re-taught two or three times. Given the suggestion that re-teaching may be less important, removing the eight-word limit could be a solution to let the children teach the robot many times in a different way than via re-teaching. Another option would be to let the children decide themselves how many times they want to teach a word. The word does not fade anymore. Instead, each time a word is taught, it simply adds up to the existing mix of colors on the robot screen. This way words/colors can appear several times on the robot's face. On the one hand, this adds more freedom to the activity and thus (arguably) fun. On the other hand, this feature could escalate easily via children trying to teach a single word over and over again and not paying attention to the story. Additionally, the ending of the activity should be redesigned because that came forth of the eight-word-limit.

The issue that taught words were not remembered more often could also be explained in terms of learning by teaching. As was already explained during the discussion of the results, the learning by teaching aspect could be redesigned to more closely resemble the method of J. Martin, described in [19]. This could be done by extending the activity with a preparation phase during which the children learn the words, in order to teach them later in the activity. An important choice here is whether or not the children will have access to the words they had to learn, during the activity. If not, arguably important child-child communication could be stimulated because the children will remember different words from this preparation phase. The activity could be designed so that children need each other's knowledge to achieve better results during the activity.

From the results, it was suggested that child-child communication could be important for word learning. A user test could be done to test the influence of this. For example the group of children could be split in two. One group follows the same user test procedure as done for this project, the other group comes together afterwards to discuss the activity and the words they learned. This would be the influence of active information exchange that is tested. To test the more informal, "passive" information exchange of children talking to each other, the test circumstances could be changed. One of two groups would not be allowed to talk to each other about the activity until after the test. In practice this may be hard to pull off because children can do it anyway and they may hear things from children of the other group. A test like this should be considered very carefully.

Also to research the effect of gender and age, research could be done by dividing the children in groups. For gender, duos of children could be made in three ways, boy-boy, girl-girl and boy-girl. The making of duos would need to be done carefully because whether or not the children in the duos are good friends may play a role for activity execution and teamwork. This can be suggested based on observations of the user test. Arguably duos of the same gender have a higher chance to be good friends. The proportion of friends in duos should be the same for all three groups. For age, different groups of a school can be tested. It is important that these groups are from the same school. If not, unwanted differences between the two groups of children could play a role. For example, different school curricula or different socio-economic class of the children('s parents). It would be interesting to test a group 6, group 4, and group 2. These are ten, eight and six-year-olds respectively. Many hypotheses can be tested in such a test. For example, it is expected that ten-year-olds will tell little to no stories on their own, because they may not be engaged anymore by this activity. However, they may still learn many words because they may better understand the educational goal of the activity. On the other hand, six-year-olds are expected to be more engaged by the activity. They may tell many (parts of) stories on their own but may not understand the educational goal. They could be too engaged by the fun aspects and learn less because of this.

Several suggestions were made based on the influence of the role. These can be verified with more testing. Crucial would be to not let the children choose themselves what role they take. To test influence of the role, the exact same user test as already done can be executed, but then with random role assignment, for example by tossing a coin. Also interesting would be to test if intelligence or enthusiasm indeed play a role in role selection. This could be done by getting more insight in the children's school performance, via the class's teacher. For enthusiasm, the children could note themselves on the pre-test how enthusiastic they are for the coming activity. However, this might not work. Children might not yet be capable to fully reflect on how they think of something and just give maximum enthusiasm score. Lastly, it can be tested how the activity would proceed if nothing was told about any role, if the children were free to interact with both the robot and the tablet.

From the user test became clear that supporting the puzzle-aspect would be the most obvious change to the activity that is not about further testing existing features. Unintentionally, many children perceived the activity as a puzzle where the objects needed to be exchanged to put them on the "correct" location. As told, children found this an engaging challenge that evoked storytelling as reasoning for their actions. The robot's texts were perceived as hints for the puzzle. The activity did not fully support this. For example when encountering the sea, the robot said "oh I would like to swim, it is so hot in France". Children reacted that the robot was already at the beach, or that the sea object was already at the beach (location). Some robot texts did not suggest a clear location to go to, for example the turtle that wanted to join the adventure, or the car that could be used to drive anywhere. The activity could be changed in many ways to better support this puzzle. Extensive brainstorming would be needed for this. Some suggestions are for example to fit the activity's theme to the puzzle. It could be about a mystery or a detective. The robot's texts and the locations could be used to communicate hints with. The object's colors could be used to suggest the solution (all objects of the same color at one location). J. Schell's chapter on puzzle design would help here [17]. However it is important that the educational value of the activity remains. The puzzle should also contribute to this, or at least not hinder it.

A very interesting design option to further investigate is changing the balance between digital and tangible interaction. Currently, teaching is done fully digitally, via the tablet UI. Also the learned words are digitally represented on the robot. All this could be made tangible instead. Suggestions are to let each word be represented by a small chip/tile and to let the robot collect these. Also a tangible interface could be designed representing the robot's memory. This would need to be filled with tangibles of the learned words. More idea generation will result in more creative ideas. On the other hand, by making more, or everything, digital, the discussion arises why a robot would be better than a virtual system. It would be interesting to look into this by developing a similar activity, but then fully digital.

From an own experience it can be said that words are very effectively learned when there is a strong need to actually use them. For example learning the French ‘court’, which means ‘short’, when going to a hairdresser in France. A word will be remembered very well in this way, because it is related to an experience where its use was necessary. The activity could be designed with a stronger focus on using the words, instead of learning them. In fact, already a certain need to use words was tried to design, by making the elephant need to know the French translation in order to use a word. Knowing the French translation was not made a strict necessity however. This may still not be a good solution, because it limits which words can be used, since a limited number can be taught to the robot. Effectively achieving this need may require the activity to become more complex and/or dialog-intensive. Nonetheless it may be worth to experiment with.

Finally, and arguably most important, the effectiveness of the activity can be tested versus a normal word-learning method. The group of children would need to be split in half. One half would do the activity for ten minutes. The other half would be presented with a list of the fifteen words and their translations and would have to learn them for ten minutes. The effectiveness of the test could be tested by comparing the after-test results of the two groups. Of course here as well, the children should be divided in groups fully randomly.

8.2 Personal reflection

“*So, on which topic of this list you would like to spend 500 hours?*” is the question that was asked me before the start of the project. I had to think about what I liked the most. I had to consider what would keep me interested for such a long time. While many topics were interesting, I had to filter the “endlessly interesting” topic out of all the “interesting for 100 hours” topics. This project immediately caught my attention because firstly it would allow me to be in contact with users, secondly it would give me much design freedom, and thirdly the topic is about stories and games. As a child, I wrote stories as soon as I could write, and I spent many hours in the *Warcraft 3 world editor* to create my own game missions and worlds. I spent full weeks on creating my own strategies in a card game. Of course I played with LEGO bricks too. After reading the project description I noticed how I immediately started to think creatively about what kind of playful activity I could make. I felt the same inner drive to create as I felt as a child. In my experience, fundamentals of what captivated me as a child, form the basis of my interests now. In other words, this assignment triggered my fundamental interest. One that has kept me busy for hundreds of hours already, 500 more would be no problem.

Besides the interesting topic, I was finally able to do a solo project. Every design decision was under my control. There was no longer a dependence on lazy, unmotivated and/or undisciplined teammates. I do not mean this negatively to them, they are probably just not yet doing what they really like, that happens to everyone. Equally important, I could fully focus on this one project. No other (uninteresting) courses, projects or activities would take away a big part of my time, only small parts. Except a machine learning course then, interesting but time-consuming. The combination of all these factors made an environment in which I was very motivated to make a great project, a crown on top of my Bachelor. By doing the project, I learned more about my own style of doing work. I like to focus on one assignment.

This way I do not get mentally distracted and I can put all my thinking-power in one thing. In addition, I should try to find people with the same passion as I have, to not be dependent on unmotivated people. Most important, I am now convinced of what interests me the most: the iterative process of creating an experience that is built by elegantly combining many factors into a coherent whole. Whether this is in the form of creating a deck of cards full of strategies, writing a story or designing a user product, the same interest lays at its basis. For my further education, I am now surer of the direction I should go in, where I should lay my focus.

Now that I have reached the end of my Bachelor study, I can reflect on it in terms of the project. The Bachelor has prepared me well for the final project that I have done. I learned concrete skills like programming, using electronics, rapid prototyping, design principles, scientific writing, user testing, statistics, and presenting. I also built much experience on soft skills like (multidisciplinary, international) team work, communication with stakeholders like the client and users, critical thinking and taking into account ethical aspects. Many of these skills I was able to use during the project. Therefor the project could certainly be seen as a final piece of work where everything comes together and all learned parts of a design process fall on their place. The project helped to better realize the importance and role of everything learned during the study by putting theory into practice.

One thing I encountered during the project that I had less experience with is data analysis. I learned about statistics, and I learned about gathering user data, but I did not learn about using the data to evaluate hypotheses. I found myself struggling with drawing conclusions based on the data, and with deciding what to do with the data in order to get the interesting information out of it. I learned that structured preparation of the user test is important for effective data analysis. By critically thinking about the concrete goals, hypotheses and how to evaluate them in terms of the gathered data, better overview can be maintained. In other words, after the user test the overview was lost, because the test was not well enough prepared in terms of goals for the project as a whole.

Another learning moment I had when I was stuck in the technical jungle of the current system. I started too late with finding out how the system was built. I started when I actually wanted to start using it to implement my idea. A big lesson I learned is that next time, I should start familiarizing myself with the technology much earlier. This way I will be able to build a high-fi prototype more fluently. In this project, I should have started with the technology already during the state of the art phase. I should have started to learn about Android app development in Android Studio, and about the structure of the current system. The only difficulty with this is that it may not yet be clear what exact technology to learn at such an early phase in the project.

As just became clear, there are always points for improvement. However, reflection on why something was a success, is at least equally important. In order to continue with a winning strategy, I should be aware of what I am actually doing that gives me the success. I can point out some (in my opinion) key contributors to the success of this project. Firstly the weekly meetings with the supervisors. In the first meeting, a standard moment in the week was set for a weekly meeting. These meetings resulted in very clear communication with the supervisors about the project, and in fast-paced project progress with weekly goals to achieve. The meetings allowed the supervisors to give much useful feedback during critical moments. With less meetings, feedback can easily come too late. In addition, there is less motivation to work hard because the (informal) deadline of the meeting occurs less often. Secondly, the execution of three user tests allowed many design iterations and resulted in a decent activity in the end. Such user tests are possibly the most effective way to identify design flaws and possible ways for improvement. A user test with a paper prototype on just a hand full of children can be more informative than a whole week of thinking and doubting between choices. Thirdly, the availability and usage of help from others. Alejandro, a supervisor, has played a crucial role during the realization of the high-fi prototype. Without his help, certain technical problems would have taken much longer to solve. Also my uncle has contributed greatly. His intelligence is unique. With his world-class software skills, he solved a key problem I was facing during realization. His knowledge was mostly needed during data analysis however. While I was a bit lost, he knew exactly what to do. Finally also the voice actor should be mentioned. She kindly offered her help and did not want anything in return for all the voice samples. Fourthly, the primary school has to be thanked a lot. Communication was great and acquiring the parents' consent to do the test went surprisingly smoothly. Finally, self-reflection has played a key role in the success. In the first place, during project selection, I was able to clearly identify what interested me the most and why. In the second place, during realization, I was able to realize in time that continuing to use the coBOTnity system would probably cause me a lot of trouble because of a lack of time. Because of this, I was able to switch to Processing in time.

This project was a great experience. I was able to put my knowledge from the Bachelor into practice in a way that really interests me. I am proud of the result.

9. References

[1] O. Mubin, C. J. Stevens, S. Shahid, A. Al Mahmud and J. Dong. (2013) "A review of the applicability of robots in education" In: *Technology for Education and Learning*, Vol. 1, 2013, p. 1–7. [Online] http://roila.org/wp-content/uploads/2013/07/209-0015.pdf

[2] S. ter Stal (2017). *Designing an Interactive Storytelling System for Children Using a Smart Toy* [Master Thesis, University of Twente] http://essay.utwente.nl/73287/1/ter%20Stal_MA_EEMCS.pdf

[3] M. Lindberg, K. Masson, B. Johansson, A. Gulz and C. Balkenius. (2017) "Does a robot tutee increase children's engagement in a learning-by-teaching situation?" in *Intelligent virtual agents* p. 243-246 [online] https://link.springer.com/content/pdf/10.1007%2F978-3-319-67401-8.pdf

[4] C.C. Chase, D.B. Chin, M.A. Oppezzo and D.L. Schwartz. (2009) "Teachable agents and the protégé effect: Increasing the effort towards learning." in *Journal of science education and technology* august 2009, vol. 18, issue 4, p. 334-352 [online] https://link.springer.com/article/10.1007/s10956-009-9180-4

[5] A. Krzywinski and W. Chen (2015) "Hi Robot: Evaluating robotale" in *Proceedings of the 2015 ACM International Conference on Interactive Tabletops and Surfaces, ITS 2015,* p. 367-372 [online]https://dl-acm-org.ezproxy2.utwente.nl/citation.cfm?doid=2817721.2823507

[6] A. Druin, J. Montemayor and J. Hendler. (1999). "Designing PETS: A Personal Electronic Teller of Stories" in Proceedings of the SIGCHI conference on Human Factors in Computing Systems, Pennsylvania, USA p. 326-329 [online] http://www.cs.umd.edu/hcil/trs/99-25/99-25.pdf

[7] J. Werfel. (2014) "Embodied Teachable Agents: Learning by Teaching Robots" in *New Research Frontiers for Intelligent Autonomous Systems* at 13th International Conference on Intelligent Autonomous Systems, Venice, Italy [online] http://people.seas.harvard.edu/~jkwerfel/nrfias14.pdf

[8] J. Kory and C. Breazeal (2014). "Storytelling with robots: Learning companions for preschool children's language development" in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, Vol. 2014-October, 2014, p. 643-648 [online] http://ieeexplore.ieee.org.ezproxy2.utwente.nl/stamp/stamp.jsp?tp=&arnumber=6926325&tag=1

[9] C. Chang, J. Lee, P. Chao, C. Wang and C. Gwo-Dong (2010) "Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school" in *Educational Technology & Society*, vol. 13 n. 2 p. 13-24 [online] https://eric.ed.gov/?id=EJ895653

[10] D. Wood and H. Wood. (1996). "Vygotsky, Tutoring and Learning" in *Oxford Review of Education.* p.5-16. https://www.learntechlib.org/p/82050/

[11] M. Gordon, E. Ackermann and C. Breazeal (2015) "Social Robot Toolkit: Tangible Programming for Young Children" in *Proceedings of the tenth annual ACM/IEEE international conference on human robot interaction*. Portland, Oregon, USA P. 67-68 [online] https://dl.acm.org/citation.cfm?id=2702001

[12] J.M.K. Westlund and C. Breazeal. (2016). "Transparency, Teleoperation, and Children's Understanding of Social Robots" in ACM/IEEE International Conference on Human-Robot Interaction Vol. 2016-April, p. 625-626 [online] http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=7451888

[13]C. Tonzar, L. Lotto and R. Job. (2009). “L2 Vocabulary Acquisition in Children: Effects of Learning Method and Cognate Status” in *Language learning* vol. 59, issue 3, p. 623-646 [online] http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9922.2009.00519.x/full

[14]A.M. Rosenthal, C. Strassmann and N.C. Krämer. (2016) “Robots or Agents – Neither Helps You More or Less During Second Language Acquisition” in *International Conference on Intelligent Virtual Agents* p. 256-268 [online] https://link.springer.com/chapter/10.1007/978-3-319-47665-0_23

[15]A. Yamaç and M. Ulusoy (2016). The Effect of Digital Storytelling on Visual Memory and Writing Skills. Comput. Educ., 94(C), p. 298-309. [online] https://www.iejee.com/index.php/IEJEE/article/view/145/142

[16]P. Vogt et al. (2018). “Guidelines for Designing Social Robots as Second Language Tutors” in *International Journal of Social Robotics* p. 1-17 [online] https://link.springer.com/epdf/10.1007/s12369-018-0467-6?author_access_token=aOo5TFaX9vEDiBU4jfxFLve4RwlQNchNByi7wbcMAY6UvAslVbxLmt8FqrW6dx_4pTwjKI_kdw4cQUZQnQj_EUxrF57GOkqdAmGCseBnKy66FsL_c1aXHt2FA6yWYJ16SQ8kwg7LApzqxb3AH0onxg%3D%3D

[17] J. Schell. *The art of game design: a book of lenses*, Burlington, USA, Morgan Kaufmann publishers, 2008.

[18]J.K. Westlund, G. Gordon, S. Spaulding, J.J. Lee, L. Plummer, M. Martinez, M. Das and C. Breazeal. (07-03-2016) “Lessons from Teachers on Performing HRI Studies with Young Children in Schools” in *ACM/IEEE International Conference on Human-Robot Interaction*. Vol. 2016-April, 12 April 2016, p. 383-390 [online] http://ieeexplore.ieee.org/document/7451776/

[19]J. Grzega. (04-07-2008) “Learning by teaching, the didactic model LdL in University classes” in *Journal of education for teaching* vol. 34, 2008, issue 3, p. 167-175 [online] http://www.joachim-grzega.de/ldl-engl.pdf

[20] D. Hood, S. Lemaignan, P. Dillenbourg (02-03-2015) “When children teach a robot how to write” in *ACM/IEEE International Conference on Human-Robot Interaction* . p. 83-90. [online] https://infoscience.epfl.ch/record/204890/files/hood2015when.pdf

10. Appendices

Appendix 1 Information brochure and consent form for parents

INFORMATIEFOLDER
ROBOTACTIVITEIT IN GROEP 4 VAN BASISSCHOOL DE DOELAKKERS IN WEEK 24, 2018

Beste ouder, uw kind wordt uitgenodigd om op school mee te doen aan een leuke en hopelijk leerzame robotactiviteit in het kader van de Bachelor afstudeeropdracht van student Gijs Verhoeven (Universiteit Twente, Creative Technology). In deze folder vindt u hierover meer informatie.

Inhoud en doel van de activiteit
Het doel van de afstudeeropdracht is om een activiteit te ontwikkelen waarmee basisschoolkinderen een tweede taal kunnen leren op een interactieve manier. De activiteit gaat over een olifant, een robotje, die op avontuur gaat in Frankrijk en daar Franse woordjes leert in een verhaal context. De kinderen bepalen wat de olifant beleeft op zijn avontuur en helpen hem met het leren van Franse woorden via een interface op een tablet. Hieronder ziet u een plaatje van de robot die gebruikt gaat worden en een olifant voor zal stellen, op een speelveld.

Het verloop van de activiteit en het verzamelen van gegevens
Eerst zal de activiteit kort aan de kinderen worden uitgelegd en de robot zal worden geïntroduceerd. Ook zal worden gevraagd naar voorkennis van de Franse taal. De kinderen zullen de activiteit doen in duo's of alleen, buiten het klaslokaal. Dit duurt 15 tot 20 minuten. De planning is om alle testen te doen op één dag. De reguliere lessen worden die dag zo min mogelijk verstoord. Mits u als ouder/verzorger daarvoor toestemming geeft, zullen tijdens de activiteit de volgende gegevens verzameld worden: de student zal observeren hoe de kinderen de activiteit doen en notities maken. Ook zal hij foto's en video opnames maken van interactie met het prototype. Leeftijd en geslacht zullen worden genoteerd, geen andere persoonlijke gegevens. Na afloop van de activiteit zullen een paar simpele vragen over de activiteit worden gesteld aan de kinderen. Een dag na de activiteit zal kort worden getest welke Franse woorden de kinderen hebben onthouden.
De notities, video's en informatie zullen worden gebruikt om het concept/prototype te verbeteren en zullen alleen door de student en de supervisor worden gebruikt en alleen voor deze afstudeeropdracht. Foto's voor in het eindverslag en de eindpresentatie worden zo genomen dat de kinderen in geen geval

te herkennen zijn en er wordt gefocust op interactie met het prototype. Er zijn geen risico's aan de activiteit verbonden en de gegevens zullen worden vernietigd na afloop van het onderzoek (juli 2018).

TOESTEMMINGSFORMULIER

Door dit toestemmingsformulier te tekenen geeft u toestemming dat uw kind mee mag doen aan deze activiteit en dat gegevens zullen worden verzameld en gebruikt zoals in de informatiefolder is beschreven. U bent niet verplicht te tekenen. Als u niet tekent, doet uw kind mogelijk niet mee en wordt er in ieder geval geen data verzameld. Als hij/zij heel graag wil en er is tijd over, mag hij/zij wel met de robot spelen, maar zonder dat er data wordt verzameld. Als u wel tekent is uw kind niet verplicht om mee te doen, dat mag hij/zij zelf kiezen op dat moment. Ook mag hij/zij tijdens de activiteit stoppen, als hij/zij dat wil. De klasdocent heeft veel ervaring met de kinderen en kan ook zeggen dat de activiteit voor een kind niet doorgaat. Bijvoorbeeld als hij/zij zich niet gedraagt in de klas, of als iets anders op dat moment belangrijker is. Voordat de activiteit plaatsvindt, zal die worden bekeken en gekeurd door de basisschool.

Hebt u nog vragen of bent u geïnteresseerd in de resultaten van het onderzoek, neem dan contact op met de student, Gijs Verhoeven.

☐ Ik ben over dit onderzoek volledig geïnformeerd en geef toestemming dat mijn kind hieraan mag meedoen in week 24, 2018. Ik geef toestemming voor het verzamelen van gegevens voor onderzoek, zoals beschreven in de informatiefolder.

Naam kind:

Naam ouder/verzorger:

Datum:

Handtekening:

Contact informatie

Student: Gijs Verhoeven (g.a.g.verhoeven@student.utwente.nl), 0611853386
Supervisor: Dr. Mariët Theune (m.theune@utwente.nl), 053-4893817.

Universiteit Twente, Drienerlolaan 5, 7522NB, Enschede.

Als u niet tevreden bent over de wijze waarop dit onderzoek is uitgevoerd, kunt u uw klachten richten tot de Ethische Commissie van de Faculteit EWI (Elektrotechniek, Wiskunde en Informatica) van de Universiteit Twente, tel: 053 -489 6719, email: ethics-comm-ewi@utwente.nl.

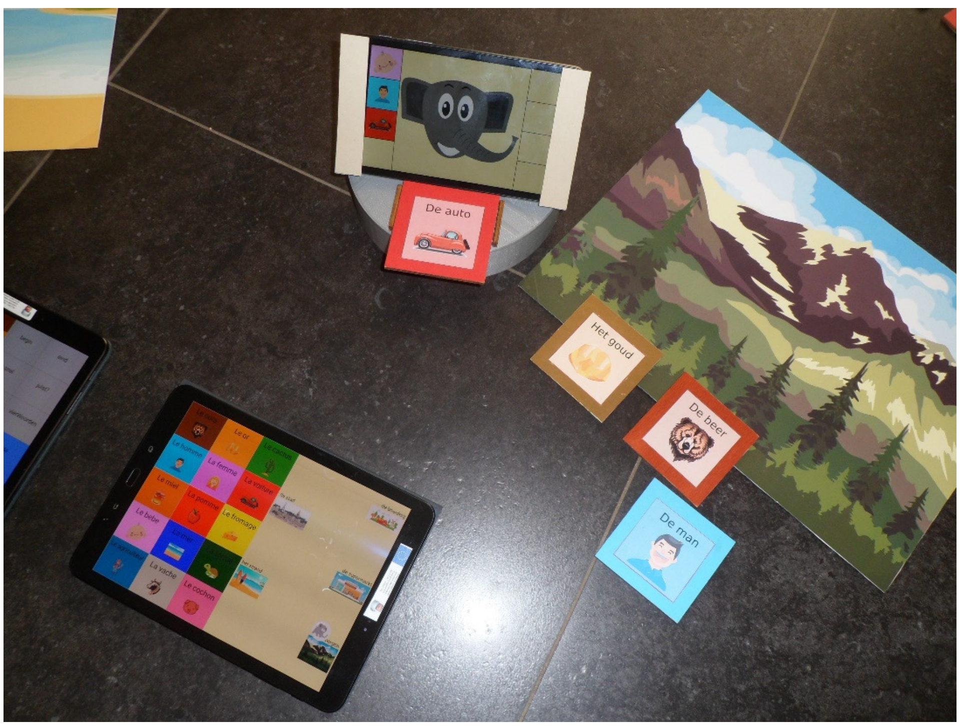
Appendix 2 The Reflection report

Reflection report

-

Social robots in primary education

The current state - the near future - the far future



Reflection for Creative Technology

Gijs Verhoeven

S1757571

15-06-2018

Table of contents

Introduction

This graduation project is about the design of a second language learning and storytelling activity for primary school children with a social robot as main character. In the activity the robot represents an elephant that goes on an adventure to France. There he can visit several locations and he can encounter a variety of objects (like a bear and a cactus). The children decide how the adventure/story proceeds and therefor what places the robot will visit and what objects he will encounter, out of the given options. The robot needs to learn some basic French words during his adventure and the children will teach him. The aim is that the children will eventually learn some French themselves. Firstly by teaching it to the robot during this storytelling activity, secondly by being involved and engaged in the activity in general. This approach makes use of 'learning by teaching'. This means that someone learns by teaching it to someone else.

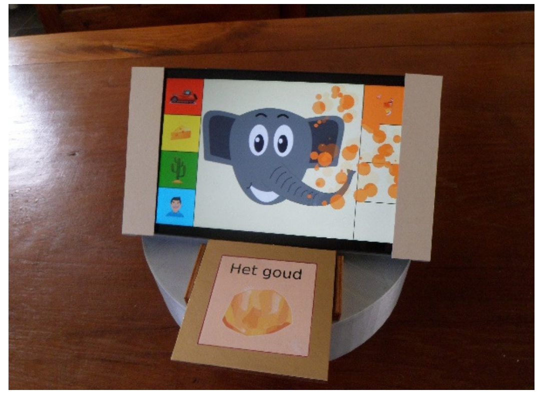


Image 3 The robot used for the project

Image 2 The different object tiles

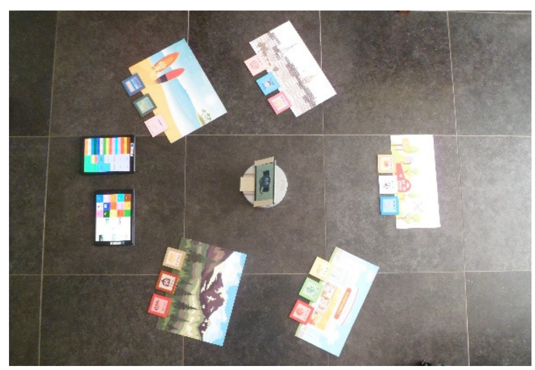


Image 3 The whole setup

On image 1, the robot can be seen. He carries gold with him, has already learned four words and is happy because he just learned his fifth. This is accompanied by a colorful explosion. On image 2, the variety of objects can be seen, represented by images on wooden tiles. On image 3, the whole setup can be seen. There are five locations: the beach, the mountains, the supermarket, the farm and the big city. Each location has three objects. The robot can carry an object and move it to another location.

The robot is being controlled via two tablets. One tablet is for the children. On this tablet, they can order the elephant to what location to move and they can teach him the words. The other tablet is controlled by a researcher from a small distance, for tele-operation. Tele-operation in this context means that the robot is being (partly) controlled by the researcher without the children knowing this, in order to fake more autonomous behavior. In this case the researcher controls the robot's emotion and speech. There is a variety of speech samples and there are three different faces/emotional expressions (normal, happy, sad). The researcher decides when the robot says what, based on the child's choices and interaction during the activity.

Thus, this graduation project focusses on some form of child-robot-interaction. While designing the robot and its behavior, it is important to make it appealing for, and accepted by, children. In other words, it is technology that intentionally wants to trigger social and emotional bonding with children. In order to successfully design a child-robot interaction activity from an ethical perspective, an elaborate discussion on the ethical side of social robots in primary education will take place. First some definitions will be given. For example, *what is a robot? What is a social robot?* Then, three different situations will be discussed: the current state, the near future (+- 15 years from now) and the far future. Related questions that will be addressed are: *How might a future classroom look like, when social robots have been integrated on large scale? What is the current state of technology and robots in the classroom? How will the teacher be affected by robot integration in his/her classroom? How to deal with aspects like robot authority, responsibility and child specific data collection?*

Some definitions

In order to discuss the topic of social robots in primary education, some working definitions will be given to make sure that the different notions concerned are clear. First off, there is the 'robot'. As defined in [1] by P. Lin et al. a robot is "*an engineered machine that senses, thinks and acts".* In other words, a robot should have sensors, processing ability and actuators. This definition rules out the fully (remote)-controlled machines, since those do not think. Note that this definition does not state that robots have to be electromechanical. Virtual robots or (partly) biological robots also count. Now what does 'think' mean exactly in this context? In [1], this is defined as the ability to process information from sensors and other sources, and to make decisions autonomously and act based on the information. 'Autonomy' is defined here as the ability to operate in a real-world environment without any form of external control for an extended period of time.

Within the category of 'robots', there is the sub-category of 'social robots'. In [2], C. Bartneck and J. Forlizzi define the social robot as a (semi)-autonomous robot that interacts and communicates with humans by following the behavioral norms expected by the people with whom the robot is intended to interact. Note that a social robot has to be at least semi-autonomous. If it would be fully controlled by a person, the robot would not be a robot in the first place, according to above definition of 'robot', and the robot's social behavior would merely be an extension of the human controlling it. Important here is the communication with humans. Robots that communicate only with other robots are not considered to be social. The social behavior does not have to be cooperative. For example competitive behavior in a game played together is also counted as social.

In this report, the notion 'primary education' is used according to Dutch standards. This means that primary education is about children of all intelligence levels, and of an age of four until twelve. At a primary school there are eight groups. Each child starts at group one and progresses one group each year. This makes that in each group, the children have about the same age. According to M. van den Bogaerdt in [3], the average amount of children per group is twenty-three. In practice, this amount can vary between about ten and thirty. Often each group has only one teacher, who is sometimes helped by a second teacher or an intern.

The current state – technology as a tool for education

For the past decades, technology has quickly merged into the everyday lives of people. Major technological inventions like the internet and smart phones have rapidly changed the world. Also on primary education, technology has its influence. For example traditional blackboards are often replaced by 'smart boards' (image 4) and books are getting replaced by tablets. These technologies have many advantages. For example there is less paper waste, photos and videos can be shown in an instance, study materials can be made interactive, more exciting and specified towards the child's needs, there is instant internet access to search for information, it can motivate shy students to participate, and tedious tasks can be automated, as pointed out by M. Rosso in [4]. One could very well argue that children benefit a lot from these advantages provided by technology.

Image 4 A smart board

There are disadvantages too, however. A major one being the dependence on technology. If the smart board malfunctions for whatever reason, teaching will become very hard for the teacher. Think about simple things like the audio does not work, there is no internet connection, the screen stays blue or the tablet has run out of power. Sometimes there may be a quick fix, but when occurring regularly, the lessons get quite disturbed. Then there are the major break downs, when some technology simply does not work anymore for a longer period of time. In such a case, people may have to move to another class room or go back to the good old blackboard and books, until the problem is fixed. Someone has to do this, and this is an important point: Once some technology is being used in the classroom, there should be someone that has knowledge of that technology and that can repair problems as soon as they occur.

Even more important, the teacher has to be able to use this new technology (to its full potential). In order to do so, (s)he has to acquire some technological knowledge. For some teachers, this could be a big challenge. People could be very bad with technology. Older people can generally handle new technologies worse as well. In other words, some excellent teacher may be prevented from properly doing his/her job, simply because of the introduction of new technologies to the class room. While these technologies could add much value to education when used well, they could also harm education when teachers have not the required technical knowledge.

Another problem is on the financial side. Technologies are often expensive. Not all schools have the money to buy these technologies on large scale. If a single tablet costs €500,-, buying them for a whole class will cost more than €10 000,-. Continuous maintenance of the devices after purchase, as discussed above, will also cost money and time. When money is the deciding factor for schools to buy education enhancing technologies, the rich schools will benefit more than the less rich. This puts the less rich on a disadvantage and increases the gap between the rich and the poor. Even when the technologies are not used to their full potential, the mere fact of having them can already motivate parents to choose to put their child on that "modern" school. Besides this, children (from less rich parents) may not have access to these technologies outside the school class and may therefore not be able to practice at home. Public places are not always the solution since there may be queues and software cannot always be downloaded on public machines.

Another disadvantage is the distraction that technology can cause. In the first place, children can go to whatever website they want. They could play some game, watch a funny video or check social media while they actually should be doing their interactive mathematics activity. In the second place, just the fact of being allowed to 'play' with modern technology, can amaze children to such degree that they pay less attention to the actual subject, as mentioned by Rosenthal et al. in [6]. This is more the case when robots come into play, as will be discussed later. In addition, social interaction and verbal communication is also key in a child's development. Teachers need to be careful to practice this as well, and not focus only on the activities with technology. Finally, there is the increased opportunity to cheat by using the internet, and a risk of finding information from sources that are not trustworthy, as pointed out by M. Rosso in [4].

The current state – social robots as tools for education

Many of the above-mentioned pros and cons of technology in the classroom also apply to introducing robots in the classroom. On the one hand, robots are engaging, increase interactivity, can be specified towards specific needs and could do tedious tasks. On the other hand, teachers will have to understand there teaching tools in order to use them properly, robots are expensive, can be distracting by being 'too fun' to play with, and may reduce human-human interaction. In fact, many of these pros and cons are much more present with robots than with other technology. For example, a robot will require even more technical knowledge from the teacher. A smartboard or tablet and their application is similar to a smart phone, to which many are familiar. Using a robot however, may be completely new territory. Children too, these days, are likely to be familiar with a tablet (at home), but a robot is something new and exciting for most of them. The robot could take up too much of their attention due to all its interactive features. There could be less attention for the subject and other social activities.

While using technology in the classroom is becoming the standard, the use of robots is still at its infancy. An example of an early initiative in this field is Digihelden [5]. This is a small company that teaches technological subjects on primary schools. Children can play, experiment and practice with modern technologies like 3D printers, interactive playgrounds, programming concepts, and also robots. They use the robot 'Dash & Dot' (image 5). This responsive and interactive little fellow can be combined with Blockly, an environment for children to learn programming concepts. The children can make the robot move, sing, dance and look around while learning about for-loops, if-statements and variables.

Image 5 The Dash and Dot robot

Digihelden is an example of using robots and other technologies as tool in the classroom to learn about the technologies themselves and to prepare the children for a future where robots may be omnipresent. The men behind this initiative are not only skilled teachers, but also very knowledgeable and enthusiastic about modern technology. In this case, it works fine, but such a combination of skills may be rare to find. Primary school teachers (mostly women these days) are generally less skilled at technology, and technical people are often not born to be a teacher. In the near future, when robots will more often be used in the classroom, this could likely become a problem. Later in this report, near future scenarios will be discussed more elaborately.

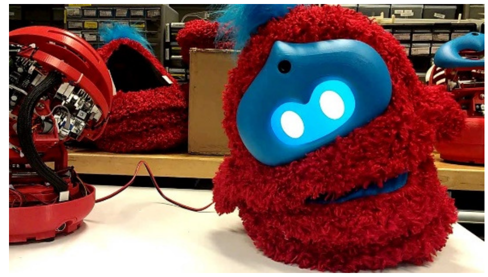


Image 6 The Tega robot

While Digihelden is an example of using technology as a tool to learn about technology, robots can also be used to learn about several other subjects. The physical presence and interactive features of certain robots make them an engaging and convincing companion. For example to learn a second language with, to learn social skills with or to learn about the environment. An example of a robot made to interact with children is the Tega robot (image 6). A smartphone forms the brain and face of the robot. The body is fluffy and can move via an electromechanical system inside. Researchers around the world are developing robots like Tega to research child-robot interaction. While technology advances and research continues, the robots are not quite ready yet to be produced and introduced on large scale on schools. This is a reason that they are not yet seen often.

Within this field of child-robot interaction, many challenges remain. For example decent speech recognition for children does not exist yet. Next to this, artificial intelligence has not yet reached the level to make robots that can fully autonomously guide a learning activity and decide an action for every possible action a child may do. This is only on the technical side. On the interaction side, there remain many questions too. It is still unclear how children perceive a certain robot: a living being, an object, or something in between. Young children (age 4) may believe that the robot lives, but could also just pretend to believe this during their play. How do children with autism perceive a robot? How should a robot respond when the child is bored, or does not understand the activity? How should a robot sense this in the first place? All these questions need answers in order to design an effective and ethically correct learning activity for children with an autonomous social robot. In order to establish child-robot interaction, the child should understand the robot, and the robot should understand the child. If this is not the case, miscommunication, annoyance and many other undesirable effects could occur which disturb the activity and the learning process.

To summarize, having technology in the classroom is becoming the standard. This has several advantages as well as disadvantages. On the one hand, there is instant internet access and lessons can be made interactive, exciting and specified towards a specific child. On the other hand, the teacher will have to learn more about these technologies in order to use them effectively in class. This can be a big challenge. Also the gap between the rich and the poor may increase. Robots as teaching tools in the classroom are under development and subject of research, but not ready yet for large-scale production. Development and research takes time. Although robots share many advantages and disadvantages with modern technology, they have the potential to be used as tool for education for a variety of subject. Digihelden is an initiative by people who are skilled teachers and have knowledge about technology. They teach children about robots and other modern technologies.

The near future – social robots are becoming part of primary education

In the near future, about fifteen years from now, problems like the above-mentioned could be solved or solutions could have been improved. Armed with improved autonomy and social features, robots can enter the classroom on large scale. With robots in the near future, improved versions of Tega, or this graduation project's robot are assumed. The ones that are currently under development and being researched. These robots will have better interactive and communicative features than nowadays' robots, but will still have their limitations regarding decision making in certain situations. The fully autonomous teacher assistant or student friend may take a little longer to be developed, as will be discussed in 'the far future' chapter.

Image 7 The Toyota kirobo mini robot

It can be expected that robots will not only enter the classroom, but will show up at all kinds of places and will become more present in general. In fact, one could say that the smartphone/tablet revolution may repeat itself for robots. First, phones were used to call and text. Later, they became multi-purpose devices accessible for everyone. In a similar fashion, robots could become more accessible and multi-purpose in the future, allowing them to be used by the big public in many different ways. A multi-purpose social robot may become part of the household in fifteen years. Social companion robots are already being developed. The Toyota kirobo mini for example (image 7). When robots start to enter people's lives, and especially children's lives, some modern ethical concerns regarding robots may fade away automatically. This statement needs some explanation.

Image 8 The Nao robot

The fact is that children learn very fast. Nowadays, three-year-olds already know how to interact with a tablet, while at first, the tablet was probably not even designed for this target group. Children learn by playing, by doing, by trying. If a companion robot is at their disposal, like the Toyota kirobo mini, Nao (image 8), or the commercially available Cozmo (image 9), they will learn how to interact with it, the do's and don'ts, its possibilities. Here the parents will play a role as well. It is their task to raise the child. Contemporary sentences like "don't eat cookies above the keyboard!", "don't grab the dog by its tail!", "first wash your hands before you touch the tablet!" can easily exist for robot interaction as well. One big assumption is made here: robots will enter households at the same time as they enter classrooms. This could be likely to happen, given that robot development progresses gradually and at some point, people will start to buy them for all kinds of purposes.

An ethical concern of these days, as pointed out by S.G. Tzafestas in [7], that could turn out to be less impactful is the awareness of the robot's capabilities and function, to a child. Better feedback from the robot, combined with pre-knowledge from the child could already take away possible misuse, disappointment or misunderstanding. For example, if the child knows that the robot has a certain role, and is not equal to a teacher or other human being, the child may less likely expect things from the robot that it cannot do. So even though the robot is designed to mimic human behavior, the child could be aware of its specific function and limitations, and still be engaged. This understanding can come forth from previous experiences with robots, at home for example. However, here also the teacher has a role to inform the children about the robot's role and capabilities.

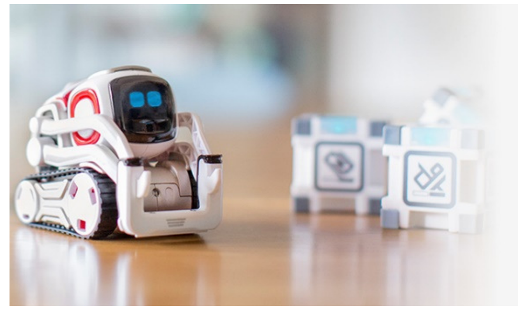


Image 9 The Cozmo robot

Robots with a cute look and interesting behaviour could make children attached to the robot, after they interacted with it several times over a longer period. This could be harmful when the robot breaks down, is being removed or the child is not allowed to play with it anymore. Especially when the reason for removing it is not clear, stress could be the result. Also this may be less likely to occur when the child understands that the robot is a tool for education and not a friend for life. It is important however, to allow children to say goodbye to a social robot if they want that. Sudden loss of a friend is never pleasant. One could also argue that dealing with loss is a good life lesson for the child. It is not always fun, there are times of goodbye as well. Also here, there is an important role for the teacher as person between the child and the robot. A social connection should be established, and even so be stopped via proper guidance by the one in charge.

The one who is in charge does not always need to be the teacher. The robot could be designed to have some authority over the child as well, for example to improve structured education. The issue arises who controls the duration, difficulty and type of activity. For example if the child wants to stop because it is tired or needs to go to the toilet. The robot should have been designed to allow this up to a certain extent, or it should be plausible that the teacher interrupts and stops the robot. Here the role of the teacher can be discussed. The robot should be related in such a way to the teacher, that the children listen to the robot, but still understand that the teacher stands above the robot. The teacher should be able to control the robot while preserving the authority image that the children have of the robot. If not, children may misbehave and not listen to it. A robot could convincingly be introduced as the teacher's assistant or friend, for example.

As may seem obvious from the above text, the role and requirements of the teacher will change. The tools that the teacher will have at his/her disposal will gradually become technologically more complex. In the current state, teachers may already struggle with tools like a smartboard or a tablet. In fifteen years, autonomous social robots like the one from this graduation project, Toyota's kirobo mini or Cozmo can be used as tool for education in the classroom. In the first place, the teacher should be aware of the (long term) effects that a robot can have on a child, the capabilities and limitations of the robot, and the role of the robot with regards to the teacher him/herself. In the second place, the teacher has to be able to solve (small) technical problems that might occur with the robot, so the dependence on technology has less chance to turn out negatively. Examples of small technical problems could be that the robot does not understand the child's action, gets stuck internally on the software side, or some sensor does not respond well. The teacher can try to solve these problems by resetting the robot, for example, or turn him on and off, or go to a "teacher" program/interface (as part of the robot software package) where he/she might identify or report the error, or apply other knowledge about robots. While doing this it is important that the image/fantasy/role-model of the robot is preserved for the child, in order to keep the child engaged later. The teacher could say that the robot needs a break, accidentally got hurt, or does not understand it for a moment. Something that may break the child's fantasy is taking apart the robot in front of the child. If the robot face is made with a tablet, that is masked by the robot's "fur, or skin", it would be strange for the child to suddenly see that the robot face is just an app on a tablet that can be taken apart from the body. Similar kinds of debugging should happen behind the scenes.

In fifteen years, not all teachers may be equally familiar to robots and how they work just via own experience with robots. Before robots can be used, teachers themselves should be taught. As mentioned above, the social impact as well as technical principles have to be explained. While children may see a robot as some kind of "alive" or as friend, the teacher should see it as a tool, and stand above the robot-child interaction in that sense. Basic notions that a robot gathers information via sensors, processes it, and can execute a certain number of tasks based on the input, should be clear. In addition, knowing things like sensor range and other sensor characteristics, possible robot behaviours and limitations, can all help the teacher to understand why a robot sometimes will not respond in the way a child expects or wants it.

Even after a teacher is being taught the principles of robots and child-robot interaction, he/she could still struggle with, or be insecure about, the modern technology. One could say that these people will have a disadvantage, and therefor may be less likely to get a job as teacher. This could hurt education, however, as these people could be excellent teachers without technology. On the other hand, you could argue that having technological knowledge is simply a vital skill in the future, when technology may be everywhere. If a teacher does not have this, it will not only have a disadvantage as teacher, but also in other situations in life. Not being able to handle technology could become a widespread disadvantage which may reduce the chance on a job in general, not only as teacher. In this case, people will simply have to go with the flow and learn about these new technologies. An example of these days: When you do not have a smartphone, because you are a bit unsure about data, privacy, and how it all works in general, you could feel isolated when all your friends and family do have it. Events are being planned, funny videos are being shared, all via the smartphone. Even though you do not like it, this could be a reason to accept the situation and just buy one and figure it out. In short, robots will enter the lives of people in many ways, not only in the classroom. This way, less technical people may be "forced" to go with the flow and learn the basics of robots anyway.

When robots become regular devices in people's lives, similar to smartphones/tablets these days, the extra effort that a teacher has to put in, to learn about a specific robot used at school, may be not much. So in the end, the requirement for a teacher of having technological knowledge may turn out to be less of an issue, given that people will be more surrounded by technology everywhere. A teacher of the future that learns about a specific robot at school could be compared to a teacher of these days that has to get familiar with a specific app used at school on the tablet. Here also the role of the robot developer comes into play. It is their task to understand their target group and to build a user friendly robot for the child, but also for a teacher who has to be able to manage the child-robot interaction.

A concept like Digihelden may be a good solution as well. Especially in the period when robots can be used for education, but do not yet show up everywhere. This transition period could be between now and fifteen years. This concept works in a similar way as the gymnastics teacher on the primary school. One or two people organize gymnastics lessons for each class, throughout the day or week. The normal teacher, who may be less suited for these wild and acrobatic activities, can have a break while the expert on this area takes over. This is exactly how Digihelden works, but than for technology. This solves the problem of the normal teacher's lack of technological knowledge, and still the children learn about, and with, robots. During these lessons, the normal teacher can observe, or just participate him/herself, and learn something new as well. Besides this, it is likely that teacher education will adapt to the new tools that will be available for the future teachers. Lessons on child-robot interaction could fit well in the curriculum, as well as technical lessons. Also this could be done by people like Digihelden.

To summarize, in the next fifteen years it is expected that robots of increasing complexity will gradually show up at all kinds of places in people's lives. This can be learned from the evolution of the (smart) phone and tablet. The classroom will just be one of the places where this happens. Due to this, children as well as teachers, will gradually grow more familiar with robots. Children will learn more about the robot's role, capabilities and how to interact with it. This can happen for example at home, they can learn from their parents. Teachers (being adults) will learn more about the technical side, but also about child-robot interaction. A concept like Digihelden could grow in popularity, because they can educate and support teachers on the technical field, and also educate the children in special technology lessons. Modern ethical concerns on child-robot interaction may fade away as complex social robots will gradually enter children's lives and not suddenly be there. It is in the nature of a child to learn and adapt fast. It is the responsibility of the parents and teachers to be well informed on child-robot interaction and basic robot technology. It is the responsibility of robot developers and researchers to design user-friendly robots, for the teachers as well as for the children.

The far future – autonomous social robots as part of primary education

In the far future, which is more than fifteen years from now, autonomous social robots could grow in complexity to a degree that they can take on other roles than 'tool' in the classroom. Mainly because of continuous progress in the fields of artificial intelligence and child-robot interaction, this would be possible, technically and socially. While in the near future social robots can be used in certain scenarios (like an English language lesson, or math), in the far future they could be permanently present in the classroom. This will be possible because these complex robots will be able to understand many different kinds of incoming sensor data and speech and act accordingly. The robot could be framed as assistant of the teacher. A possible task could be to make sure all children pay attention, or do not cheat on a test. For example, the robot could supervise the classroom while the teacher has to educate a child with special needs. The other way around could also be an option: the robot plays with/teaches the child or children with special needs, while the teacher is busy with the rest of the class.

In the same fashion as finding information on the internet is an important skill these days, in the far future, communication with a robot may be an important skill. An interesting topic would be a technology lesson, given by a robot itself. The children could be educated by the robot, and interact with it in this way. The robot could explain about his own sensors, inner workings and components. In a certain way this can be compared to a cultural lesson from someone of a totally different culture. For example, a guest lesson from a Japanese who explains the Japanese norms and values. To continue on different cultures, a robot could be introduced as Englishman in the classroom. This robot would be able to understand Dutch and English, but it can only speak English. It could be used to practice English with. As described by Chang et al. in [8], children are not as hesitant to speak to robots in a foreign language as they are when talking to a human instructor. In addition, students can practice with the robot as often as they want. The robot will not get bored or run out of time, unlike a human instructor.

Another interesting property of advanced robots in the classroom is their data collection, as pointed out by J. Werfel in [9]. In a role as tutor, the robot could collect data about individual children's (learning) behaviour. After some time, the robot could analyze and learn from this data and adjust its tutoring behavior towards the specific needs of the child. Behavioral data could also be analyzed to identify patterns amongst children and the robot could adjust itself to a child as a certain type of learner. A robot could identify difficult subjects for a child and go more into detail on those subjects. In other words, a robot can adjust itself as tutor to improve learning for the child. The robot could develop a profile for each child and adjust its behavior per child. This can also be used for the above-mentioned robot roles as teacher assistant.

However, the more user-specific data a device collects, the more important it is that the data is stored in a safe place. No child (and no parent) wants its scores and behavioral data to be in the hands of other people without permission. Besides that, it may be better that the child itself does not know how it is being classified by the robot. It may be demotivating to learn that your robot companion classifies you as "exceptionally bad at math, requires much extra attention", for example. On the other hand, it is possible that the general opinion on data collection and privacy changes drastically in the far future. Available data keeps growing, and in the far future there will be many robots collecting personal data. Consider for example a robot companion for households. A robot like that may know exactly the weekly patterns when people are home and when not, conflicts in the household, and secrets. The amount of data being collected by Google for example, can be called ‘scary’ as well. Still many people agree to their end user agreement while not knowing exactly how their data is being used and by whom. “Everyone does that, everyone has Facebook/Google, I guess it is safe then…”, is the argument. It is likely that robot data collection is perceived in the same way in the future. Just as with modern technology, it is the responsibility of the robot developer to make safe software systems that are hard to get hacked into. This is also very much in the developer’s own interest because people will not buy robots/products in general that can easily be broken into and hacked.

Talking about the teacher of the far future, again things will change. First off, it is very unlikely that a social robot will completely take over the job of teacher. As mentioned before, human-human interaction is vital in a child’s development, possibly even more so in its early years. Besides that, human teachers have often many years of experience and have dealt with countless situations regarding children and education. For a robot to reach this level autonomous knowledge and decision making capabilities, we would go beyond the far future and into the unforeseeable future, which is extremely hard to predict for.

The process of the near future may repeat itself. Teachers will have to expand their technical knowledge in order to teach in a changed environment. However, again robots everywhere will gradually become more complex and it is again the responsibility of the robot developer to deliver user-friendly products. One could also argue that in the future, the self-repair capabilities of robots may improve, and that technical knowledge will be less needed. The robots may be able to “survive” on their own.

Given the presence of these advanced robots, it may become true that important child-teacher and child-child interaction will occur slightly less. How bad this actually will be, is debatable. In a world where robots and humans live together, the classroom should represent the same environment and prepare for the future. In that sense child-robot interaction in the classroom may become a very important sort of communication too.

To conclude, in the far future highly advanced robots could take on several roles in the classroom. They could be able to communicate naturally and adapt themselves based on user-data. Data collection is debatable but it may be perceived in a similar way as data collection of smart devices these days. As robots will gradually grow more complex, children and teachers will have to adapt gradually as well. However, robots will grow more autonomous and may require less error fixing by the teacher. Children may spend more time with robots and less with other children or the teacher, but this will represent the outside world of the far future in which communication with and understanding of a robot is an important skill.

References

[1] P. Lin, K. Abney, G. Bekey. (19-01-11) “Robot ethics: Mapping the issues for a mechanized world” in *Artificial Intelligence* volume 175, Issues 5-6, April 2011, p. 942-949 [online] https://www.sciencedirect.com/science/article/pii/S0004370211000178
[2] C. Bartneck, J. Forlizzi (2004) “A Design-Centred Framework for Social Human-Robot Interaction” in *Proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication* Kurashiki, Okayama Japan September 20-22,2004. P. 591 – 594 [online] https://ieeexplore-ieee-org.ezproxy2.utwente.nl/stamp/stamp.jsp?tp=&arnumber=1374827
[3] M. van den Bogaerdt (21-12-2016) “Gemiddelde groepsgrootte nagenoeg stabiel” [online] https://www.vosabb.nl/tag/groepsgrootte/
[4] M. Rosso (27-04-2016) “Advantages and disadvantages of technology in the classroom” [online] https://www.linkedin.com/pulse/advantages-disadvantages-technology-classroom-mike-rosso/
[5] J. van de Hout, R.J. Pastoor, C. van Laarhoven https://www.digihelden.nl/
[6] A.M. Rosenthal, C. Strassmann and N.C. Krämer. (2016) “Robots or Agents – Neither Helps You More or Less During Second Language Acquisition” in *International Conference on Intelligent Virtual Agents* p. 256-268 [online] https://link.springer.com/chapter/10.1007/978-3-319-47665-0_23
[7] S.G. Tzafestas, “Ethical issues of socialized robots*”* in *Roboethics. Intelligent systems, control and automation: science and engineering volume 79.* Switzerland, Springer international publishing, 2016, pp. 118 – 121
[8] C. Chang, J. Lee, P. Chao, C. Wang and C. Gwo-Dong (2010) “Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school” in *Educational Technology & Society*, vol. 13 n. 2 p. 13-24 [online] https://eric.ed.gov/?id=EJ895653
[9] J. Werfel. (2014) “Embodied Teachable Agents: Learning by Teaching Robots” in *New Research Frontiers for Intelligent Autonomous Systems* at 13th International Conference on Intelligent Autonomous Systems, Venice, Italy [online] http://people.seas.harvard.edu/~jkwerfel/nrfias14.pdf