# Evaluation

## Project description

The system is designed for kindergarten children, ages from 5 to 6. Its purpose is to provide an interactive way of understanding a new story. When using the system, users shall understand the message of the story and moreover, gain knowledge in several domains like: mathematics, natural sciences, vocabulary, grammar.

## Requirements summary

The system has four main requirements. Ordering them by priority, it results in the following list:

* Understand the story
* Learn grammatical constructions explained in context
* Expand vocabulary
* Practice mathematics.

All system features should be done in an interactive way, so that children are kept entertained. The story telling needs to keep children engaged. This can be achieved by using sound effects, like imitating animal voices from the stories or stop briefly once or twice and ask rhetorical questions.

New words shall be introduced given similar examples, synonyms. They should be explained in detail. After the words lesson, children should be evaluated through a game, to make sure they have a clear understanding of the new words.

The grammatical constructions should use words from the story, not to lose the main purpose of the system, to learn a new written story in a fun and exciting way.

Practicing mathematics will be done taking into consideration children abilities to understand mathematical concepts.

In addition to functional requirements, there are a few important usability requirements. Given the system is going to be used by children themselves, it should have appealing images and colors. They should be enough to engage them but not to disturb them and make them unable to focus. The system shall provide intuitive navigation, suggestive icons and audio since children are not able to read.

The system must be consistent. One action should mean the system throughout the whole application. This way, each child will quickly get familiar with the application. Also, the system must give appropriate feedback for each action done by the user.

## Overview

In the evaluation of the system three techniques are used:

* Heuristic evaluation
* Real users evaluation
* Accessibility evaluation

The heuristic evaluation was done by the kindergarten teacher. This evaluation was not done on the final version of the system but on an intermediate version. Given the resources available, the evaluation was done at the kindergarten, where the kindergarten teacher interacted with the application, and also asked for the opinions of some of the children. This evaluation was needed in order for us to know in what direction to continue with the development.

The evaluation done with the real users, in our case kindergarten children, took place at the university. We needed a quiet place where single users or pairs of users would interact with the system and provide feedback. Together with the team and the children in the room was also the professor. In order to receive feedback, the children were asked to rate the application on a scale from 1 to 5, 1 meaning they did not like the application and 5 meaning they enjoyed it. Besides this, we also observed the interaction between the children and the application because this is a way of capturing immediate feedback to different actions that children would not be able to provide afterwards.

The accessibility evaluation was done both automatically and manually. For the automatic evaluation we used Chrome extension of WAVE Evaluation Tool. It facilitated web accessibility evaluation by providing a visual representation of accessibility issues within every page. Because no automated tool can tell if a page is accessible or not, we evaluated our application manually too, by using WCAG 2.0 checklist.

## Heuristic evaluation

The heuristic evaluation was done by the kindergarten teacher at the kindergarten. An intermediate version was presented to the teacher by the team while children were also present given that this was the only way of conducting the evaluation. Children were asked to focus on something else but nevertheless, they were attracted by the sounds and images presented in the application.

During the evaluation, the kindergarten teacher interacted with the application and provided us with instant feedback about whether the children will be able to understand the concept in the way they are presented, the navigation, the media contained in the application.

Given the information provided by the teacher we learned the following:

* Children may not understand the concept of ‘big’ when presented with two animals that on screen are more or less of the same size; a way to rephrase would be to use the word ‘large’ instead
* Children must be forced to learn the story before any games could be played since all the games are based on the story
* Any aspects of the story need to be presented in the order of the initial story
* The audio media presented in the application should not differ from one recording to another. For instance volume, voice, speech style should be consistent.

After the evaluation from the teacher, we learned that we needed to present the story in a totally different way. Initially not all the story was included and children could pick at random what parts of the story to listen. In the final version, the full story was presented in an interactive way ordered as the initial story. Also, the navigation to and from the story was adjusted in order for the children to not be able to play games without listening the story, and after each game they can choose what they want to do next, listen the story again or select another or the same game.

Because of how children understand concepts, we also needed to adjust some questions presented in the game to make sure children understand what they are asked to do. This also helped in developing the other games.

In the heuristic evaluation, we asked the teacher whether or not the design of the application is suited for the children because she knows the children of that age best, what works for them and what doesn’t, what they have learned and what knowledge and capabilities they possess.

## Real users evaluation

For all user-centered software products, evaluation is critical for finding out how users interact with a design so that it can be improved. However, it is important to bear in mind that methods appropriate for evaluating interactive products with adults do not necessarily work for children. For example, children with limited language skills might not provide much insight during traditional usability evaluation methods such as think-aloud or structured interviews. Also, children are more likely to give superficial answers to questions which they do not understand. They are also highly susceptible to suggestions and have a poorer recall for events from their memory.

After a small research on evaluating usability on children interactive products, we collected useful data that highlighted what children do and want. All facts gathered were taken into consideration during evaluation.

Firstly, children need control. They feel empowered when they feel they can make their own decision. This is the reason why during evaluation we let them play by themselves, without giving instructions. Also, we designed the system such that at every step, the child can decide what he wants to do next, listen to the story or choose any preferred game. These decisions also led to meeting an important usability requirement, which stated that children shall be able to use the application without needing assistance.

Secondly, children like to spend time with other children, so they usually form groups around one piece of technology. Having this in mind, we brought only one laptop for two children, so they can play together and help each other. Moreover, whatever the children are asked to do for the evaluation process, it is important to fit their time span. The younger they are, the shorter the process should be. For the evaluation, we managed to fit to 30 minutes timespan.

Furthermore, children should be assisted during the evaluation. All team was present to the process. However, we tried to have the minimum possible impact on the child and let him lead the conversation or the interaction with the product. Also, we tried to make the process fun and exciting. We started by asking questions to get to know them. To make them feel comfortable, we asked about their day, their mood. Then we moved to warm-up questions related to our application. For instance, we asked about their favorite animal, if they ever went to the zoo, if they have any pets at home or at their grandparents.

The setting for our evaluation process was favorable. The two children were excited to play games. Our team was between the first applications to be evaluated, so they were not tired yet. They were likely to develop a relaxed relationship with the team.

The evaluation took place in an empty classroom at university, so children could feel comfortable. Our teacher was also present, which was a plus, since they knew her already. For giving feedback to each application, the children had 5 options, 5 colored smiley faces, to express how much they like the games, from a sad to a happy face.

After observing them during the evaluation process, we collected plenty of useful information. The process started with them listening to the whole story by Marin Sorescu, “In gradina zoologica”. We modified this part after presenting the prototype to the kindergarten teacher. The final version included audio of the whole story, as it is originally written. They were patient, silent and very receptive. They were entertained and it was very easy to notice that they enjoyed the story.

One of our requirements was that after listening to the story, the system should allow them the possibility to listen again or play one the five games available. It is important to mention that the system contains four full interactive games and one vocabulary lesson, where users listen to a few words and their explanation. Children did not want to listen the story again, they were very excited to play games, so that they even played one twice.

Another requirement was for children to figure out by themselves how to use the application. We validated the fact that the system was well designed for this requirement, since children discovered by themselves how to listen questions again. Also, for the vocabulary games, they missed the explanation that this game is played using arrows, not mouse clicks and they managed to find out alone how it works.

The math game was designed such that the correct number of animals were painted on the screen. We noticed that sometimes, children did not listen carefully the questions, but started counting the animals on the screen to find the correct answer. As a result, we will try to find an efficient way to make them listen the question before answering.

The audio for vocabulary lesson had lower volume than the rest of the audios. The children were quiet and tried to be attentive. However, when playing the vocabulary game, they did not remember the word “ursuz”. As a result, we will change the audio to be consistent.

We managed to finish the evaluation process in about 30 minutes, when children started to feel tired. The timespan fitted perfectly the evaluation of the whole application.

At the end, after playing all games, they concluded the one with habitats was the most difficult one. The only game they played twice was the math game.

## Accessibility evaluation

The accessibility evaluation was done automatically, using the Chrome extension of WAVE Evaluation Tool. It provides visual feedback about the accessibility of the web content by injecting icons and indicators into the page. Although it cannot say if the application is accessible, it is still useful for error identification. This tool indicates accessibility errors that need to be fixed and highlights accessibility features that probably improve accessibility of the application.

Applying the Wave Evaluation Tool to each page of our application, there were found a few errors. They all were referring to missing alternative texts for some images. No other types of errors regarding accessibility were found. The tool also signaled some alerts that can improve accessibility. They referred to the fact that some actions do not appear to be accessible to both mouse and keyboard users. For the application to be fully accessible, interaction should be device independent. In our case, some games require just mouse interaction (e.g. the habitat game) or just keyboard interaction (e.g. the vocabulary game).

Wave Evaluation Tool includes many checks for compliance issues from Section 508 and WCAG 2.1 guidelines. Because it cannot check all of the issues in these guidelines, we have done a manual evaluation too, verifying if WCAG 2.1 checkpoints were respected in our application. We discovered that not all audio outputs were found in text format too. Also, for audio that play automatically for more than 3 seconds, there is no available mechanism to pause or stop the audio, neither a mechanism to control audio volume.

Also, there were aspects covered by this tool, which needed a human evaluation. For example, Wave reveals the existent alternative texts, but we had to ensure that they were equivalent and appropriate.

## Summary

The advantages of the evaluation are that we get to really learn what works and what does not for the users that the application is intended for. Whatever we can imagine without knowing for sure was easily disproved in the heuristic evaluation done by the teacher. It is also very fulfilling to see children use and actually like the application that we made for them.

Knowing more about the evaluation process helps us better organize and be more prepared when giving our system for an evaluation. With more resources the evaluation can be done with more organization, in more steps and using more techniques.