A Voice-Enabled Game Based Learning Application using Amazon's Echo with Alexa Voice Service

A Game Regarding Geographic Facts About Austria and Europe

Leonardo Bilic, Markus Ebner, Martin Ebner (✉)
Graz University of Technology, Graz, Austria
`martin.ebner@tugraz.at`

Abstract—An educational, interactive Amazon Alexa Skill called "Oster-reich und Europa Spiel / Austria and Europe Game" was developed at Graz University of Technology for a German as well as English speaking audience. This Skills intent is to assist learning geographic facts about Austria as well as Europe by interaction via voice controls with the device. The main research question was if an educational, interactive speech assistant application could be made in a way such that both under-age and full age subjects would be able to use it, enjoy the Game Based Learning experience overall and be assisted learning about the Geography of Austria and Europe. The Amazon Alexa Skill was tested for the first time in a class with 16 students at lower secondary school level. Two further tests were done with a total of five adult participants. After the tests the participants opinion was determined via a questionnaire. The evaluation of the tests suggests that the game indeed gives an additional motivational factor in learning Geography.

Keywords—Game-based learning experience, geography, educational, interactive, voice enabled, speech assistant, Amazon Alexa

1   Introduction

Through the constant progress in artificial intelligence and especially regarding voice-enabled services it seems like every day the possibilities for new applications in regard to speech-controlled systems would increase. Service providers like Amazon's Alexa, Apple's Siri, Microsoft's Cortana and many more constantly get new updates and with those new features. Why should we stop at simple controls like asking for a weather report, making a phone call, buying groceries online or other services such voice assistants provide at this very moment? This question led to the idea and further to the development of the "Österreich und Europa Spiel / Austria and Europe Game". The motivation was to build a fully speech-controlled game with which the audience might be able to learn and enjoy the Game based learning experience.

As stated by both Malone and Plato [1] the idea to use games to improve the intrinsic motivation of students within the learning process is not new. Malone started al-

ready in 1980 with his previously mentioned PhD thesis. Today we face different challenges which legitimate the work on this topic. As Böckle et al. fittingly states [2]:

"Today's challenge follows the interests and creativity of an individual student. One of the most interesting fields in this research is Game Based Learning (GBL), which is very similar to Problem Based Learning (PBL), where a specific problem scenario is embedded within a play framework (Barrows & Tamblym, 1980). Despite the widespread recognition of the advantages of using games in elementary and secondary education as well as in higher education, little evidence can be found on the use of digital and/or online games." [2]

There are not many games which are widely popular known for voice-enabled devices and if there are such they are mostly restricted to a certain age or by only having one language. It was intended to be multilingual from the beginning and was implemented in German first and English afterwards. First it was intended to have a narrow target audience regarding age however the final decision was to make it available and interesting for everyone.

2 The Game

2.1 Concept

At first the Amazon Alexa Skill was analyzed and designed. The game should be similar to other Trivia, Quizzes and as such an interactive game based learning experience. Malone states further [1] that there are different types of factors which need to be fulfilled in order to provide an intrinsically motivating environment. The factors are:

- **Challenge:** Basically, the challenge within the game can be both extrinsic as well as intrinsic. A simple comparison between intrinsic and extrinsic motivation would be a competitive Multiplayer community versus Singleplayer experiences which a player plays only for his or her own enjoyment.
- **Fantasy:** Malone describes this as the ability of a theme to embody or encourage using one's own fantasy.
- **Curiosity**: Novelty, complexity, surprisingness and incongruity are just a few concepts which Malone states here.

Not all aspects were implemented to all user's satisfaction however, the overall feedback was rather positive.

Additionally, it has to be mentioned that this should be a Game Based Learning experience and as such a main goal was to make sure that the user not only plays the game but also learns about the subject while doing so. It does not matter if the user answers right or wrong after the question he or she will get additional information on the asked question. The questions for the game were formed by research in a Geography book [6] as well as in an Atlas [5].

A question that arose during the planning phase was what would legitimate the game compared to other already existing games and possibly other researches. The main points are listed below:

- **Focus on Austria and Europe**
- **Game modes**: There are different game modes. See section 2.3 Game modes for further details.
- **Visual feedback**: Another feature implemented is the possibility to visualize the audio input and output. Thus the user could experience the game both visually and audible.
- **Multilingual experience**: One of the most important aspects is that the game is multilingual and through that more accessible to a broader audience.
- **Trend on digitalization in education in Austria**: As stated in the "Nationaler Bildungsbericht 2018, Band 2. Fokussierte Analysen und Zukunftsperspektiven für das Bildungswesen" in the chapter 8 "Bildung im Zeitalter der Digitatlisierung"[4] more and more primary schools already use computers for E-Learning.

Although one might think that multilingual support might not be of such importance one has simply to look at localisation in commercial video games. As Bernal-Merino states in his book [3] "Translation and Localisation in Video Games":

"The game publishing industry is slowly realising the crucial part that the localisation of multimedia interactive entertainment software, a.k.a. game localisation, plays in boosting sales globally, opening new markets and expanding franchises. Nonetheless, some companies (developers and publishers) still seem to be unable to fully integrate best localisation planning and practices into their workflow, and academics conducting research in this field are also thin on the ground which does not help to improve the situation."

As it gains more demand to be able to use a variety of information and communication technologies and voice-enabled technologies are due to the fact that they are quite new a niche market the importance of the project and its research is assured.

2.2 Technical background

The service itself was implemented self-hosted. There is the possibility to implement applications with the Amazon Web Service (AWS). However the final decision was to develop the game self-hosted with the Flask-Ask framework as this provided more freedom overall.

The big picture of how components work with each other can be seen visualized in figure 1:

- The actor gives a voice command to the device. The device can be arbitrary as long as the Amazon Voice Service can be installed on it. It was tested on Android smartphones, a Raspberry Pi and the official Amazon Echo (2. Gen.).
- The device sends the information gathered to the Amazon Voice Service.
- Here the audio gets forwarded to the Endpoint defined within the Amazon Alexa Skill. This is the start of the HTTPS communication.

- The next step is the Intent Recognition Algorithm.
- Amazon's Server forwards the result then to our server. Then the internal server logic handles the intent accordingly to the implemented protocol. For example if a question was asked and the user answered then the internal logic would determine if the question was right, wrong or if the user said that he or she didn't know.
- The output from the server will be forwarded to the Amazon Voice Service.
- Here the output gets transformed from text to audio.
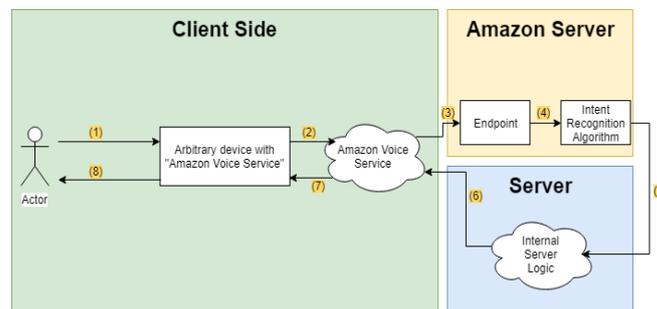- Finally, the device answers the actor on his or her initial command.

Fig. 1. Screenshot of a HTTP Communication during a game execution

As can be seen in figure 1 the third step is a transition from the client's side to Amazon's server and will be redirected later to our local server. This is done via HTTPS. Of course, HTTPS is not a perfectly secure way to communicate as it can be hacked with strategic Man in The Middle Attacks with tools like ARP Spoofing, DNS Spoofing, Sniffing and SSL Dump as Chomsiri states [7].

2.3 Game modes

The initial design intended to have two different modes. The first mode was called the "Quiz game". In this game mode the user would get a question and four possible answers with one of which the user had to choose of, e.g.: "How many states has Austria?" with four possible answers.

The second mode was called "Relations game". As the name already suggests this game gave the player two objects in relation to each other and the player had to figure out on which a certain adjective applied. An example for a question would be "Which lake is bigger?" with two lake options given.

The different game modes satisfy different of Malone's previously mentioned motivational factors with different weight.

The questions for the different game modes and their answers were stored in a XML file, which the server reads upon its setup phase.

2.4 Evaluation

The game was first tested in a secondary school where two teachers and 16 students tested the application. The students were divided in four groups of three students and one group of four students. Later three adult subjects tested the game at the Graz University of Technology. The under-aged subjects tested the German version as it is their native language. The adults however tested the application in English. A vocabulary sheet was provided in case some words might be unknown to the testers as English was not their native language. However, afterwards nobody stated that the vocabulary was a problem. Nobody was allowed to use the visual assistance provided by the cards that were implemented. As such the whole test was only perceived via hearing. Before the actual test started a disclosure was given that none of the names or the individual results will get published and for the students also that they will not be graded.

Before the interview every participant was asked the questions he or she had answered wrong again to see if they paid attention to the information given by the device afterwards. The evaluation had six statements and each had to be answered in a scale with points from one up to five where five is the highest score of approval and one the lowest. The groups of students evaluated together and had to discuss internal which the final score they wanted to give to certain statements was. The adults however all evaluated individually. The result of the evaluation can be seen in table 1 for the student groups and in table 2 for the individual adults who participated.

Table 1. The evaluation results from the students

Statement	Group 1	Group 2	Group 3	Group 4	Group 5
The application was easy to use	5	5	5	5	5
The game was fun to play	5	4	4	5	3
I could imagine playing the game in my free time	4	5	3	3	2
The game's questions were easy to answer	5	4	3	5	5
I would like to play the game again	5	4	5	5	4
I have learned something while playing the game.	3	2	3	4	3

Table 2. The evaluation results from the adults

Statement	Adult 1	Adult 2	Adult 3	Adult 4	Adult 5
The application was easy to use	5	5	5	4	5
The game was fun to play	4	5	4	4	4
I could imagine playing the game in my free time	4	5	3	2	1
The game's questions were easy to answer	2	4	4	3	4
I would like to play the game again	4	5	2	3	2
I have learned something while playing the game.	4	5	4	4	3

3 Conclusion and Future Work

After completion of all those milestones many insights were given on what can and should be done in the future of the project as it is crucial to keep working on the ap-

plication. One aspect which needs to be tested is if the server is able to operate on different operating systems as well. This however is not as big as a priority. At a closer look it seems that Amazon itself has some bigger problems with their voice-enabled devices and services. There are a couple of problems:

- **Wrong evaluation**: As many participants of the evaluation phase commented Amazon understood their spoken words wrong although they did speak loud and clear in a calm environment.
- **Accelerating speech speed bug**: The issue with Amazon suddenly speeding up the given speech was not reproducible. It occurred on different sentences and sometimes it did not occur at all.
- **Overall speech speed**: One has to mention that it is unfortunate that the speed of spoken words of Amazon's Alexa cannot be regulated.
- **No possibility for interactivity with visual feedback**: The cards provided by Amazon give the user a possibility to both have an audio and visual feedback experience. However there is left unused potential as it is not possible to interact also with the visual feedback. An example for using it would be that during the game the user could look and interact with a map.
- **Intent recognition:** The intents get recognized by Amazon at their server but there is no way to observe on how this is done which makes it rather difficult to work on unwanted behavior.

The experience overall regarding the program's and with that the server's logic in Python using the Flask-Ask framework was very pleasant as it provided a simple to use framework with almost no problems at all. One has to mention however that it is very unfortunate that one cannot access the language of the device connected. For the future it would be definitely of interest to implement more different subjects as this was requested by some external testers. Further should be analyzed if other voice service providers would be more adequate as Amazon's Alexa proofed to have a lots of issues or to use a self-implemented voice-enabled service self which upon the game would be built on. Another feature which might be interesting is to save the data of users in a database such that not only the current session is used for the game as this information is volatile. Amazon's Alexa does not provide for Skill developers to differentiate between persons upon their voices which technically should be possible. It would also be of interest for the project's good to test if a combination of both voice and more classic input technologies like mouse and keyboard would increase the overall satisfaction of users.

4 Acknowledgement

5 References

[1] Malone W. Thomas. (1981) "What makes things fun to learn? A study of intrinsically motivating computer games." https://doi.org/10.1145/800088.802839

[2] Böckle Martin, Ebner Martin, and Schön Martin. (2007) "Game Based" Learning in Secondary Education: Geographical Knowledge of Austria."

[3] Bernal-Merino A. Miguel. (2014) "Translation and localisation in video games: Making entertainment software global." Routledge. ISBN: 978-1-3157-5233-4 https://doi.org/10.4324/9781315752334

[4] Brandhofer Gerhard, Baumgartner Peter, Ebner Martin, Köberer Nina, Trültzsch-Wijnen Christine, and Wiesner Christian. (2018) "Bildung im Zeitalter der Digitatlisierung." In: Nationaler Bildungsbericht Österreich, Band 2. Leykam. ISBN: 978-3-7011-8118-6

[5] Hölzel Eduard. (2015) "Grosser Kozenn-Atlas." Hölzel. ISBN: 978-3-85116-607-1

[6] Mayrhofer Gerhard, Posch Robert, and Reiter Isabell. (2015) "GEOprofi – Geographie und Wirtschaftskunde für die 5. Schulstufe. Vol. 6" Veritas Verlag. ISBN: 978-3-7058-8415-1

[7] Chomsiri Thawatchai (2007) "HTTPS hacking protection." https://doi.org/10.1109/AINAW.2007.200

6 Authors

Leonardo Bilic is an Austrian Computer Science Master's degree student at the Graz University of Technology. Besides he works part-time as both a tutor at TU Graz and as an intern at NXP Semiconductors.

Markus Ebner, is currently working as a Researcher in the Department Educational Technology at Graz University of Technology. He deals with e-learning, mobile learning, technology enhanced learning and Open Educational Resources. His focus is on Learning Analytics at K-12 level. In addition, several publications in the area of Learning Analytics were published and workshops on the topic were held.

Martin Ebner, is with the Department Educational Technology at Graz University of Technology, Graz, Austria. As head of the Department, he is responsible for all university wide e-learning activities. He is an Assoc. Prof. on media informatics and works at the Institute of Interactive Systems and Data Science as senior researcher. For publications as well as further research activities, please visit: http://martinebner.at. Email: martin.ebner@tugraz.at