Abstract—This paper presents a game with the aim of developing persistent knowledge inside a Cultural Heritage exhibition. The authors developed a game called Solis Curse, that was used as a way of testing the knowledge acquired during the museum visit. Similar to an entertainment quiz, the game presents a set of questions with progression levels of difficulty. Through the analysis of the responses and some others qualitative and quantitative metrics such as the difficulty level and time, the players score is calculated and compared with a global ranking ladder. Beside putting to test the visitors knowledge, the game also offers the possibility of speech interaction with an embodied conversational agent that steers the game and helps the user to successfully complete the task at hand. Based on interdisciplinary multimedia technologies and cultural contents, Solis Curse supports an entertainment platform for a younger audience, which may be regarded as a mean to convey cultural heritage contents, conducted by the motto of learning by playing. The results of this research point out the benefits of having game as an aid to the learning process and also shows the impact of multi-modal interaction on the user immersion.

I. INTRODUCTION

Computer games are nowadays an important part of most childrens lives and a increasingly part of or culture as a whole, this factor aroused the interest of the academic community regarding games and their applications, since they had proven to be motivating activities whose applications can go far beyond simple leisure [1]–[3]. Several studies point out the benefits of using games as a way of supporting learning and knowledge transfer. Serious games can give teachers and parents the ability to reach students where they live, bypassing many of the challenges associated with restructuring the education system from inside out [4].

Through the use of computer games its possible to enhance experiential learning, inquiry-based learning and self-efficacy, several authors [5], [6] defend that the use of these paradigms in computer games can produce a significant increase of learning outcome, compared with traditional methods. The key factor involved in the development of knowledge-games is the design process, games should be designed considering pedagogic aspects, so they can be integrated as a motivating element of the game [7]. As Malone said in [8] games should follow three main design principles: challenge, fantasy and curiosity. Based on these guidelines, this paper presents a case study where the impact of a game in the development of persistent knowledge inside a Cultural Heritage exhibition was analysed. The game Solis Curse was created as part of a cultural exhibition related to Santa Engracia church, National Pantheon, in Portugal. The main goal of the exhibition (with the theme The Building Works of Saint Engratia: the Pantheon under the Republic) was to show the vissitudes of this monument, which took 400 years to construct. The game was used as a way of testing the knowledge acquired during the exhibition visit. Similar to an entertainment quiz, the game presents a set of questions with progression levels of difficulty. Through the analysis of the responses and some others qualitative and quantitative metrics such, as the difficulty level and time, the players score is calculated and compared with a global ranking ladder.

Beside putting to test the visitor knowledge, the game, implemented in Unity 3D, also offers the possibility of a different user-computer interaction. Since speech is the most natural way of humans to communicate, we endow an Embodied Conversational Agent - ECA - modelled in 3D Studio Max, using automatic speech recognition (ASR) and speech synthesis engines (TTS) and a spoken dialogue system based on natural language processing techniques. We provided the ECA with personality, oriented to establish empathy towards specific audiences, allowing an easier transfer of civic values, like the importance of maintaining our cultural heritage, while strengthening their educational purposes. At this stage the use of these techniques is fairly simple, but with large possibilities of evolution in the next releases. Based on interdisciplinary multimedia technologies and cultural contents, Solis Curse supports an entertainment platform for a younger audience, which may be regarded as a mean to convey cultural heritage contents conducted by the motto of learning by playing. The results of this research point out the benefits of having game as an aid to the learning process and also show the impact of multi-modal interaction on the user fulfillment. In next section we will present a game overview and in section 3 the different components of the game. Section 4 presents a user evaluation of the game and in section 5 some conclusions are presented.

1http://unity3d.com/
II. GAME DESCRIPTION

This game is results from the intersection of several areas of knowledge, such as cultural heritage diffusion, multimedia game technologies and speech and natural language processing techniques. In this section, an overview of the game will be described.

The construction of the Church of Santa Engrácia, currently Portugal's National Pantheon, lasted for more than 400 years, taking many building campaigns over the years to completely finish it. This led to the common belief that this church was cursed. This was mainly due to the story of Simão Sílis, a New Christian unjustly accused of vandalizing the church on the 16th century, who fated that the temple was to never be completed.

The Solis' Curse goal is to build a 3D model of the Pantheon, divided in five different levels. Each level represents a different question, and answering it correctly will get the player through to the next level, revealing a new piece of the model. As mentioned before, there are five levels, thus there are five questions to answer. Following Malone’s guidelines [8], the questions have an incremental level of difficulty to immerse and challenge the user. The model should be concluded as quickly as possible, since the time total will influence the score. In case of a wrong answer, the game ends and the players are offered the chance to replay the game from the beginning. A set of three different lifelines is always available on a question scene, and they can only be used once. In order to engage the user into the game plot, the scenery alludes to the times of Simão Sílis. According to [9], this game falls into the category of the virtual museum, mainly because the different backgrounds are digital representations of a cultural heritage site. Alike the Egyptian virtual temple [10], there are six rooms linked by corridors. To accompany the player through these medieval rooms, there is a virtual character as a narrator. This storyteller represents a cleric, who was the judge that condemned Simão Sílis. In each room transition, which corresponds to a new level, a portion of the monument's history is told. The player does not physically exist in the game, so there is not a direct interaction with the game scenery, only with the virtual agent through the question-answer speech recognition setup. Thus, this means the player is not a character in the game. The narrator, through first and third person views, assures the programmed passages between rooms. The reasoning behind this decision was to limit the gameplay time, in order to allow the maximum possible number of players. In a freely roaming scenario, each player would take an increased amount of time exploring the scenery, not focusing on answering correctly. This game is a complement for an exhibit that receives both Portuguese and foreign visitors, therefore a bilingual setup is crucial. In the main menu, the visitors can choose between Portuguese and English. After choosing the language, there is a cut-scene (with a black background) in which the ECA briefs the player, telling him the legend behind the game. The narrator divulges the story of Simão Sílis, as an historical context to the game. The rules are also stated and the game begins. There is another cut-scene with a developed background of the monument, showing the transition between the atrium and the library. During this course the narrator was another monologue, as if he was talking to himself, speaking about the monument, so that the player has more information to answer to the questions correctly. After this, the player is confronted with the first question. The agent reads the question and four options, and only one is correct. The player was one minute to answer correctly, and he may utilise the lifelines to help him. There are three different lifelines: 50/50, which removes two wrong answers; More time, which slows the chronometer by 30 seconds; ECAs help, giving the player the right answer. The player is able to answer in a multimodal way: by touching the screen; by speech recognition, stating the answers letter (A, B, C or D), or by reading the correct answer. Giving the correct answer will trigger a new cut-scene, where a piece of the monument appears in the last room, and the ECA moves to the next area. After answering correctly to every question, the ECA goes into the final room, where the completed monument is displayed. The game ends and the player is directed to the scoreboard, where he can insert is name and view is total points and place.
III. Game Components

This section will present the different game technology components. Summing up, the main game platform was built using Unity 3D, using imported scenarios and character from 3D StudioMax. The face animation of the virtual agent was done using FACE [11]. The agents spoken dialogue platform has an ASR and TTS engines, and a QA module integrated in the game engine environment.

A. Game environment and development

Unity 3D was the chosen game engine for several reasons. Its easy to learn, mainly due to the intuitive interface and the availability of written resources and tutorials. This instinctive interface allows a good ECA animation, by creating animated routes in corridors. This program enables the user to do games with a great visual quality and realism, which lack sometimes in serious games. This software was a very good compatibility with image, audio, video, 3D modelling and true type fonts programs. Also, it is compatible with different platforms, such as Windows and Mac OS, and Android and iOS for mobile platforms. This game was presented on the exhibition using a kiosk concealing a computer with a touch screen and a microphone and speakers system. Also, it is possible to download the game on the museums website.

For this game, several important modules were created. The main module for this game, the game manager, controls the going and outgoing information data. This system controls the sub-modules, which one with a specific function. Depending on the input type, the system will react using different components. If the player answers the questions using the touch screen, the quiz module sends this data, using the game manager, to the QA component. The answer is confirmed and presented to the player on the main screen. If the user responds with speech, it is necessary to use an integrated speech recognition system, which communicates with the game manager, explained further in this section.

The module responsible for the ECA path animation is the movement module, where all the information about the length and directions that the agent needs to walk through the entire scenery are kept. The game manager knows when to trigger these animations, depending on the game progress.

As mentioned before, 3D StudioMax was used to model the ECA and the backgrounds. As Malone said in [8] the Fantasy factor is extremely important in this kind of games. Therefore, the backgrounds are medieval-themed, and every room is different. The player was an historical perspective from the organization and decoration of the rooms, which are all different. The question timing and cut scenes are designed in a way that the user was enough time to enjoy the scenery, without it becoming dull.

B. Virtual Agent

This agent is based on a spoken dialogue system, using ASR and TTS systems, 3D facial animation and natural language processing techniques. Through this multi-modal interface, based on concurrent input modes, there is a more natural interface with the game.

This virtual agent platform receives the speech input through an ASR system (AUDIMUS) [12], which sends the speech through a Language Interpretation module. Finally, the speech goes to the QA [11], module to be verified as the correct answer. The result is sent to the TTS system (DIXI) [13], that generates an audio file with phonetic and temporal tags. The phonetic tags are transformed into visemes, which are used to generate facial animations along with temporal tags.

Usually, this speech recognition system and generated facial animations run independently from the game platform. This causes the facial animation to be done offline, and not in real time. In this game, the speech recognition and audio generation were integrated in the game engine, unlike most serious games. Facial animation was not included in the game engine development, which will be altered in the future. Our goal is to generate cut-scenes and in-game scenes in which the ECA is able to have any sort of facial animation, generated in real time.

AUDIMUS is a hybrid speech recogniser that combines the temporal modelling capabilities of Hidden Markov Models (HMMs), with the discriminative pattern classification capabilities of multilayer perceptrons (MLPs). This same recogniser is being used for different tasks, based on a common structure, with different components. The acoustic models were adapted for a microphone input device. The system is using a specific language model, with a limited number of words in the vocabulary, for this task. In the future, the system will be more accurate; the language model will have an increased complexity.

The Language Interpretation module is responsible for extracting the intentions of the users utterances. As mentioned before, the user can answer a question by using the touch screen or by speech, using a module that receives an input from ASR, concurrently. When this module receives input on the touch screen, it sends the answer chosen by the player to the QA module. This module checks for the correct answer. If the input came from the ASR, there are a few language analysis steps until it finds a Speech Act, which codifies the question or answer from the player. This Speech Act, the output of the Language Interpretation module, is sent to the QA module.

Our dialogue manager consists of several main modules, including the Behavioural Agent (BA). We use frames to represent both the domain and the information collected during the interaction with the users. The Language Generation module has to find the best way to express what the Behavioural Agent has decided. The TTS module (DIXI) is a concatenated based synthesiser. This framework supports several voices and two different types of unit: fixed length units (such as diaphones), and variable length units. This latter data-driven approach can be fine-tuned to a limited domain of applications, by altering the design of the corpus. This was the case for this game, where several recordings with the voice emphasis were made.
The exhibition lasted from June to November, and more than 1400 visitors played Solis Curse. The system evaluation should incorporate technical and usability assessment.

The technical evaluation has to comprise the game manager module, all its sub modules (mainly the SDS system and the facial animation module), and the linkage between them, to ensure availability and performance requirements. These requirements are difficult to evaluate, and it can only be done by a list arranged by the development team. These requirements were met pleasingly, which reflected in a good user appraisal. The users reflected only about the usability evaluation. There is no absolute satisfaction metrics, only comparative ones by user interviews and forms. More than 50% of visitors used speech as the main approach to answer, which is something that they found very stimulating and remarkable. In this setting, there aren't many examples in Portuguese museums using speech recognition. Answering using the touch screen was also a welcomed solution, mainly due to the fact that makes the game more interactive and aesthetically appealing. Thus, the users were thrilled with the multi-modal way to answer the questions, agreeing that the visual interface had a high quality standard.

The ECA played a huge factor in the visitors satisfaction. If the player is not able to finish the game, the agent encourages him to review the exhibit through a cut scene. When interviewing the users, 70% of the inquired said that they felt more stimulated to take another look to the exhibit, in order to replay the game and have a better result.

V. Conclusion

The visitors fit in different age scopes and literary backgrounds, as well as diversified historical awareness of the monument. Our goal was to develop a game in order to meet the needs of all visitors, which influenced many game structural requirements. The data we gathered from the players confirmed that these aims were met. Also, the main purpose for this game is to improve cultural heritage awareness, which occurred when the visitors re-examined the exhibit to get a better score in the game. As mentioned above, there aren't many examples of serious games or speech recognition in Portuguese museums, so this was seen a breakthrough for raising cultural heritage consciousness. It was our objective to enrich the exhibit with multimedia contents, in order to enhance each visitor's experience.

The main technical goal was to introduce more people to the usage of speech recognition in serious games. As stated before, the visitors responded satisfactorily to this feature. In the future, we hope to further develop this technical element. Also, we expect to be able to animate the ECAs face in real time, as mentioned in the Game Components section. We expect to further develop this area of expertise in upcoming projects.

VI. Acknowledgment

We would like to thank the major entities and people that participated in this project and helped make this game possible: VoiceInteraction for making available the Voice engines and Spoken Dialogue System; The National Pantheon team for allowing us to develop this game and public raise awareness; The Exhibition Coordinator for presenting all the contents for the questions and answers; Our fellow colleagues, who helped vastly towards the development of this game; To the people that helped and advised us in various steps of the making of this article. Their support is greatly acknowledged.

REFERENCES


Thank you for reading.