Evaluation of semi-automatically generated accessible interfaces for educational games

Javier Torrente a, b, *, Manuel Freire a, Pablo Moreno-Gera a, Baltasar Fernández-Manjón a

a Complutense University of Madrid, Spain
b University College London, UK

A R T I C L E   I N F O

Article history:
Received 7 July 2014
Received in revised form 3 January 2015
Accepted 5 January 2015
Available online 12 January 2015

Keywords:
Accessibility
Games
Educational games
Interface adaptation

A B S T R A C T

The increasing body of evidence supporting the use of videogames in educational settings (usually referred to as serious games) is pushing their deployment across different areas of the educational system. However, this increased adoption also raises serious ethical issues: videogames are one of the least accessible forms of multimedia, and if education is to embrace serious games, there is an imperative need for universal accessibility in serious games to prevent a digital divide. However, producing accessible games is expensive and effort consuming, and serious games development already fare with limited budgets. In this work we explore the potential impact of the (semi-) automatic adaptation of game interfaces as a way to facilitate accessible game development (and thus trim the cost down). We propose a game interface model optimized for point-and-click adventure games, a popular genre among serious games that we have used to perform different semi-automatic adaptations in a game. We have tested the resulting adapted game with end users with specific disability profiles. Our tests discovered that the automatic adaptations produced usable games that retained part of their attractive, although different usability issues had a negative impact on the user experience. We also discuss the origins of such limitations and propose possible remediation actions, as well as a refined interface model.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Education is a Universal Human Right (United-Nations, 1948). As new technologies are brought into the classrooms, equality principles and eventually laws require that they be made accessible to all students to prevent the infringement of this right (see for example section 508 (IT Accessibility & Workforce Division (ITAW), n.d.)). This should also be the case of educational games, commonly referred to as serious games, which are rapidly gaining acceptance and represent a promising educational tool for the near future (Johnson, Adams, & Cummins, 2012; Johnson et al., 2013), as evidence proving positive impact on students’ achievements (Papastergiou, 2009; Sadler, Romine, Stuart, & Merle-Johnson, 2013) and motivation continues to grow (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Hwang & Wu, 2012). However, the current levels of accessibility in videogames (both educational and recreational) are still relatively low compared to those of Web-based technologies commonly used to support learning (Briere et al., 2005; Westin, Bierre, Gramevos, & Hinn, 2011; Yuan, Folmer, & Harris, 2011), and far below acceptable standards. This aspect is especially relevant in ageing populations, like those in most western countries, where the percentage of people affected by any sort of disability is rapidly growing ((WHO), 2011). Lack of accessibility also prevents wider adoption of serious games for the education of students with special needs (Yuan et al., 2011), one of the areas where educational games have shown greater potential but which is still in an early stage (Durkin, Boyle, Hunter, & Conti-Ramsden, 2013). In this area, serious games have been effectively used to teach alcohol-affected children about fire and street safety (Coles, Strickland, Padgett, & Bellmoff, 2007), to improve visual-integrative abilities and sensory integrative functioning in children with Down syndrome (Wuang, Chiang, Su, & Wang, 2011), or to improve students’ chances to getting and keeping a job (Lanyi & Brown, 2010). In order to generalize the use of serious games for special education it is indispensable to improve their level of accessibility. Besides, it is necessary to identify those types of serious games that better combine accessible design and educational value for students with special needs.

* Corresponding author. Complutense University of Madrid, Spain.
E-mail address: jtorrente@e-ucm.es (J. Torrente).

http://dx.doi.org/10.1016/j.compedu.2015.01.002
0360-1315/© 2015 Elsevier Ltd. All rights reserved.
One of the most relevant causes of the lack of accessibility in serious games is the non-trivial cost of implementing accessibility into an existing game, an effort-consuming process that may affect multiple aspects of the game. Typically this would require interventions affecting art resources, game design, difficulty adjustments, puzzles, language used, gameplay mechanics, and the different underlying game technologies (Grammenos, Savidis, & Stephanidis, 2009; Ossmann, Miesenberger, & Archambault, 2008). For this reason, game accessibility solutions are usually game-specific and therefore hard to scale and reuse across games. And that is partly why it is rare to find educational games that take accessibility into consideration.

With the goal of increasing the accessibility of serious games we advocate for providing developers with software tools that take care of as much of the process of making a game accessible as possible. We denominate the approach as “semi-automatic” because the developer will always need to deal with some tasks, but the goal is to reduce these manual adaptations to a minimum.

In this paper we discuss a first approach focused only on improving the accessibility of the game interface. Although making a game accessible requires more than providing alternative user interfaces (Grammenos et al., 2009), we consider this work a necessary first step before more ambitious approaches that deal with other aspects of the games can be considered.

The article is structured as follows: in Section 2 we describe the general approach, setup the specific scope of this work, and provide some general context about point-and-click games necessary to understand our work. In Section 3 we describe the general interface model proposed. Section 4 contains a case study where we evaluate the approach, while Section 5 discusses a potential set of improvements that may yield improved results. In Section 6 we describe some related work and finally, in Section 7, we summarize our conclusions and outline the next natural steps in this research area.

2. Approach, scope and context

In this section we briefly introduce our approach (Section 2.1.). We also discuss the actions we have taken to narrow down the scope of this work (Section 2.2.). Finally, we provide an overview of point-and-click games from the point of view of their interface and the accessibility barriers and opportunities they offer. This overview is important to understand the game interface model described in Section 3.

2.1. Approach

Our approach to reduce the overhead of accessible game development is to (semi-) automatically adapt the user interface, taking into account the special needs of each user profile, requiring a minimum amount of input from the game author. However, it is still necessary to make simple additions such as providing alternative synonyms for command recognition, making final tweaks and revising the adaptations made by the system. After revising and complementing the automatic adaptations, different versions of the same game are available for delivery to the end-users, each one configured with a different interface (see Fig. 1).

The application of this approach to the adaptation of serious games to different accessibility profiles is a significant challenge for different reasons: (1) The capabilities of persons with disabilities vary a lot for each individual, depending on the type and degree of disability, (2) games have some of the most complex interfaces in modern software, and (3) it is significantly more difficult to make a game accessible than a desktop or web application (Archambault, Ossmann, Gaudy, & Miesenberger, 2007).

2.2. Scope

We have taken two actions to narrow down the scope of this work. First, we have focused on classic “point-and-click” adventure games as the target game genre. Videogames are an heterogeneous media and it is not possible to propose a solution that covers all of them at once. We focus on this genre because it has been identified in the literature as having the higher potential for serious and educational applications due to its strong narrative underpinnings (Adams, Mayer, MacNamara, Koenig, & Wainess, 2012; Garris, Ahlers, & Driskell, 2002; M D Dickey, 2006; Michele D. Dickey, 2011).

Second, in our proposal we consider only physical disabilities, focusing on three profiles: blindness, low vision and reduced mobility. Physical disabilities are the most relevant for the purpose of this work since most of the barriers they find in videogames are related to the interface. Cognitive disabilities (another relevant area of research in accessible games (Torrente, Del Blanco, Moreno-Ger, & Fernández-
standard verbs, their synonyms, and the standard names for all in-game objects and characters.

3. Proposed adapted game interface model for point-and-click adventure games

Classical point-and-click adventure games had their golden age in the 90’s, with titles like Monkey Island® or Myst®. In these games the user moves the mouse cursor around the game scene to find elements to interact with. When hovering over an interactive element, visual feedback is provided. For instance, to indicate an active element, the cursor may change shape, the element’s borders may become highlighted, and a label with an element description may be temporarily revealed. Users can decide, based on this feedback, to further interact with the element (e.g. clicking on a character may launch a conversation, and clicking on a door may trigger a menu with options such as open, close, or examine for clues) (Fig. 2).

From an accessibility perspective, this type of interface presents advantages and drawbacks. The main accessibility problems presented by these games are the reliance on sight to explore the game world, and the need for both sight and fine upper-limb motor control in order to use the mouse for pointing and clicking with precision. Both of these also are common accessibility problems in other game types (Archambault et al., 2007).

In contrast, a positive aspect is that point-and-click adventure games do not present many other typical accessibility barriers, such as fast pacing, absence of subtitles (these games are close-captioned very often), or the use of time pressure to provide challenge (Bierre et al., 2004). Additionally, narrative games (both linear or with multiple branches) are well suited to textual interfaces. Indeed, text-only adventure games predated the graphical, point-and-click variety. Adventure games are therefore uniquely suited to supporting accessibility, offering a clear separation between game actions and interface which allows the use of multiple interface variants (Grammenos, Savidis, & Stephanidis, 2007).

3. Proposed adapted game interface model for point-and-click adventure games

Since it is possible to detach interface and in-game tasks in point-and-click games, we propose a revised interaction model where the same game world can be affected by different alternative input methods, and its internal state can be represented in different alternative forms (Fig. 3).

The next sections describe in more detail how this interface model works for the user profiles selected. Also a reference implementation of this model based on the eAdventure platform can be found in (Torrente, Del Blanco, Moreno-Ger, Martinez-Ortiz, & Fernandez-Manjón, 2009). The implementation of the game interface was possible because the eAdventure authoring platform provides an explicit representation of the game model, which can therefore be processed and adapted.

3.1. Interface model for blind users and users with reduced mobility

Both blind users and users with reduced mobility will experience problems for input using the mouse, either due to missing feedback (blind users cannot perceive where they are pointing to) or due to lack of fine motor control, respectively. In addition, blind users have problems to perceive feedback from the game if it is transmitted through graphical output (Yuan et al., 2011).

In both cases, players formulate short commands in natural language: (e.g. “grab the notebook”, or “talk to the character”). An interpreter reads the commands, executes them (if they pass a syntactic and semantic validation), and provides suitable feedback using the appropriate channel: auditory for blind users through a built-in text-to-speech engine, and text for users with reduced mobility. As depicted in Fig. 3, unrecognized commands are never sent to the game world, which expects only valid game actions.

Command processing is driven by a grammar that defines valid commands, combined with a list of synonyms for relevant verbs (actions) and nouns (interactive elements) that includes built-in synonyms for common words (e.g. “use”, “grab”, “talk”) and synonyms specified by the game author for each game element. The grammar is automatically generated from each game description, based on the game actions defined for each of the interactive elements available in each game scenario. An additional set of game-independent vocabulary provides access to always-available interactions such as opening menus, skipping dialogue lines, or exiting the game altogether.

The grammar in Table 1 can be automatically generated for each game without human intervention. This includes the provision of synonyms for those verbs that are common to all games (e.g. grab/take, examine/inspect, etc.), resulting in a grammar that includes the standard verbs, their synonyms, and the standard names for all in-game objects and characters.

![Feedback loop interacting with a point-and-click adventure game.](image)

Fig. 2. Feedback loop interacting with a point-and-click adventure game.
However, human intervention is required to facilitate further context-based synonyms for in-game elements (e.g. door/gate). While it would be feasible to use dictionaries to further automate these definitions, it would not be desirable since the appropriateness of synonyms may depend on the specific context (e.g. there may be a door and a gate on the same scene) and narrative (e.g. a character that represents a talking cat called Alfred could be referred to as the cat or Alfred). In addition, the in-game scene descriptions may need to be adapted to each profile, since the required level of detail and actual use of language may be different when targeting users with reduced mobility or blindness. However these adaptations are simple to make once the developer is provided with a support tool that loads the results of the automatic modification and allows the developer to fine tune as needed.

3.2. Interface for users with low vision

Major barriers for low vision users are related to having interactive elements that blend into the background or elements and fragments of text that are too small (Bergel, Chadwick-dias, Ledoux, & Tullis, 2005). Additionally, color-blindness can result in nominally different colors blending into each other, and is especially problematic when color is used to encode important attributes or to convey relevant information.

A number of strategies to address these barriers are proposed. First, text size and small game elements are significantly enlarged. Second, a special rendering mode is used to increase the contrast of interactive elements over game-scenario backgrounds. Luminosity of interactive elements is increased using a light green filter. A dark purple filter is applied to all other areas, decreasing their brightness, resulting in a high-contrast scheme (Fig. 4).

These effects, as well as additional adjustments in font sizes and colors used for cursors, buttons and menus can all be performed automatically, and so they are in our reference implementation. In some cases, human intervention may be required to identify unexpected outcomes of the automatic adaptations (e.g. artifacts generated by the automated border generation). But still this intervention is very cost effective compared to the manual generation of those adapted assets.

4. Case study

As a preliminary evaluation of the interface model proposed, we conducted a case study, which is discussed in this section. The main objective of the case study was to observe how these simple semi-automatically generated interfaces could be used in a real scenario with real users, and how far it is possible to reach with limited human intervention in the generation of accessible educational games.

4.1. Method, participants and settings

Fourteen volunteers with diverse characteristics were recruited. Five of these users had no disability and played the standard version of the game, three had reduced mobility (but good vision), three were blind, and the final three had low vision and used the high contrast mode. The number of users recruited, although limited, is consistent with available recommendations for usability evaluation through user observational methods (Fernandez, Insfran, & Abrahao, 2011).
Average age of participants was 35.64 (SD 9.64), with a minimum of twenty-one and a maximum of 55. Most of the participants were females (eleven out of fourteen; 78.57%). Nine of the participants held a college degree or a PhD, four had only completed high school and only one had no formal qualifications. They were asked to rate their own computer literacy from one to five, resulting on 3.07 on average (SD .616) and a median of three.

Participants were individually briefed on the experiment and asked to sign a standard consent form, in which they were informed that only anonymous results would be published for research purposes, and that they were free to abandon the experiment at any time without penalty. The participants were requested to fill a short questionnaire including demographic and general computer literacy questions. Next, users were provided with an adapted version of the same game to play for roughly 40 min allowing them to play more or stop earlier if desired. Users with no physical disability played the original game with no adaptations. The same facilitator that provided the briefing stayed with the participant during the whole session, providing advice when requested and occasionally asking participants to reason out loud when they appeared to be stuck.

All sessions were recorded in high-definition video, capturing the screen and the voice of the participants and the facilitator (who was free from taking notes during sessions). At the end of each session, regardless of whether the game had been finished or not, participants were asked to fill a survey on their experience. Post-game comments to the facilitator were also captured on the video soundtrack. An empty, quiet classroom was used for recording, and long sessions were often split to allow the participant and the facilitator a few minutes of rest.

Video recordings were analyzed using a user observational method similar to the one described in (Moreno-Ger, Torrente, Hsieh, & Lester, 2012). Following this method, users interact with the application individually. They are prompted to “think aloud” (verbalize all their thoughts and feelings just as they come) while completing a set of tasks proposed by the researchers. All sessions are recorded (including both screen and user’s face expressions). Two different researchers analyze the video recordings, tagging events on two different dimensions: user-related (if the user is learning/reflecting/etc.) and system-related (if the event is about the layout and user interface/content/technical error/etc.). Then all researchers meet and revise all annotations, reconciling any discrepancies found until a single agreed events document is produced. This document is used to decide on potential changes to be made in the application during the next development cycle.

4.2. Materials

4.2.1. The game “My First Day at Work”

The eAdventure educational game that participants were asked to play, “My First Day at Work”, was developed in collaboration with Technosite, an offshoot of Spain’s largest disability-oriented NGO, ONCE (http://www.once.es). The game places the protagonist as a newly-hired office worker, arriving for his or her first day on the job. Participants can choose their game avatar among a cast of four characters, each of them with different accessibility requirements (blind avatar, avatar in wheel chair, avatar with hearing impairments and a fourth avatar.

<table>
<thead>
<tr>
<th>Textual command</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examine the wall (1)</td>
<td>Describes of the object “wall”, if it exists in the scene.</td>
</tr>
<tr>
<td>Go to the door (1)</td>
<td>The student’s character in the game attempts to move towards a place called “door”.</td>
</tr>
<tr>
<td>Use keys with locker (1)</td>
<td>The game attempts to use the “keys” object on the “locker” object</td>
</tr>
<tr>
<td>Name items in the scene (2)</td>
<td>Lists the scene elements that have already been discovered and are therefore available for examination or interaction</td>
</tr>
<tr>
<td>Open options menu (2)</td>
<td>Pauses the game and show the options menu.</td>
</tr>
<tr>
<td>Describe (the) scene (2)</td>
<td>Provides a description of the scene as a hint for the student.</td>
</tr>
</tbody>
</table>

Fig. 4. Left: Standard visualization of a game scene from eAdventure game ‘Eating Out’. Right: Adapted visualization of the scene for people with low vision. High contrast rendering mode is applied darkening the background and highlighting the interactive elements.
with no apparent disability). Player character selection is also used to configure the user interface. For example, if the user selected the blind avatar, the game is set up with the interface described in Section 3.1.

The game is structured into two chapters. The first chapter has no back-story, and is intended as a short tutorial covering the interaction techniques the player will need (which are different depending on the character selected), so he or she can get used to the interface. It consists of a single small puzzle involving a table, a key hidden inside a book placed on this table, and a keypad where the key must be entered in order to proceed to the actual game.

Upon entering the main chapter, the player is greeted with an explanation of the setting for the main adventure. Next the player arrives at the company’s front desk, and is instructed to find the Director of Human Resources at her office, and given a series of tasks. There are no separate levels, and the order followed to complete tasks does not affect the game in most of the cases. The game ends when the character visits the Human Resources Director after completing all tasks. Fig. 5 depicts an overview of the game-structure, used when reporting user times in Fig. 7 and Fig. 8.

Game-play in this chapter also differs slightly according to player character. For example, the blind character is expected to perform additional setup steps to customize her in-game workplace as she would need to do in real life (e.g. get a screen reader installed in her computer). These adaptations were not fully automatically generated, as they involved minor modifications in the game design and had to be integrated into the game metaphor to favor immersion (Torrente, del Blanco, et al., 2012).

While a speed run through the game can be performed in less than fifteen minutes, typical game-play ranges from 30 to 60 min (see Fig. 7), depending on the interface used. In this case study participants were limited to around 40 min of gameplay, and completing the game was not considered mandatory, so some stopped playing before completing it.

The game contains a total of 25 scenes with approximately 40 scene items, four playable characters, seven non-playable characters (NPCs) with twenty interactive conversations, and eleven inventory items.

4.2.2. End-user questionnaire

A four-item Likert questionnaire was handed out to the users after each test session. The questionnaire had nine common items for all users, plus four additional items for users using one of the adapted interfaces (see Tables 2 and 3). These items were formulated to evaluate either what was the general feeling of the users about their gameplay experience or the usability of specific aspects of the game interface.

Two different scales were built upon the common section of the questionnaire (Table 2):

- **G**: Rate for the Usability of the game. Aggregation of all nine elements: G.INT-G.MEC. (G.LOST was reversed first, resulting in item G._LOST = 5-G.LOST).
- **G.UX**: Rate for the User Experience of the game. Aggregation of four items related to user feelings about the gameplay experience: G._LOST, G.COM, G.FRUS and G.FUN.

A Cronbach’s alpha test was conducted to assess the reliability of both scales, resulting in .696 for G.UX and .847 for G. Two items with low item—total correlation factors (<.5) were removed from both scales (G._LOST and G.TEXT), achieving and increased Cronbach’s alpha of .867 for G.UX (three items) and .905 for G (seven items).

4.3. Results

This section provides the results obtained organized in three subsections: a quasi-quantitative analysis of the results of the end-user questionnaires; a qualitative analysis of end user comments, and finally and heuristic evaluation produced by usability experts that reviewed the recording sessions.

4.3.1. End user questionnaires

Average scores for the scales G (usability) and G.UX (gameplay experience) were measured and compared across user profiles, using the average score of users with no-disability group for comparison. As Fig. 6 and Table 4 show, rates provided by Blind and Low vision users were similar to users with no disability. Only users with reduced mobility provided lower scores.

In general, both usability and user experience were positively evaluated by the users.

Individual analysis of all items in scale G reaches a similar conclusion. Table 5 provides the median for each item, highlighting any data above (+) or below (−) the group of non-disabled users. There are almost no differences observed compared to the control group but in the reduced mobility group, which consistently scored lower.

Data collected on items that were specific to each disability profile are reported on the next tables. Recognition of speech commands achieved the lowest scores, which may be the reason why users with reduced mobility provided lower scores.

4.3.2. Annotations

Participant sessions were recorded on video. The videos only captured user interaction (the screen of the participant’s computer), spoken comments and interactions between the participant and the facilitator. Full transcripts were generated. The following events were coded and time-stamped:

- **User interaction with the game:**
  - **good, not useful** (for instance, choosing to go back to a completely visited area),
  - **exploratory** (systematically explore all available objects or actions in a scene, without a clear goal), and
  - **not recognized** (for instance, when spelling errors or unexpected syntax was used in textual or spoken interfaces).

- **User interaction with facilitator:**
  - user requests help,
  - facilitator provides help, and
Fig. 5. Map of the game. The third chapter encompasses the main tasks.
Fig. 6. Aggregate general perception by profile type (left) and aggregate usability perception by profile type (right). Participants within the blind profile rated the game as highly in the G category as those in the non-disabled profile, and almost as highly in the G.UX category, while limited-mobility players rated it significantly lower in than other profiles in both categories.

Fig. 7. Time spent in tasks by user profile (labels at far right). The first vertical dashed line at T13 marks the end of the tutorial, which was only available for users in the Blind and Reduced mobility profiles. The line at T82 marks the end of the game itself, reached by only five of the fourteen users, all of them from the No Disability and Low Vision profiles. The vertical dotted line at Th1 marks the start of the help task (available at any time after T24 is finished). Large delays in this task were due to users exploring available help questions, and often having problems finding the exit from the task itself. T33 was particularly difficult for certain profiles, as was T23 (the first conversation). Compared with T25 (second conversation), a strong learning effect can be observed.
A set of data-points for u06 is missing due to a corrupted video file. Several tasks could be performed in multiple orders; for instance, in U13, tasks T51-T54 were performed after T74 and before T81b; additionally, some tasks could be skipped without penalty. Only five participants managed to finish the game, all of them from the Low Visibility and No Disability profiles. Due to problems with voice-recognition and text-input interfaces, users from the Blind and Reduced mobility profiles reached similar results; however, feedback by both groups is very different.

- **general conversation**

- **Game events:**
  - session start/end,
  - user changes scene,
  - user finishes task.

All video tagging and coding was performed by a single author, avoiding inter-coder agreement issues, but also raising the possibility of coder errors, omissions or interpretation issues (Boring & Gertman, 2005; Law & Hvannberg, 2004). Figs. 6 and 7 contain visualizations generated from the coded transcripts of the user videos, totaling around twelve hours of gameplay time and over 3000 events. Both visualizations were generated using R’s ggplot2 package.1

4.3.3. Discussion of gameplay issues

Within the low-vision and no-disability profiles, only one user of each group failed to complete the game in time. None of the blind or limited-mobility groups came close. This is a major setback, although it does not necessarily mean it is caused by the adapted interfaces. Moreover, much can be learned from analyzing the videos looking for the root causes.

The traces for users from the Reduced Mobility (RM) profiles are remarkably similar. All users became stuck in approximately the same moment, and decided not to progress any further. The greatest cut-off is T34, where players encountered a long in-game e-mail that described the tasks that they were expected to accomplish during the rest of the game. Having invested upwards of twenty minutes to reach this point, and given the difficulties interacting with the voice-recognition input, they decided not to continue playing. Blind users (B) found no accuracy problems (being limited only by their actual keyboard accuracy), but still had problems formulating correct commands, and also abandoned shortly after T43. The large difference in their feedback of the game experience (see Fig. 6) is due to the perceived cause for slow progress, rather than the progress itself. For the blind, the interface behaved mostly according to expectations, while for reduced mobility users, who had a clear idea of exactly what they wanted to do but had problems achieving those goals due to poor voice recognition and error feedback, the interface was to blame for not allowing them to achieve their goals. Nonetheless, poor voice recognition accuracy was not a problem related to the interaction model proposed, but of the open-source voice recognition software used, which was also untrained.

Blind and low-vision players could use the “actions” command to see a list of available actions. However, many available actions were only hinted at. For example, the command “actions” could comment on the existence of three objects in the scene; but it was up to the user

---

1 http://cran.r-project.org/web/packages/ggplot2/index.html.
Table 2
Wording for the general section of the end user questionnaire.

<table>
<thead>
<tr>
<th>Item id</th>
<th>Wording</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.INT</td>
<td>I was able to interact with the game without experiencing any problems.</td>
</tr>
<tr>
<td>G.TXT</td>
<td>Text descriptions and dialogues helped me to understand the game mechanics.</td>
</tr>
<tr>
<td>G.LOST</td>
<td>I felt lost very often and I did not know how to continue.</td>
</tr>
<tr>
<td>G.HELP</td>
<td>Support and help contents provided by the game are appropriate.</td>
</tr>
<tr>
<td>G.COM</td>
<td>I felt comfortable using the game.</td>
</tr>
<tr>
<td>G.FRUS</td>
<td>I felt frustrated using the game.</td>
</tr>
<tr>
<td>G.FUN</td>
<td>The game was fun.</td>
</tr>
<tr>
<td>G.MEC</td>
<td>The game mechanics were easy to understand for me.</td>
</tr>
</tbody>
</table>

* Indicates items that were reversed for result analysis.

Table 3
Wording for the disability-specific section of the end user questionnaire.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Item id</th>
<th>Wording</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Interface (Reduced Mobility)</td>
<td>S.COM</td>
<td>Always knew what commands introduce.</td>
</tr>
<tr>
<td></td>
<td>S.REC1</td>
<td>Voice recognition was efficient.</td>
</tr>
<tr>
<td></td>
<td>S.FEED</td>
<td>I was able to notice when the game had recognized a command.</td>
</tr>
<tr>
<td></td>
<td>S.REC2</td>
<td>The game has recognized my commands.</td>
</tr>
<tr>
<td>Keyboard Interface (Blind)</td>
<td>K.COM</td>
<td>I knew what commands or phrases type in.</td>
</tr>
<tr>
<td></td>
<td>K.REC1</td>
<td>The game has recognized my commands.</td>
</tr>
<tr>
<td></td>
<td>K.FEED</td>
<td>I knew when the game was ready to accept new commands.</td>
</tr>
<tr>
<td></td>
<td>K.FEED2</td>
<td>I knew when the game had recognized a command.</td>
</tr>
<tr>
<td>High Contrast Interface</td>
<td>H.ELM1</td>
<td>I could easily recognize objects and characters on the screen.</td>
</tr>
<tr>
<td></td>
<td>H.TXT</td>
<td>I could read the texts clearly.</td>
</tr>
<tr>
<td></td>
<td>H.BLG</td>
<td>I could always tell what character was talking.</td>
</tr>
<tr>
<td></td>
<td>H.ELM2</td>
<td>I could distinguish interactive elements from non-interactive elements.</td>
</tr>
</tbody>
</table>

Table 4
Medians of user experience and general usability for each user profile.

<table>
<thead>
<tr>
<th>Profile</th>
<th>General usability (G) [Range 7–28; 7 items]</th>
<th>Evaluation of user experience (G.UX) [Range 3–12; 3 items]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Blind</td>
<td>Low Vision</td>
</tr>
<tr>
<td>Blind</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low vision</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Reduced Mobility</td>
<td>23</td>
<td>17</td>
</tr>
<tr>
<td>No Disability</td>
<td>23</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 5
Median for every item on the common section of the questionnaire. Medians above and below the control group are highlighted. This table clearly shows that the speech interface underscored compared to the other three.

<table>
<thead>
<tr>
<th>Profile</th>
<th>G.INT</th>
<th>G.TXT</th>
<th>G.LOST</th>
<th>G.HELP</th>
<th>G.CTRL</th>
<th>G.COM</th>
<th>G.FRUS</th>
<th>G.FUN</th>
<th>G.MEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blind</td>
<td>3.00</td>
<td>(–2.00)</td>
<td>(–2.00)</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Low vision</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Reduced Mobility</td>
<td>(–2.00)</td>
<td>3.00</td>
<td>(–2.00)</td>
<td>3.00</td>
<td>3.00</td>
<td>(–2.00)</td>
<td>(–2.00)</td>
<td>3.00</td>
<td>(–2.00)</td>
</tr>
</tbody>
</table>
to first say “objects” to know what they were, and then actually examine each object in turn to probe it for additional actions that could be performed. Certain complex commands would show up in “actions” only after exploring the relevant objects.

Error reporting was far from detailed. In particular, it was very common for players to forget to dismiss conversations or messages before sending in new commands – approximately 63 commands were ignored for this reason, as the only valid action when a message or conversation was active was to say “advance” (or click a mouse button or arrow-key on the keyboard). When commands failed, no details were provided as to the reason of the failure. At this point, players often requested clarification from facilitators in figuring out what had failed. Indeed, players with Blind and Reduced Mobility profiles requested significantly more assistance from facilitators (helped 25 times on average during their sessions) than players with the LV or No Disability profiles (helped on average in 3.7 occasions). As expected, slow progress throughout the game tasks also resulted in more casual conversation with facilitators, with over 3.2 times the amount of general-conversation events in Blind/Reduced Mobility profiles as compared to Low Vision/No Disability. This also prevented users in Blind and Reduced Mobility groups from completing the game, since the amount of time available was limited (60 min).

4.3.4. Examples of interface problems found in transcripts
Several users tried to rely on dialogues with in-game characters to “help them out”, since in the real world asking people is a very effective way of gaining information. However, the game’s conversations were limited in scope, and sometimes, as in the following transcript, contained breaches of protocol:

- U01 (Blind profile, issuing text command): talk to receptionist (who had just told her to visit the Managing Director’s office to present herself)
- Receptionist (game character): I have nothing more to say
- U01 (to facilitator): “This is totally unrealistic! Nobody expects a blind person to find her way around an unknown office without offering guidance!”

This problem, although it may have had a significant impact in the whole experience, was related to the adaptation of the game design and not to the interface, which is the subject of the study.

Users that had not been able to play with accessible videogames often had encouraging comments regarding the game.

- U01 (Blind profile, to facilitator): “hey, this is fun. Can I take a copy to play at home?”

Lack of context (the interface did not remember previous incomplete commands) and contextual help (the interface could not suggest possible continuations, or diagnose the cause of errors) was a significant hurdle for participants in the Blind and Reduced mobility groups.

- U07 (Reduced Mobility profile, issuing spoken commands): talk (fails)
- Facilitator: you have to specify with whom
- U (command): managing director (fails)
- F: now you have to specify the verb
- U (command): talk to the managing director (fails; extra “the”)
- U (command): talk to managing director (ok – conversation starts)

In some cases, intrinsic task difficulty (such as assuming familiarity with certain concepts) was a greater problem than expected.

- Facilitator: do you know how to use an e-mail reader? (U15, Low Vision profile, has spent a few seconds looking blankly at the screen)
- U15: yes, sort-of … actually, I haven’t used e-mail for a long time
- F: do you know what Powerpoint is?
- U: actually, no.

4.3.5. Remarks to interviewers during debriefing
Participants were asked to comment on positive and negative aspects and things that they would improve in the game; facilitators elicited the following responses:

Participants commented positively on

- Exploratory nature; liked not knowing what was in store (U1 Blind)
- Novelty; never played something like this before (U2 Blind)
- Accessibility; very few accessible games (U3 Blind)
- Unexpected outcomes during the game; for instance, a certain co-worker seems to be willing to date (U4 Low Vision)
- Design and simplicity (U6 Low Vision)
- Avatars for people with disabilities; general accessibility (U15 Low Vision)
- Small touches of humor (U8 No Disability)
- The need to focus in order to interact correctly with other characters (U11 No Disability)
- Good dialogue scripts, which should make it fun also for players with disabilities (U13 No Disability)
- It is not only educational, it is an actual game — that is, it is fun to play (U14 No Disability)

Suggestions and negative comments were
5. Proposed amendments to the model

The most important part for the scope of this paper is to analyze game interface issues and how these can help us improve the interface model proposed. It is interesting to point out that a considerable number of the issues identified in the case study are related, partially or totally, to the game design and not to the game interface, which is the subject of study. For example, problems related to the “help guy” character are a consequence of a poorly designed conversation that did not take into account previous topics discussed with the character. While this was a problem all users suffered, it was especially frustrating for blind and reduced mobility users. Other game-design issues can be explained by the experience users had playing videogames. For example, while some users struggle to understand how the game worked, others rapidly understood the mechanics and found the pace too slow.

Also implementation-related issues arose in user interviews (e.g. problems with the voice recognition software), which require technical interventions (e.g. including a more mature voice recognition package).

Most interaction problems for users within the blind (Blind) or reduced-mobility (RM) profiles were due to players not knowing what actions to perform to achieve a given goal, or having problems when attempting to communicate these actions to the game engine. Many of interaction problems could be solved implementing hierarchical menu-driven action selections, as found in the suggestions of all limited-mobility users. The same approach would be of use for blind users — especially during the first minutes of interaction, when players are faced with the dual challenge of learning the gameplay and the interface simultaneously.

Therefore, in our amended model, we propose a new menu-driven interaction mode for profiles Blind and Reduced Mobility. This mode would be the default for new users, until they felt comfortable enough to write or dictate commands directly, and would allow menu navigation similar to that of standard voice-driven automated call centers (including numerical selection of options by order of presentation), augmented with typical text-to-speech controls to skip items, control output speed, and repeat the last utterance. User familiarity with these systems, currently highly widespread in company call centers, is seen as an added benefit. As an example, a possible menu would be the following (triggered by pressing a key or saying “menu”):

- Actions, ten available
  - Use, telephone or fax machine (select to choose which)
  - Take telephone
  - Talk to Jane
  - Examine, four targets (select for list)
- Persons: Jane (collapsed due to having a single option available)
  - Examine
  - Talk to
- Objects, three in scene
  - Potted plant (can examine)
  - Telephone (can use, take, or examine)
  - Fax machine (can use, examine)
- Tasks
  - Your current task is to send a fax; say “change” to change
  - There are three hints available for this task
  - You have completed five tasks, say “completed” for list (list follows)
  - There are ten tasks remaining, say “remaining” for list (list follows)
- Game
  - Stop using menu by default; you can activate it again saying “menu”
  - Help for voice commands
  - Save or restore game
  - Quit

Before being sent to the game, commands would be shown or spoken briefly to the user. This would help users learn the correct commands that would allow them to bypass the menu system entirely. Additionally, menus would be displayed on the screen for reduced mobility players in a similar way to in-game conversations or decisions, providing a more consistent interaction throughout the game. For reduced mobility users, additional visual feedback can be provided at low development cost by highlighting scene objects as they are
mentioned, mimicking the discovery process of “hovering the mouse over the scene” frequently identified in the group of users with no disabilities when encountering a new game-scene.

Notice that a new type of commands is being proposed: the “Tasks” submenu was not available in the previous version of the interface. We feel that its addition would allow players to track their progress and request context-dependent help. This would attempt to address the increased amount of facilitator support requested by certain user profiles observed during our experiment, as in many cases, facilitators limited themselves to restating the player character’s context.

Significant improvements can be made in error handling and reporting. In our amended model, we propose additional error-related output, indicating the exact point of the command that caused the problem and, if available, possible continuations. For example, we envision the following interaction (which could only arise in free command mode; menu actions would always be correct in a syntactic sense):

- Alice (player): talk to Bob
  Interface: Bob cannot be talked to; you can talk to Peter or John

Or

- Alice (player): talk
  Interface: you can talk to Peter or John

The addition of a menu system and of improved error handling and reporting will require only small changes to the game platform, and most importantly, no changes at all to the games themselves. The addition of a tasks submenu and hints for the tasks would require additional game-hooks and authoring tool support, but we feel that it would offer increased support to players who may become lost during game-play.

6. Related work

We believe this work is innovative and the first of its kind to the best of our knowledge, since automatic adaptation of game interfaces for accessibility purposes has not been proposed elsewhere. However, there are several works that have deeply influenced us which deserve discussion.

We have taken ideas from a body of works that deal with automatic generation of interfaces, which has a long history in other fields (Boutekkouk, Tolba, & Okab, 2011; Falb et al., 2009). Also state-of-the-art accessible game interaction has been considered and integrated, being the literature reviews provided by Westin (Westin et al., 2011) and Yuan (Yuan et al., 2011) especially helpful. The high contrast adaptations performed are also influenced by Westin’s Terraformers (Westin, 2004). We have also integrated ideas from other authors that have identified accessibility barriers users with physical disabilities face, and proposed guidelines to overcome them (Game Accessibility Guidelines, 2012; MediaLT, 2006; Ossmann, Archambault, & Miesenberger, 2008). Finally, a very inspiring work is the Unified Design of Universally Accessible Games (Grammenos et al., 2007, 2009), a methodology developed by Grammenos et al. for developing universally accessible games that enforces designing game tasks that are not bound to any specific interaction, and which has also been adopted by other authors (Garcia & de Almeida Neris, 2014).

7. Conclusions and future work

We have described a case study that tested user perceptions, perceived usability, and performance on an automatically adapted game for three physical disability profiles: blind, low vision, and reduced mobility, along with a group of users with no apparent physical disabilities for comparison. From the point of view of performance, two out of three low vision participants managed to finish the game within time, while none of the participants in the reduced mobility or blind profiles advanced into the second half of the game. These profiles had in common the use of a text (or speech-recognition) interfaces for command input.

Most to the problems found are not a consequence of the interface model proposed, but of technical errors or flaws in the game design. In this sense, this experience confirms that users with physical disabilities need not only to have an adapted interface, but also an adapted game design including the alternative branches of the plot and enhanced character support. The semi-automatic adaptation of game design and content, which was not the purpose of this work, will be explored in the future.

As for the interface model proposed, the detailed study of video recordings and user interviews has allowed us to identify the problems with our automatic adaptations, leading to several important changes to the adaptation model we proposed.

It is striking that the general perceptions and usability is markedly different between participants from the reduced mobility and blind profiles, given that they were using the same game interface. In particular, blind participants rated the system and its usability much higher than the reduced mobility participants. We attribute some of these differences to the novelty of accessible video games from the point of view of our blind participants, coupled with the unacceptably high failure rate of speech recognition encountered by users of the reduced mobility interface. However, we understand that the main culprit is the underlying textual command system and its lack of visual feedback for users with unimpaired sight.

The potential of this approach to trim down the cost of implementing accessibility into serious games has already been discussed in (Torrente et al., 2014), were the development of the first version of “My First Day at Work” is studied. The first version created without taking into account disabilities required roughly 4 person-months of development effort. Then, the solution proposed was applied to make the game accessible for the target disabilities discussed in this paper (blindness, low vision, reduced mobility), requiring only two additional person-weeks of development effort. This effort was dedicated to adding descriptions for the text-to-speech module and synonyms for the natural language input processing module, which are relatively cheap tasks. The development also required increases in resources and effort in the tutorials (10.32% increment) and art resources (8.96% increment).
As a final remark, we would like to highlight the fantastic acceptance of the game, which shows people with disabilities are avid to have more content like this available and justifies further research in the field. Also, the sort of adaptations proposed and the game developed could be applied to other types of physical disabilities, like deafness. Cognitive disabilities fall out of the scope of this work.

Acknowledgments

The e-UCM research group has been partially funded by Regional Government of Madrid (eMadrid S2013/ICE-2715), by the Complutense University of Madrid (GR13/4-21340), by the Ministry of Education (TIN2013-46149-2-C1-1-R) and by the European Commission (SEGAN 519332-LLP-1-2011-1-PT-KA3-KA3NW, H2020-RAGE 644187).

We also want to thank Mercedes Turreiro and Manuel Ortega Moral from Technosite for their support of this work, and also to all the users that kindly volunteered to participate in the case study.

References


