

Available online at www.sciencedirect.com**ScienceDirect**

Procedia Computer Science 100 (2016) 935 – 940

Procedia
Computer Science

Conference on ENTERprise Information Systems / International Conference on Project
MANagement / Conference on Health and Social Care Information Systems and Technologies,
CENTERIS / ProjMAN / HCist 2016, October 5-7, 2016

Audio description in Interactive Television (iTV): proposal of a collaborative and voluntary approach

Rita Oliveira^{a*}, Jorge Ferraz de Abreu^a, Ana Margarida Almeida^a

^a*CIC.Digital/DigiMedia - Department of Communication and Arts, University of Aveiro, 3810-193 Aveiro, Portugal*

Abstract

Audio description (AD) is a media feature especially designed for visually impaired people to enable the access, via audio, to the description of visual content, such as television programs or movies, in which a narrator orally describes the visual scenes that are not perceptible by these users. Without this feature, unfortunately not available on a large scale, visually impaired users (VIU) have major difficulties in taking the best benefit of television contents.

Currently, the production of AD is a costly and complex process, achieving only the minimum requirements mandated by what TV law obliges. However, the current technological convergence paradigm provides great opportunities to enhance this process, leveraging the number of AD and bringing relevant socio-economic impacts.

In this context, this paper aims to describe some key steps to support the participatory development of a cross-platform prototype, which enable collaborative and voluntary creation of AD in an inclusive approach. A special focus is given to a preliminary test aiming to understand how the target audience responds to the designed collaborative strategy. Based on a qualitative evaluation, we found that volunteers were satisfied with the procedure used to create AD and visually impaired users felt that this AD assisted them.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the organizing committee of CENTERIS 2016

Keywords: Collaborative audiodescription; Interactive television; Accessibility; Inclusive design; Visual impairment.

* Corresponding author. Tel.: +351-234-370-389; fax: +351-234-370-868.

E-mail address: ritaoliveira@ua.pt

1. Introduction

Audio description is prone to be applied in various spaces and contexts, being only necessary to adapt the methods and technologies to the characteristics and specificities of the corresponding situation [1]. For this reason, there are several forms of audio description. The presented research proposes a new approach for the creation of TV audio descriptions supported in the recording of real time audio descriptions (as it happens in performing arts with AD, e.g. theater and dance) rather than in the common process where the AD is pre-prepared. This new approach relies, therefore, on volunteers (e.g. VIU's relatives that usually help them during television viewing) who do not necessarily have trained skills. It is worth to say that this approach is especially relevant for non-linear TV programs, for instance those available from Catch-up TV services (this time-shift service allows previously broadcasted programs to be available for later consumption - usually during 7 days or more) or video-on-demand (VOD) services [2]. The motivation behind this proposal is the poor offer of TV programs with audio description, namely in Portugal, where VIU can only access a few hours of audio described content per month. In this framework, this research aims to study new strategies to increase the number of audio described content fostering the access of visual impaired users to television programs. To achieve this, a cross-platform prototype of voluntary and collaborative creation of audio description is being specified.

In the following sections we present the state of the art related to this field of research and explain more deeply the prototype. Next, we present the first study performed in order to analyze the target audience and the research context, pointing out the methodological process, the selected findings and its discussion. The paper concludes with a summary of the lessons learned and the work to be done in the future.

2. State of the art

In Portugal, Neves [1] is a precursor of several projects that encourage the use of audio description in various domains. Oliveira et al. [3] has some work in the field of interactive television, proposing an adapted system to VIU and with support for audio description. In Europe there are also audio description incentive projects, such as the European project 'ADLab' [4] that involves countries like Portugal, Italy, Spain, Belgium, Poland and Germany. Also in Europe the 'HBB4ALL' [5] is an important project that aims to promote a wider availability of accessible media aids to all viewers with special needs (like audio description, subtitles and sign language). The project aims to make a cross-platform production and distribution of accessibility features more cost-efficient and more flexible and easier to use, benefiting users and also broadcasters and media producers. In Spain, the 'AudescMobile' application was created to allow the access to AD of several types of audiovisual productions, using audio fingerprinting (analysis of a sound excerpt recorded by the application) [6]. Also in Spain, Szarkowska and Orero [7] addressed a problem related to audio descriptions: the absence of standards for loudness or AD mix. In Brazil there are projects related to the promotion of audio description; for example Campos [8] suggests a mechanism based on semantic web application for the automatic creation of audio description for movies and the study from Bizelli and Machado [9] that shows the benefits of using audio description in digital television. In the United States of America a free and experimental tool was developed as part of an academic research allowing to add and access audio descriptions, on YouTube videos, through a web tool [10]. As audio description is financially expensive, the design of a volunteer model to create it becomes clearly advantageous. This is aligned with several European initiatives that take advantage of volunteer models for creation of audiobooks [11] [12] [13], audio description book illustrations [13] and audio newspapers [14].

3. Prototype

The concept for the collaborative and inclusive approach that will support the development of the cross-platform solution for iTV will be based on the following steps: i) recording of the audio description to be carried out by the volunteers synchronously with the television content; ii) easier edition/elimination of unsuitable content; iii) uploading the final result into an audio descriptions repository; iv) assembly of several contributions to create a more complete audio description; v) evaluation of the final audio description by the community; and vi) provision of the audio description through an iTV application (interconnected with a catch-up TV program or other non-linear

source that makes the respective television program available). It is also foreseen to provide the possibility of selecting the audio description by various criteria (e.g. name of the author or group of authors). Figure 1 depicts a usage scenario of the prototype.

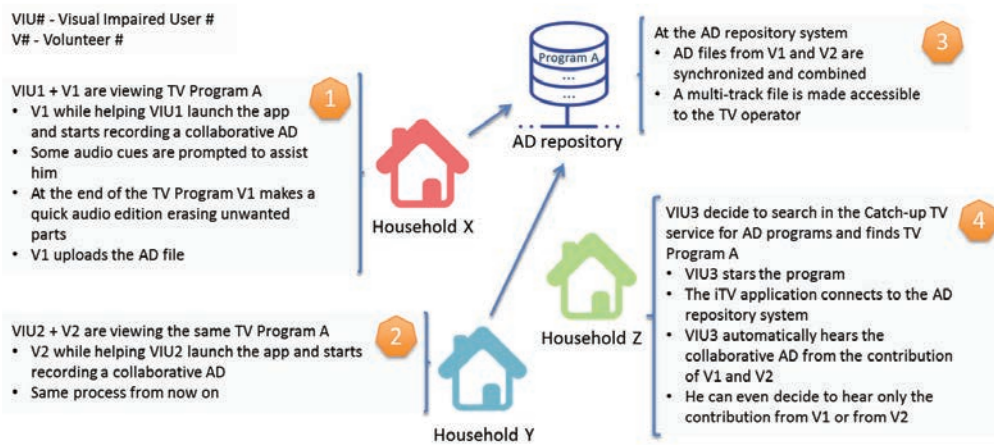


Figure 1 – A usage scenario of the prototype

With the aim to achieve the mentioned goals, it was essential to present both volunteers and VIU with a preliminary version of the prototype dynamic (cf. next section). This participatory design approach is of paramount importance to acquire, at the early stages of the conceptualization and design of our system, important hints to improve the functional requirements and to settle the basis for the technical implementation of the system over a commercial pay-TV infrastructure.

4. Methods and Results

The first phase of this research aims to study the end users, including the receptivity and the opinion of both volunteers (to create the audio descriptions) and potential users (with visual impairment). A preliminary test was conducted, followed by exploratory interviews applied to these two user profiles to be later completed with other data collection moments. Regarding volunteers, these interviews allowed us to understand their opinions about the proposed strategies to be used when recording the audio descriptions (whether through voice recording or using drafting text files played from a synthesized speech engine). In what concerns VIU, it is important to get insights if they prefer an individual or collaborative approach and what are the difficulties that these users experience when accessing current audio description services, as well which strategies they use to solve those problems. In the next subsections these steps are described allowing a better comprehension of the opinion of volunteers and VIU about the creation and utilization of real time and spontaneous audio descriptions. It is worth to say that this sample is very limited in order to obtain generalizable results, so this preliminary study only provides a proof of concept.

4.1. Methodological Process

The preliminary test was supported by a methodological approach organized in four stages and involving two groups of people and a specifically selected TV program to be audio described. A simulated environment was created to allow mapping the concept of the future prototype. Two households (A and B) were identified according two selection criteria: having at least a visually impaired TV user (who usually benefits from spontaneous audio descriptions made by a relative, e.g. a son, brother or a grandchild) and, cumulatively, having this relative as a volunteer to create an AD. The TV program selected was two episodes (not seen by the users) of a popular soap opera broadcasted by a Digital Terrestrial Television (DTT) channel. The test involved the following stages:

- 1st stage – volunteers were asked to record an audio description using a video excerpt (with 5 minutes) of one of the episodes (it was available through a proprietary VoD[†] service offered by a commercial IPTV provider). In order to allow them to become familiar with the procedures, they were informed that they could record the audio description with headphones and in an individual ambience. The video excerpt was the same in the two households[‡];
- 2nd stage – volunteers were asked to record an audio description of a video excerpt with (5 minutes) of another episode of the same soap opera (it was available through the same VoD service) accompanied with the visual impaired user allowing him/her to be part of the process. The video excerpt in the two households was different[§]. Audio descriptions recorded by volunteers at this stage were post-mixed with the video excerpts of the selected TV content and provided in the IPTV VoD service for the next stage^{**};
- 3rd stage – the video with the AD performed at household A (in the 2nd stage) were made available to the VIU of household B and vice-versa;
- 4th stage – all participants of the two households were interviewed.

4.1.1. Participants' characterization

The participants in our study (two volunteers and two VIU) were all woman. One of the volunteers is a student and the other has a professional activity, having both less than 40 years old. Both of the VIU are retired and have more than 65 years old; in both cases their visual impairment was acquired and its level is high: one is blind and the other has severe low vision. All participants are clients of the same IPTV provider.

4.1.2. Interview protocols

A semi-structured interview protocol was used to gather participants' opinions about the creation of real time and spontaneous audio descriptions. This type of interview was chosen in order to facilitate data processing and to allow some flexibility to the interviewees. Two interview protocols were prepared: one for the participants (volunteers) who recorded AD and the other for the participants with visual impairment (VIU) who experienced the audio described videos.

The volunteers' interview protocol was divided into four parts, integrating overall thirteen questions: Part I – General Data: participants were asked to say their age and occupation; Part II – Television Consumption Patterns: participants were asked to answer to some questions related to their TV content consumption patterns (e.g.: how much time they spent watching TV in the week); Part III – AD Creation Experience: in this part, participants were asked to report the difficulties they had in the AD recording and to suggest modifications related to the experienced procedure; Part IV – AD Creation Platform: functional questions related to the cross-platform that will be developed were asked to participants (e.g.: if they preferred to use an artificial speech generation tool instead of their own voice).

The visually impaired users' interview protocol was divided into five parts, having eighteen questions: Part I – General Data: participants were asked to say their age, occupation and also their visual impairment type (low vision or blindness) and how long they have this disability; Part II – Television Consumption Patterns: participants were asked about the time they spent watching television and also about the use of audio description (e.g.: if they watch TV programs with audio description); Part III – Difficulties in viewing TV content: participants were asked about their difficulties in viewing TV content and how they believe these problems could be solved; Part IV – AD Use Experience: participants were asked to report the difficulties they had in understand the performed AD and to compare professional and voluntary AD approaches; Part V – AD Creation Platform: functional questions related to the cross-platform that will be developed were asked to participants (e.g.: if they preferred to listen AD created through an artificial speech generation tool instead of a natural voice).

[†] This is a special VoD service that allows IPTV clients upload their own videos and make them available to others

[‡] The original video can be accessed here: <https://6q0qjk.s.cld.pt>; and correspondent audio descriptions here: <https://pydz6k.s.cld.pt>

[§] The original videos can be accessed here: <https://49gl2v.s.cld.pt>; and correspondent audio descriptions here: <https://u26no8.s.cld.pt>

^{**} The videos can be accessed here: <https://3xv6w4.s.cld.pt>

4.2. Selected Findings

The analysis of the interviews revealed interesting findings, which are following presented in this section.

Concerning the volunteers' interview, participants spend about 1 hour per day watching television in the week and more than 2 hours in the weekend. The television genres they most like are information, soap operas, series and movies. They frequently help their relatives with visual impairment in viewing television programs: one of the participants stated that helps her relative more than 3 times per week and the other said it depends of the requests, but normally she helps her several times per week. Concerning the AD narration experience, participants felt no difficulty in its creation/recording and suggested no modifications in the procedure. Both participants indicated their willingness to use, in the future, a platform that enables the voluntary and collaborative creation of audio descriptions. The two participants thought that a tool that sonorously helps them in the insertion and conclusion of the AD excerpts (for example through short audio prompts) was useful, as it would avoid creating audio descriptions superimposed to the characters' dialogues. One of the participants also said that the audio prompts would have to appear in advance (about 3 to 5 seconds for the person not to be anxious or hasty). Both participants said that they would not like to use an artificial speech generation tool instead of their own voice, because they consider that a natural voice is always more noticeable. Further, one of the participants stated that, despite being easier to use the artificial voice, she thinks the users of the audio description will benefit more from a natural voice.

Regarding the visually impaired persons interview, both participants spend between 2 and 3 hours per day watching television in the week. In the weekend the number of hours increases to more than 3 hours watching television. The television genres they like more are information and soap operas. The participants said that, sometimes, they need that someone in their family helps them clarifying what is happening in the program they are watching. Curiously, both stated that are helped less often than their relatives reported. Both participants do not often watch programs with audio description: one of them stated that watched two or three times and the other stated watched only one time. Nevertheless, they would like to have access to more television programs with audio description and assigned a high degree of usefulness to it. The main problems the participants feel when watching TV are: i) not being able to read subtitles; and ii) do not understand certain things that happen in TV programs. To address these problems, they both think that spoken subtitles and audio description are good solutions. Regarding the voluntary AD they experienced, no participant felt difficulties in understanding the heard AD excerpt. Comparing the professional AD participants known and the AD they heard, the voluntary AD seems more attractive to them. One of the participants justified her answer saying that she had a feeling that it was more detailed. The other participant considered the voluntary AD more attractive, because she did not have a good experience when she had contact with the professional AD. After hearing an example of artificial voice, both participants mentioned not to want hear an AD created through a synthesized voice instead of a person's voice, because they think this last one is more natural and pleasant. However, one of the participants said that the artificial voice also seems a viable option when it is not possible to use the natural one. Finally, they considered that a TV service in which they could choose and activate voluntary AD for various programs would be useful, because it would assist them to understand what they do not notice and/or understand in programs.

4.3. Discussion

During the preliminary test, family members of visually impaired persons have demonstrated a great predisposition to help persons with this type of impairment while watching TV. They already are used to help their relatives during TV watching and stated they will use their own voice to create a more natural and perceptible AD in a future service. Visually impaired people spend a significant amount of time watching television, and soap operas are one of the television genres that they watch more. This fact is important under the scope of our study because this type of programs is most commonly adapted to audio description. We believe that the few reported situations in which interviewed VIU said that they don't use this type of support from their relatives are related to the strong desire of autonomy that characterizes persons with this type of impairment. We also consider that audio description may be still less used by the participants because they have a TV commercial solution instead of the Portuguese DTT service (the only one that is prepared to the audio descriptions emitted by the national public channel) and also because, as it was said before, there is a poor dissemination of this solution. The identification of audio description

as one of the solutions to the problems when participants watch TV, the high degree of usefulness assign to this feature and the preference of the voluntary audio description instead of professional AD are good signs that encourage and support the future use of a service that will offer voluntary AD. This predisposition to use the service is also confirmed with the answers to the last question of our interviews, in which participants stated that this future service would be helpful because it could help them realize what they do not understand in the programs.

4. Conclusions and Future Work

The creation of new strategies and services supported by universal access principles is an important contribution to increase digital literacy and to promote the inclusion of people with special needs. Over the last years, several research works [3] [4] [5] were focused on the study of strategies and practices for enhancing the access and use of audio description, considering the needs of VIU. Our study underlines these goals as it intends to study new strategies for the inclusion of VIU, considering specifically the challenges presented by the Television paradigm, towards the development of a cross-platform prototype for creating collaborative audio descriptions that can be accessed by VIU through an iTV application. The first study conducted, described in this paper, allowed us to verify that people who volunteered to create audio descriptions were satisfied with the tested procedure and willing to future use a platform that enables the voluntary and collaborative creation of AD. The VIU who benefited from voluntary audio descriptions also agreed with the concept of the service to be developed, and considered that a TV service in which they could access this type of AD would be useful. The presented results are of most relevance to enable the development of the forthcoming steps of this research in which a more detailed and extensive study will be conducted in order to include a largest and more diverse group of participants. Later then, with the obtained results, we will conceptualize all the requirements of the cross-platform of AD, considering several challenges such as creating a promotion mechanism (for example through a ranking of the best AD narrators) and implementing basic guidelines/best practices for the proper implementation of suitable audio descriptions.

References

1. Neves, J. *Imagens que se Ouvem – Guia de Audiodescrição*, Instituto Politécnico de Leiria: Leiria; 2011.
2. Abreu J, Nogueira J, Becker V, Cardoso B. Survey of Catch-up TV and other time-shift services: a comprehensive analysis and taxonomy of linear and nonlinear television. *Telecommunication Systems* 2016; **62**(1):1-18. doi:10.1007/s11235-016-0157-3.
3. Oliveira R, Abreu J, Almeida M. Promoting IPTV Accessibility for Visually Impaired Users: Implementation of an Adapted Service. *Procedia Computer Science* 2014; **27**: 113-122. doi: 10.1016/j.procs.2014.02.014.
4. ADLAB Project. Audio Description: Lifelong Access for the Blind. ADLAB Project Website 2014. Available from: <http://www.adlabproject.eu/>.
5. HBB4ALL. Hybrid Broadcast Broadband For All - Connected TV Accessibility. HBB4All Project Website 2016. Available from: <http://www.hbb4all.eu/>.
6. Organización Nacional de Ciegos Españoles (ONCE). *Audesc Mobile: la audiodescripción en nuestra mano*. Centro de Investigación, Desarrollo y Aplicación Tiflotécnica 2014. Available from: <http://cidat.once.es/home.cfm?id=1516&nivel=2>.
7. Szarkowska A, Orero P. The importance of sound for audio description. In: Maszerowska A, Matamala A, Orero P, editors. *Audio Description: New perspectives illustrated*. Amsterdam: John Benjamins; 2014. p. 121-140.
8. Campos V, Araújo T, Filho G. *CineAD: Um Sistema de Geração Automática de Roteiros de Audiodescrição*. Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia) 2014 - IX Workshop de Teses e Dissertações (WTD). Available from: <http://www.lbd.dcc.ufmg.br/colecoes/wtdwebmedia/2014/001.pdf>.
9. Bizelli J, Machado F. Audiodescrição na televisão digital e perspectivas de suas contribuições para o acesso à informação por pessoas com deficiência visual. In: Martins, S. et al, editors. *Tecnologías de la información y comunicación, TIC, en educación especial*. Madrid: Universidad de Alcalá de Henares; 2014. p. 77-97.
10. YouDescribe. A free accessibility tool for adding audio description to YouTube videos. YouDescribe Website 2015. Available from: <http://youdescribe.org/>.
11. Biblioteca Nacional de Portugal (BNP). *Leitura para Deficientes Visuais – Recursos áudio*. BNP Website 2016. Available from: http://www.bnportugal.pt/index.php?option=com_content&view=article&id=121&Itemid=155&lang=pt.
12. Librivox. *Acoustical liberation of books in the public domain*. Librivox Website 2016. Available from: <https://librivox.org/>.
13. Projeto de Leitura Inclusiva Partilhada (PLIP). KITS PLIP. PLIP - Instituto Politécnico de Leiria Website 2016. Available from: <http://plip.ipleiria.pt/dinamizacao>.
14. National Talking Newspapers and Magazines (NTNM). *About Our Service*, NTNM. Royal National Institute of Blind People (RNIB) Website 2016. Available from: <http://www.tnauk.org.uk/index.html>.